

ISSN 2663-3051 (Print)  
ISSN 2663-306X (Online)  
(ISSN 0555-2656 до 2019 р.)

**БІОНІКА ІНТЕЛЕКТУ**  
**ІНФОРМАЦІЯ, МОВА, ІНТЕЛЕКТ**

**BIONICS OF INTELLIGENCE**  
**INFORMATION, LANGUAGE, INTELLIGENCE**

---

---

**№ 1 (104)**

**2026**

---

---

**НАУКОВО-ТЕХНІЧНИЙ ЖУРНАЛ**  
**SCIENTIFIC AND TECHNICAL JOURNAL**

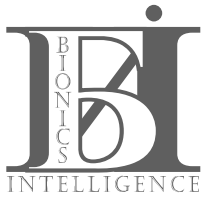
**Заснований у жовтні 1967 р.**  
**Founded in October 1967**

Засновник та видавець  
*Харківський національний університет радіоелектроніки*

Founder and Publisher  
*Kharkiv National University of Radio Electronics*

Періодичність видання – 2 рази на рік  
Publication frequency – 2 issues per year

Харків • ХНУРЕ • 2026  
Kharkiv • KhNURE • 2026



Науково-технічний журнал

«БІОНІКА ІНТЕЛЕКТУ»

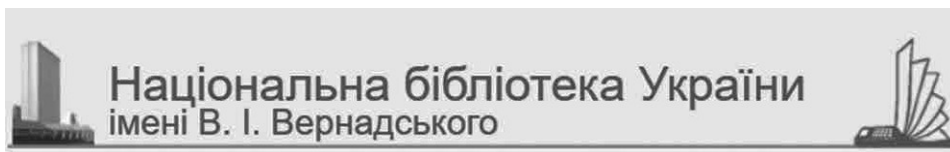
Заснований Харківським національним університетом  
радіоелектроніки у 1967 році

Реферування та індексування:

# Google Scholar



INDEX  COPERNICUS  
I N T E R N A T I O N A L



Журнал включено до Переліку наукових фахових видань України,  
в яких можуть публікуватися результати дисертаційних робіт на здобуття наукових ступенів  
доктора наук, кандидата наук та ступеня доктора філософії,  
згідно з наказом Міністерства освіти і науки України № 820 від 11.07.2016  
(внесено зміни згідно з наказом МОНУ № 920 від 26.06.2024)

© Харківський національний університет радіоелектроніки, 2026





А. І. Воржевітіна<sup>1</sup>, В. Д. Шанідзе<sup>2</sup>

<sup>1</sup>НТУ «ХПІ», м. Харків, Україна, Anzhelika.Vorzhevitina@sgt.khpi.edu.ua,  
ORCID iD: 0009-0001-0562-0191

<sup>2</sup>НТУ «ХПІ», м. Харків, Україна, vladyslavashanidze@gmail.com,  
ORCID iD: 0009-0002-9809-5572

## ФОРМУВАННЯ СИСТЕМИ ОЗНАК ДЛЯ МОДЕЛЮВАННЯ ЛІНГВІСТИЧНИХ СТРУКТУР НА ОСНОВІ МЕТОДУ КОМПАРАЦІЇ ЗА ЗАСОБІВ ТЕОРІЇ ІНТЕЛЕКТУ

У статті розглядається проблема формалізації лінгвістичних структур в рамках використання основного методу теорії інтелекту — методу компараторної ідентифікації та математичного інструментарію — апарату алгебри скінченних предикатів (АСП). Автори обґрунтовують перехід від імовірнісних методів аналізу тексту до детермінованих логіко-математичних моделей, що забезпечує високий ступінь інтерпретованості результатів. Запропоновано метод ідентифікації деструктивних і маніпулятивних лінгвістичних структур на основі компараторного підходу, що дозволяє системі виконувати абдуктивний висновок і відновлювати пропущені ознаки (лакуни) в умовах невизначеності даних. У роботі детально описано алфавіт предикатних змінних (інтенція, суб'єктні відносини, лексичні маркери) і наведено результати верифікації моделі на контрольних вибірках текстів кримінальної та маніпулятивної спрямованості. Дослідження робить внесок у вирішення проблеми «заземлення символів» (Symbol Grounding) і створення прозорих когнітивних архітектур для систем інформаційної безпеки.

ЛІНГВІСТИЧНІ СТРУКТУРИ, АЛГЕБРА СКІНЧЕННИХ ПРЕДИКАТИВ, ТЕОРІЯ ІНТЕЛЕКТУ, КОМПАРАТОРНА ІДЕНТИФІКАЦІЯ, АБДУКТИВНИЙ ВИСНОВОК, ІНФОРМАЦІЙНА БЕЗПЕКА, ІНТЕРПРЕТОВАНИЙ ШІ, РЕАВТОМАТИЗАЦІЯ МОВИ

**A. I. Vorzhevitina, V. D. Shanidze. Formation of a system of features for modelling linguistic structures based on the method of comparison using the tools of the theory of intelligence.** This article examines the problem of formalizing linguistic structures within the framework of the main method of intelligence theory—the method of comparative identification and mathematical tools—the apparatus of finite predicate algebra (FPA). The authors justify the transition from probabilistic methods of text analysis to deterministic logical-mathematical models, which ensures a high degree of interpretability of the results. A method for identifying destructive and manipulative linguistic structures based on a comparator approach is proposed, enabling the system to perform abductive inference and restore missing features (lacunae) in the face of data uncertainty. The paper describes in detail the alphabet of predicate variables (intent, subject relations, lexical markers) and presents the results of model verification on control samples of criminal and manipulative texts. The study contributes to solving the problem of symbol grounding and creating transparent cognitive architectures for information security systems.

LINGUISTIC STRUCTURES, FINITE PREDICATE ALGEBRA, INTELLIGENCE THEORY, COMPARATOR IDENTIFICATION, ABDUCTIVE INFERENCE, INFORMATION SECURITY, INTERPRETED AI, LANGUAGE RE-AUTOMATION.

### Вступ

Сучасний етап розвитку міждисциплінарних досліджень характеризується інтеграцією когнітивних наук і лінгвістики, що зумовлює необхідність переосмислення фундаментальних механізмів обробки вербальної інформації. Вважається, що алгебра скінченних предикатів (АСП) — це універсальний математичний апарат для формального опису детермінованих, дискретних, скінченних об'єктів [1–5]. У роботах [1–3] доведено, що для того, щоб мати можливість розвивати теорію інтелекту, перш за все необхідно мати формальну мову, якою можна було б математично описувати структури і функції людського інтелекту, особливо ті, що стосуються накопичення та інтерпретації знань людиною. У якості такої мови у науковій школі професора Ю.П. Шабанова-Кушнаренка та його послідовників використовується АСП. У рамках теорії інтелекту мова розглядається не просто як інструмент комунікації, а як складна

когнітивна система, що оперує специфічними конфігураціями даних. Ключовим об'єктом аналізу в даному контексті виступають лінгвістичні структури (linguistic structures) — організовані елементи і зв'язки всередині мови, що забезпечують зберігання, передачу і генерацію смислів [6–8].

Варто зауважити, що сучасні дослідження багатомовності дедалі частіше зосереджуються на механізмах автоматизації та контролю мовленнєвої діяльності. У цьому контексті особливий інтерес становлять зовнішні прояви когнітивних процесів, одним із яких є почерк — складна психомоторна діяльність людини, що згодом переходить у високоавтоматизований режим. Це дозволяє розглядати почерк як детермінований слід автоматизованих лінгвістичних і моторних програм, чутливий до змін мовної домінанти та процесів її де- і реавтоматизації.

Питання про правомірність і межі використання терміну «лінгвістичні структури» залишається фундаментальним для сучасної науки. Незважаючи

на високий ступінь абстракції, дане поняття є загальноновизнаним науковим конструктом, що виступає базисом для таких шкіл, як структуралізм, генеративна лінгвістика Н. Хомського і когнітивізм [9]. Лінгвістичні структури охоплюють багаторівневу організацію мови: від фонетичних і морфологічних одиниць до складних синтаксичних схем, дискурсивних моделей і лінгвокультурних констант. У прикладному аспекті, особливо в завданнях моделювання інтелектуальних систем, заснованих на базах знань, та комп'ютерній лінгвістики, цей термін використовується для опису формальних алгоритмів та ієрархічних зв'язків (наприклад, синтаксичних дерев), що відображають внутрішню логіку мовної здатності людини [6–8, 10].

Для глибокого розуміння процесів обробки інформації необхідно розмежувати суміжні, але не тотожні поняття: «мовну систему», «структуру мови» і «лінгвістичні структури». В академічному дискурсі ці терміни утворюють сувору ієрархічну залежність:

1. Мовна система (language system) є найширшим поняттям — цілісним комплексом усіх підсистем і правил їх взаємодії, що має динаміку та здатність до розвитку [6–8, 11].

2. Структура мови (language structure) виступає як архітектурний каркас цієї системи, що визначає спосіб внутрішньої організації її елементів [6–8, 11].

3. Лінгвістичні структури (linguistic structures) є конкретними «будівельними блоками» і стійкими конфігураціями (схеми слів, речень, когнітивні схеми), які реалізують загальну структуру всередині живого функціонування системи [6–8, 11].

Таким чином, дослідження обробки лінгвістичних структур на основі теорії інтелекту дозволяє не тільки описувати формальну сторону мови, але й виявляти механізми того, як абстрактні ментальні категорії трансформуються в конкретні мовні одиниці. У даній статті запропоновано підхід до формалізації цих процесів за допомогою апарату алгебри скінченних предикатів, що дозволяє побудувати детерміновану модель ідентифікації смислових лінгвістичних структур. Основна увага приділяється розробці та верифікації алгоритму компараторної ідентифікації, який забезпечує розпізнавання деструктивних та маніпулятивних мовних стратегій навіть в умовах неповних або завуальованих даних. Це відкриває нові можливості для створення інтелектуальних систем із високим рівнем інтерпретованості (Explainable AI), здатних до глибокого аналізу когнітивних намірів автора тексту.

### 1. Огляд існуючих рішень з обробки лінгвістичних структур на основі теорії інтелекту

Сучасний стан області обробки лінгвістичних структур характеризується переходом від суто емпіричних методів до спроб інтеграції глибокого

навчання з когнітивними та формально-логічними моделями інтелекту. Незважаючи на відсутність єдиної «універсальної теорії мовного інтелекту», аналіз існуючих рішень дозволяє виділити кілька ключових напрямків, що визначають поточний ландшафт досліджень.

Найбільш помітним досягненням останніх років стало використання великих мовних моделей (LLM), таких як GPT-4, Claude і LLaMA. Дослідження показують, що ці моделі здатні виконувати семантично складну бінарну класифікацію (наприклад, виявлення порушень прав людини в соціальних мережах за аналізом текстових повідомлень) з якістю, порівняною з людським експертним рівнем [12].

З точки зору оцінювання роботи нейромереж (LLM) не просто як технічної програми чи калькулятора, а як когнітивної системи, яка намагається відтворити інтелектуальні функції людини (такі як сприйняття, усвідомлення, розуміння текстових повідомлень, мовної ієрархії, логіку побудови речень, виявлення зв'язків, тощо), LLM демонструють феномен «емерджентного синтаксису» — здатність спонтанно засвоювати ієрархічні структури мови в процесі навчання на передбачення наступного токена. Однак архітектура цих рішень позбавлена прозорості та прямої інтерпретованості.

На противагу суто статистичним методам, когнітивно-семіотичні та гібридні підходи базуються на логічних і смислових принципах людського мислення, інтегруючи когнітивно-лінгвістичні алгоритми імітації механізмів засвоєння граматичних структур із семіотичним моделюванням мов зі складною морфологією, де рекурсія та еліпсис виступають фундаментом для інтелектуальних технологій представлення знань. Особливе місце в цьому ряді посідають логіко-алгебраїчні методи на основі алгебри скінченних предикатів і теорії категорій [13–15], що дозволяють формалізувати вилучення фактів і будувати прозорі детерміновані системи, здатні до логічного висновку в умовах невизначеності. Такий синтез підходів дозволяє подолати глибокий розрив між високою ефективністю нейромережових моделей у відтворенні синтаксичної ієрархії та їхньою неспроможністю забезпечити прозору інтерпретацію когнітивних намірів автора. Отже, критичний дефіцит інтерпретованості (ефект «чорної скриньки») та відсутність механізмів обґрунтованого висновку в умовах неповних даних роблять актуальним пошук альтернативних методів моделювання мовних знань, що поєднують математичну строгість із когнітивною глибиною.

### 2. Постановка завдань і цілей дослідження

Незважаючи на істотний прогрес у галузі практичної обробки природної мови, відсутність єдиної теорії інтелектуальної обробки лінгвістичних структур

породжує низку критичних розривів між когнітивними здібностями людини та алгоритмічними рішеннями. Однією з центральних проблем залишається зведення воедино розрізнених рівнів організації мови, від морфології та синтаксису до дискурсу та прагматики, в рамках єдиної когнітивної архітектури [16]. Існуюча невизначеність у питаннях формального зіставлення людського лінгвістичного інтелекту та властивостей сучасних нейромережових моделей вимагає пошуку нових, інтерпретованих підходів до аналізу внутрішньої структури знань. Проблеми обґрунтованості символів і семантики, а також інтеграція фактичних, концептуальних і метакогнітивних рівнів знань залишаються відкритими, заважаючи створенню прозорих механізмів вилучення правил зі статистичних даних.

Метою даного дослідження є розробка сучасної моделі обробки лінгвістичних структур, яка базується на використанні методу компараторної ідентифікації теорії інтелекту та засобів алгебри скінченних предикатів для забезпечення глибокої формалізації переходу від низькорівневих ознак тексту до високорівневих когнітивних схем [17].

В рамках досягнення поставленої мети планується вирішити завдання побудови моделі обробки лінгвістичних структур у вигляді системи рівнянь алгебри скінченних предикатів, що описують класи еквівалентності текстових об'єктів на основі їх семантико-синтаксичних характеристик [15]. Особлива увага приділяється підвищенню ефективності аналізу текстів у специфічних доменах, таких як кримінальна лінгвістика, за рахунок врахування тонких структурних особливостей і механізмів операціоналізації знань [18]. Дослідження спрямоване на подолання проблеми опису слабоформалізованої інформації, шляхом впровадження процедур компараторної ідентифікації, що дозволяють відновлювати відсутні атрибути лінгвістичних об'єктів і забезпечувати обґрунтованість і пояснюваність рішень, що приймаються інтелектуальною системою в умовах невизначеності та суб'єктивності вихідних даних [17].

### 3. Формалізація лінгвістичних ознак у базисі алгебри скінченних предикатів

Для переходу від неструктурованого текстового масиву до формальної інтелектуальної моделі ми виділяємо скінченну множину об'єктів  $U$  (текстових фрагментів) і шляхом формування системи ознак та їх значень визначаємо класи еквівалентності  $K$  на  $U$ . В рамках апарату алгебри скінченних предикатів (АСП) кожен лінгвістичний об'єкт характеризується вектором ознак:

$$X = (x_1, x_2, x_3, \dots, x_n), \quad (1)$$

де  $n$  — кількість ознак, значення яких вибираються зі скінченних алфавітів.

Спираючись на методологію класифікації кримінальних текстів [18] і принципи компараторного аналізу [17], ми виділяємо три домінуючі атрибути, що формують «інтелектуальний портрет» повідомлення. Методологічною основою такого підходу є компараторний принцип, що передбачає зіставлення систем ознак. Цікавою в рамках цього підходу є також задача дослідження почерку людини при переході від однієї мови до іншої (рідної чи іноземної). Почерк у такому підході описується як система формалізованих параметрів (розмірні, просторові, варіативні ознаки), які утворюють структурно впорядкований опис письмової мови. Порівняння таких систем дозволяє виявляти ступінь подібності або розбіжності автоматизованих програм без суб'єктивної інтерпретації. Аналогічно, у даній роботі лінгвістичні ознаки (варіативність синтаксису, щільність маркерів) розглядаються як індикатори когнітивного навантаження: зростання нестабільності параметрів інтерпретується як ознака зниження автоматизації або активної перебудови відповідних навичок.

#### 3.1. Визначення алфавіту ознак і предметної області

Система ознак будується на основі якісних характеристик тексту, які перетворюються в дискретні значення  $a_i^j$ . Тут перший (нижній) індекс  $i$  вказує на номер змінної, а другий (верхній) індекс  $j$  — на конкретне значення даної змінної. Така індексація дозволяє створити унікальну координатну адресу для кожного лінгвістичного значення в базі знань системи.

Атрибутивна змінна  $x_1$  — тип комунікативної інтенції. Даний параметр визначає прагматичну мету повідомлення. Множина значень:

$$M_1 = \{a_1^1, a_1^2, a_1^3\}, \quad (2)$$

де  $a_1^1$  (дескриптивний) — нейтральне оповідання про факти або звіт про події;  $a_1^2$  (індуктивний) — маніпулятивне спонукування до дії, характерне для вербування або шахрайських схем;  $a_1^3$  (агресивний) — виражена психологічна атака, пряма погроза або спроба придушення волі адресата.

Атрибутивна змінна  $x_2$  — структура суб'єктно-об'єктних відношень. Описує характер взаємодії між учасниками комунікації. Множина значень:

$$M_2 = \{a_2^1, a_2^2, a_2^3\}, \quad (3)$$

де  $a_2^1$  (персоніфікований) — жорстка зв'язка «Я — Ти», що вказує на адресність впливу;  $a_2^2$  (інституційний) — звернення від імені групи, організації або анонімної спільноти;  $a_2^3$  (невизначений) — повна анонімізація або розмитість суб'єкта дії.

Атрибутивна змінна  $x_3$  — Домінуючий лексичний маркер. Фіксує ключовий об'єкт інтересу в тексті.

Множина значень:

$$M_3 = \{a_3^1, a_3^2, a_3^3\}, \quad (4)$$

де  $a_3^1$  (ресурсний) – фокус на фінансах, персональних даних, криптоактивах або паролях;  $a_3^2$  (ідеологічний) – використання термінів боротьби, ворожнечі, помсти або політичних гасел;  $a_3^3$  (побутовий) – використання загальноновживаної лексики без явних деструктивних маркерів.

Формальна коректність моделі забезпечується дотриманням закону несумісності знань у межах кожного алфавіту. Оскільки алфавіти ознак  $A_1$  (інтенція),  $A_2$  (відношення) та  $A_3$  (маркери) побудовані за принципом взаємовиключних категорій, виконується умова  $x_i^{a_i^j} \wedge x_i^{a_i^k} = \emptyset$ , що унеможливорює неоднозначність при ідентифікації. Наприклад, інтенція не може бути одночасно суто дескриптивний ( $a_1^1$ ) та агресивний ( $a_1^3$ ), оскільки наявність маркерів агресії автоматично змінює статус об'єкта в ієрархії класифікації

### 3.2. Побудова предикатної моделі та приклад операціоналізації

Згідно з теорією [14], підсумкова лінгвістична структура тексту формалізується предикатом  $P(x_1, x_2, x_3, \dots, x_n)$ . Розглянемо випадок  $n = 3$ , цільова змінна  $y$  визначає належність об'єкта до одного з трьох класів: кримінально-деструктивного ( $y_1$ ), інформаційно-маніпулятивного ( $y_2$ ) або нейтрального ( $y_3$ ), кожен клас відповідає певному рівню загрози та комунікативній стратегії, закладеній у тексті. Походження цих класів базується на експертному аналізі ознак  $x_1, x_2, x_3$ .

Клас  $y_1$  – кримінально-деструктивна структура. Формується при поєднанні ознаки  $x_1$  у значенні прямої агресії ( $x_1^{a_1^1}$ ), адресності ( $x_2^{a_2^1}$ ) і корисливого інтересу до ресурсів ( $x_3^{a_3^1}$ ). Це найбільш небезпечний тип контенту (вимагання, шантаж).

Клас  $y_2$  – інформаційно-маніпулятивна структура. Походить від поєднання прихованого спонукання ( $x_1^{a_1^2}$ ), анонімності або групового впливу ( $x_2^{a_2^2}$ ) та ідеологічного навантаження ( $x_3^{a_3^2}$ ). Це контент, спрямований на зміну переконань (пропаганда, вербування).

Клас  $y_3$  – нейтрально-інформаційна структура. Заснована на описовому викладі фактів ( $x_1^{a_1^3}$ ) в рамках офіційних або побутових відносин ( $x_2^{a_2^3}$  або  $x_3^{a_3^3}$ ) з використанням загальноновживаної лексики ( $x_3^{a_3^3}$ ). Це безпечний повсякденний контент.

Приклад №1.

Розглянемо текст – «Якщо ти не надішлеш пароль від пошти протягом години, твої дані будуть опубліковані».

Аналіз інтенції – спостерігається пряма загроза («будуть опубліковані») і обмеження часу ( $x_1^{a_1^3}$  (агресивний)). Аналіз відношень – використано адресне звернення «ти» ( $x_2^{a_2^3}$  (персоніфікований)). Аналіз

маркерів – ключовими словами є «пароль» і «дані»  $x_3^{a_3^3}$  (ресурсний)).

Система класифікує даний об'єкт як кримінально-деструктивну структуру, оскільки повністю відповідає моделі шантажу:

$$P_{y_1} = x_1^{a_1^3} \wedge x_2^{a_2^3} \wedge x_3^{a_3^3}. \quad (5)$$

Приклад №2.

Розглянемо текст – «Наші прихильники вже отримують доступ до секретних архівів. Приєднуйся до боротьби за справедливість, поки не пізно!»

Аналіз інтенції – прямої загрози немає, але присутнє явне маніпулятивне спонукання і заклик до дії ( $x_1^{a_1^2}$  (індуктивний)). Аналіз відношень – автор виступає від імені групи «Наші прихильники», звертаючись до невизначеного кола осіб ( $x_2^{a_2^2}$  (інституційний)). Аналіз маркерів – текст насичений ідеологічними конструктами «боротьба», «справедливість» ( $x_3^{a_3^2}$  (ідеологічний)).

Система класифікує даний об'єкт як інформаційно-маніпулятивну структуру, оскільки не містить прямої загрози, але має ознаки когнітивного впливу:

$$P_{y_2} = x_1^{a_1^2} \wedge x_2^{a_2^2} \wedge x_3^{a_3^2}, \quad (6)$$

Приклад №3.

Розглянемо текст – «Завтра в офісі відбудеться планове засідання за результатами кварталу. Прохання підготувати звіти».

Аналіз інтенції – повідомлення має суто інформаційний характер без ознак агресії або маніпуляції ( $x_1^{a_1^1}$  (дескриптивний)). Аналіз відношень – звернення має діловий, інституційний характер ( $x_2^{a_2^1}$  (інституційний)). Аналіз маркерів – використовується стандартна ділова лексика без акценту на критичних ресурсах або ідеології ( $x_3^{a_3^1}$  (побутовий)).

Система класифікує даний об'єкт як нейтрально-інформаційну структуру. Система ігнорує його як безпечний, що не вимагає оперативного втручання:

$$P_{y_3} = x_1^{a_1^1} \wedge x_2^{a_2^1} \wedge x_3^{a_3^1}, \quad (7)$$

У разі відсутності однієї з ознак (наприклад, при розмитості суб'єкта  $x_2$ ), модель переходить до алгоритму компараторної ідентифікації для зіставлення наявних даних з еталонними кортежами бази знань.

### 4. Реалізація компараторного підходу та верифікація моделі

На відміну від традиційних методів жорсткої класифікації, компараторний підхід, що базується на теоретичних положеннях теорії інтелекту, дозволяє імітувати когнітивну здатність експерта розпізнавати прихований зміст в умовах неповноти даних. Це досягається за рахунок обчислення логічної близькості вхідного вектора до еталонних кортежів, що зберігаються в базі знань системи.

Для функціонування лінгвістичного компаратора в системі задані еталонні кортежі  $Et_{y_i}$ , які представляють собою ідеальні комбінації ознак для кожного класу (див. 3.2).

Еталон кримінально-деструктивної структури:

$$Et_{y_1} = x_1^{a_1^3} \wedge x_2^{a_2^1} \wedge x_3^{a_3^1}, \quad (8)$$

Еталон інформаційно-маніпулятивної структури:

$$Et_{y_2} = x_1^{a_1^2} \wedge x_2^{a_2^2} \wedge x_3^{a_3^2}, \quad (9)$$

Еталон нейтрально-інформаційної структури:

$$Et_{y_3} = x_1^{a_1^1} \wedge x_2^{a_2^3} \wedge x_3^{a_3^3}. \quad (10)$$

Компаратор обчислює логічну відстань між витягнутим з тексту (можливо, неповним) вектором і даними еталонами, виконуючи процедуру абдуктивного відновлення відсутніх значень.

#### 4.1. Механізм абдуктивного висновку лінгвістичного компаратора

Механізм «заповнення» пустих клітинок працює за принципом мінімізації логічної розбіжності:

1. Аналіз контекстного оточення — якщо атрибут  $x_2$  (відношень) є невизначеним, компаратор аналізує предикати  $x_1$  (інтенція) та  $x_3$  (маркер);

2. Абдуктивне припущення — у випадку, якщо інтенція визначена як «агресивна» ( $a_1^3$ ), а маркер як «ресурсний» ( $a_3^1$ ), система через механізм АСП робить логічний висновок, що найбільш імовірним (і єдино несуперечливим для класу  $y_1$ ) значенням для  $x_2$  є «персоніфікований» ( $a_2^1$ );

3. Логічна верифікація — тільки після того, як лакуна заповнена значенням, що робить предикат класу істинним  $P = 1$ , система приймає рішення про класифікацію.

Цей підхід дозволяє системі не просто ігнорувати неповні дані, а інтелектуально реконструювати прихований намір автора, що забезпечує «заземлення символів» (Symbol Grounding) у формальних логічних структурах.

#### 4.2. Верифікація моделі на розрахункових прикладах з лакунами

Нижче наведено приклади роботи лінгвістичного компаратора, де через покомпонентний аналіз тексту встановлюються значення ознак і відновлюються приховані смисли.

Приклад №1.

Об'єкт — «...слухай, якщо завтра не буде всієї суми, нарікай на себе. Я знаю, де ти живеш».

Аналіз інтенції — використання фразеологізованого виразу погрози «нарікай на себе» і створення умов психологічного тиску через згадку приватного простору «знаю, де ти живеш» вказує на ознаку «агресивний». Це маркери прямого придушення волі  $x_1^{a_1^3}$ .

Аналіз відношень — особисті займенники «я» і «ти» вказують на «персоніфікований», а імператив «слухай» свідчить про спрямований вплив конкретного суб'єкта на конкретну жертву  $x_2^{a_2^1}$ . Аналіз маркерів — ознака явно не названа (лакуна). Однак слово «сума» в семантичному полі «ультиматуму» і «термінів» однозначно вказує на фінансовий або інший матеріальний ресурс. Компаратор зіставляє загрозу і вимогу «суми», відновлюючи ознаку як (ресурсний)  $x_3^{a_3^1}$ . Результат класифікації відповідає формулі 5.

Приклад №2.

Об'єкт — «Секретні архіви вже розкриті. Приєднуйся до боротьби за справедливість, поки не пізно!»

Аналіз інтенції — дієслово в наказовому способі «приєднуйся» і маніпулятивний дедлайн «поки не пізно» — це класичні прийоми залучення до дії, що вказують на «індуктивний»  $x_1^{a_1^2}$ . Аналіз відношень — суб'єкт прихований (безлике твердження «архіви розкриті»). Лакуна відновлюється через аналіз контексту заклик: заклики до «боротьби» типові для «інституційний» впливу, де автор виступає як рупор анонімної спільно  $x_2^{a_2^2}$ . Аналіз маркерів — абстрактні ціннісні категорії «боротьба» і «справедливість», які використовуються для емоційного тригера, а не для передачі фактів, вказують на «ідеологічний»  $x_3^{a_3^2}$ . Результат класифікації відповідає формулі 6.

Приклад №3.

Об'єкт — «...зібрання в офісі, звіти за квартал потрібно підготувати до завтрашнього дня».

Аналіз інтенції — намір автора обрізаний (лакуна), але відсутність знаків оклику, погроз або закликів дозволяє компаратору припустити нейтральний «deskриптивний» на основі оточення  $x_1^{a_1^1}$ . Аналіз відношень — контекст «в офісі» і згадка «звітів», що характерно для формальної комунікації всередині організації або трудового колективу, вказують на «інституційний»  $x_2^{a_2^1}$ . Аналіз маркерів — вузькопрофесійна, але нейтральна лексика: «збори», «квартал», «підготувати», слова, позбавлені конфліктогенного потенціалу, вказують на «побутовий»  $x_3^{a_3^1}$ . Результат класифікації відповідає формулі 7.

Проведена верифікація демонструє, що лінгвістичний компаратор ефективно оперує не словами, а їх функціонально-семантичними ролями. Здатність системи відновлювати значення з лакун (наприклад, інтерпретувати «суму» як ресурсний маркер тільки в зв'язку з агресією) доводить перевагу предикатного моделювання перед простим лінгвістичним аналізом. Це дозволяє перейти від формального розпізнавання символів до інтелектуального аналізу смислових структур тексту, що знаходить своє відображення у вирішенні фундаментальних проблем штучного інтелекту.

### 4.3. Математична верифікація коректності моделі

Для підтвердження математичної коректності запропонованої моделі в базисі АСП необхідно провести перевірку дотримання фундаментальних законів теорії інтелекту: законів виключеного третього (повнота) і логічного протиріччя (несумісність). Це гарантує, що система здатна класифікувати будь-який вхідний об'єкт і при цьому виключає можливість його одночасного віднесення до взаємовиключних класів загрози.

**4.3.1. Повнота (Exhaustiveness).** Повнота моделі гарантує, що будь-який лінгвістичний об'єкт з універсуму  $U$  у межах визначених алфавітів  $M_1, M_2, M_3$  буде класифікований системою. Оскільки множини значень атрибутивних змінних  $x_1, x_2, x_3$  охоплюють весь спектр комунікативних проявів (від прямої агресії до побутового нейтралітету), то диз'юнкція всіх предикатів класів утворює тотожну одиницю, що означає, що в системі не існує «мертвих зон»: будь-який вхідний вектор або відразу відповідає еталону, або обробляється алгоритмом компараторної ідентифікації для відновлення лакуни:

$$P_{total} = P_{y_1} \vee P_{y_2} \vee P_{y_3} = 1. \quad (11)$$

**4.3.2. Виключення (Exclusivity).** Виключність гарантує відсутність суперечностей: один текст не може бути одночасно ідентифікований як два різних класи. Предикати побудовані на основі законів істинності та хибності, тобто кон'юнкція значень, що є взаємовиключними в межах одного алфавіту (наприклад,  $x_1^{a_3}$  «агресивний» та  $x_1^{a_1}$  «дескриптивний» не можуть бути істинними одночасно для одного висловлювання), забезпечує детермінованість моделі та прозорість прийнятих рішень.

$$P_{y_i} \wedge P_{y_j} = 0, i \neq j. \quad (12)$$

**4.3.3. Істинність-хибність (Validity).** Істинність підтверджується збігом результатів обчислень з експертними оцінками на контрольних вибірках. Система видає  $P = 1$  тільки тоді, коли вхідні ознаки (або відновлені лакуни) відповідають еталонній структурі класу. Будь-яка критична зміна ознаки (наприклад, заміна маніпулятивного наміру на інформаційний) миттєво змінює значення предиката на 0, що доводить чутливість та істинність моделі.

## 5. Лінгво-математичний синтез у вирішенні проблем теорії інтелекту

Запропонований метод формалізації на базі алгебри скінченних предикатів і компараторної ідентифікації дозволяє підійти до вирішення низки фундаментальних проблем, що стоять перед сучасною теорією штучного інтелекту. Перш за все, розроблений підхід пропонує ефективний шлях подолання проблеми «Symbol Grounding» (заземлення символів). На відміну від стандартних великих мовних моделей

(LLM), де токени обробляються як статистичні одиниці без прив'язки до реального значення, наше рішення забезпечує детермінований зв'язок предикатних змінних з категоріями людського досвіду. Предикат  $x_1^{a_3}$  в даній моделі перестає бути просто цифровим кодом і стає формальним описом акту агресії, що є необхідною умовою для побудови систем «сильного» ШІ, здатних до глибокого розуміння контексту.

Об'єктивним корелятом автоматизації мовленнєвих програм виступає також почерк, який може розглядатися як інструмент непрямого доступу до прихованих когнітивно-моторних станів носія мови. Запропонований підхід органічно поєднується з логіко-математичним апаратом АСП, оскільки дозволяє будувати прозорі, детерміновані моделі без втрати пояснювальної сили, що відкриває нові перспективи для дослідження динаміки мовних навичок у дорослому віці.

Важливим аспектом наукової новизни є забезпечення повної прозорості та інтерпретованості прийнятих рішень. Розроблена модель повністю виключає ефект «чорної скриньки», характерний для нейромережових архітектур. Завдяки використанню предикатних рівнянь, експерт отримує можливість верифікувати кожен логічний крок системи. Якщо об'єкт класифікується як маніпулятивний, система надає доказову базу у вигляді істинності конкретної комбінації ознак, наприклад, поєднання анонімності та ідеологічного навантаження. Така прозорість висновків має критичне значення для застосування ШІ в юридичній, слідчій та експертній практиці, де ціна помилки є надзвичайно високою.

Крім того, запропонована архітектура забезпечує безшовну інтеграцію різних рівнів знань, об'єднуючи в рамках одного математичного апарату фактографічний, концептуальний і метакогнітивний шари інформації. Використання апарату АСП дозволяє створити міст між первинною обробкою тексту на рівні лексем і високорівневим аналізом комунікативних цілей автора. Це реалізує ідеї [2] про створення єдиних інтелектуальних архітектур обробки знань, здатних до абдуктивного висновку та адаптації в умовах неповноти вихідних даних. Таким чином, теоретична значимість роботи полягає в переході від імовірнісного вгадування смислів до їх логічно обґрунтованої ідентифікації.

## Висновки

У представленому дослідженні запропоновано та обґрунтовано інноваційний підхід до обробки лінгвістичних структур на основі використання основного методу теорії інтелекту (методу компараторної ідентифікації) та математичного інструментарію алгебри скінченних предикатів. Розроблена система ознак дозволила забезпечити формалізацію якісних

характеристик тексту, перенісши суб'єктивний аналіз інтенцій та смислів у площину строгих математичних моделей. Доведено, що використання апарату АСП у поєднанні з компараторною ідентифікацією забезпечує повну пояснюваність прийнятих рішень, що є критично важливим для систем забезпечення інформаційної безпеки в специфічних доменах, таких як кримінальна лінгвістика. Впроваджений механізм відновлення лакун наділив ШІ-агента здатністю до абдуктивного висновку, дозволяючи розпізнавати приховані деструктивні структури навіть в умовах високої невизначеності або неповноти вихідних даних. Отримані результати створюють міцний теоретичний фундамент для проектування гібридних інтелектуальних систем нового покоління, де гнучкість нейромережевого навчання поєднується з жорсткою логічною інтерпретованістю експертних знань [5].

Подальший розвиток дослідження передбачає уточнення системи ознак з метою переходу від описового аналізу до абдуктивного висновку щодо процесів автоматизації та реавтоматизації мови та мовлення, враховуючи не тільки опис та обробку різноманітних лінгвістичних структур, але й формалізацію типів мовної поведінки носіїв мови при переході з однієї мови на інші, що передбачає не лише аналіз статичних структур, а й формалізацію динамічних параметрів мовної біографії суб'єкта в умовах міжмовної інтерференції.

#### Список літератури:

- [1] Бондаренко М. Ф., Шабанов-Кушнарченко Ю. П., Шаронова Н. В. Інструментарій компараторної ідентифікації. Біоніка інтелекту. 2010. №2(73).
- [2] Бондаренко М. Ф., Шабанов-Кушнарченко Ю. П., Шаронова Н. В. Ситуаційно-текстовий предикат. Біоніка інтелекту. 2010. №2(73). С. 87-98.
- [3] Бондаренко М. Ф., Шабанов-Кушнарченко Ю. П., Шаронова Н. В. Булева структура тексту. Біоніка інтелекту. 2010. №2(73). С. 99-110.
- [4] Хайрова Н. Ф., Шаронова Н. В. Інформаційно-лінгвістичні технології екстракції і ідентифікації глибинних знань в текстах: монографія. Х.: ФЛП Коряк С.Ф., 2016. 205 с.
- [5] Плехова Г. А., Шаронова Н. В. Авторський курс з дисципліни «Методи наукових досліджень» / Г. А. Плехова, Н. В. Шаронова. — Харків: ФОВ Бровін О.В. 2025.—84 с. ISBN 978-617-8238-97-1
- [6] Cholan, V., & Ponomareva, V. (2021). LINGUOCULTURAL CONSTANTS OF SLAVONIC TYPICAL TEXTS. Scientific Journal of Polonia University, 43(6), 16–22. <https://doi.org/10.23856/4302>
- [7] Oleksenko, V. P. (2021). STRUCTURAL AND SEMANTIC FEATURES OF THE NEOLEXES ON THE DESIGNATION OF THE CULTURAL AND ARTISTIC SPHERE. Opera in linguistica ukrainiana, (28), 149–168.
- [8] Bondarchuk, C. S., & Chumachenko, O. A. (2021). FEATURES OF WRITTEN LANGUAGE COMMUNICATION BY MEANS OF THE UKRAINIAN LANGUAGE (on the example of official-business, epistolary and conversational-functional styles). Alfred Nobel University Journal of Philology, 1(21).
- [9] Juanda, J. (2024). Analysis of Language Structure and Its Implications in Modern Linguistics: A Study of the Understanding and Application of Structural Linguistic Concepts. Journal of Educational and Social Research, 14(1), 226. <https://doi.org/10.36941/jesr-2024-0019>
- [10] Martin, A. E. (2020). A Compositional Neural Architecture for Language. Journal of Cognitive Neuroscience, 32(8), 1407–1427.
- [11] Boyarova, L. (2021). LINGUISTIC GLOBALIZATION: PROBLEMS OF LANGUAGE SOVEREIGNTY. In Priority areas for development of scientific research: domestic and foreign experience. Publishing House “Baltija Publishing”. <https://doi.org/10.30525/978-9934-26-049-0-23>
- [12] Smyrnova-Trybulska, E., Morze, N., & Varchenko-Trotsenko, L. (2022). Adaptive learning in university students' opinions: Cross-border research. Education and Information Technologies. <https://doi.org/10.1007/s10639-021-10830-7>
- [13] Каратаєв, О. А. (2026). Моделі та методи розроблення програмного забезпечення систем штучного інтелекту на основі інтерпретації знань: дис. ... канд. техн. Наук : галузь знань 12 : спец. 121 Інженерія програмного забезпечення. — Харків, 2026.
- [14] Дорошенко А. Ю. Інформаційна технологія інтелектуального аналізу фактографічних текстових ресурсів [Електронний ресурс] : дис. ... канд. техн. наук : спец. 05.13.06 : галузь знань 12 / Анастасія Юріївна Дорошенко ; наук. керівник Шаронова Н. В. ; Нац. техн. ун-т «Харків. політехн. ін-т». — Харків, 2018. — 177 с. — Бібліогр.: с. 149-167. — укр.
- [15] Karataiev, O., Sitnikov, D., & Sharonova, N. (2023). A Method for Investigating Links between Discrete Data Features in Knowledge Bases in the Form of Predicate Equations. In COLINS (1) (pp. 224-235).
- [16] Hahn, M., Futrell, R., Levy, R., & Gibson, E. (2022). A resource-rational model of human processing of recursive linguistic structure. Proceedings of the National Academy of Sciences, 119(43).
- [17] Cherednichenko, O., Vovk, M., Sharonova, N., & Vorzhevitina, A. (2025). Comparator-based identification of food edibility from natural language description.
- [18] Khairova, N., Kupriianov, Y., Vorzhevitina, A., & Shanidze, O. (2024). Models for effective categorization and classification of texts into specific thematic groups (using gender and criminal themes as examples). In Computational Linguistics Workshop at CoLLnS 2024. CoLLnS. <https://doi.org/10.31110/colins/2024-4/004>

Received (Надійшла) 11.01.2026

Accepted for publication (Прийнята до друку) 01.02.2026

Publication date (Дата публікації) 27.03.2026

УДК 004.89:81'322.2

DOI: [https://doi.org/10.30837/bi.2026.1\(104\).02](https://doi.org/10.30837/bi.2026.1(104).02)О. М. Юрченко<sup>1</sup>, Ю. Ю. Повесьма<sup>2</sup><sup>1</sup>НТУ «ХПІ», м. Харків, Україна, Olena.Yurchenko@kphi.edu.ua,  
ORCID iD: 0000-0002-6074-0241<sup>2</sup>НТУ «ХПІ», м. Харків, Україна, Україна, Yurii.Povesma@sgt.kphi.edu.ua,  
ORCID iD: 0009-0008-9283-0517

## СЕМАНТИЧНИЙ АНАЛІЗ КОНТЕКСТІВ РІЗНОЇ ДОВЖИНИ ЗА МЕТРИКОЮ ЛЕКСИЧНОЇ РІЗНОМАНІТНОСТІ MTLD

Семантичне розгортання сенсу речення в текст здійснюється головним чином шляхом збереження та варіації лексичних елементів, що утворюють спільне семантичне поле значимих концептів різних за розміром контекстів. Це дослідження на прикладі змісту коротких текстів (питання завдань) та довгих текстів (відповіді студентів у вигляді есе) з корпусу EFCAMDAT за допомоги міри лексичної текстової різноманітності (MTLD) розглядає лексичні методи представлення сенсу, здатні правильно відтворювати семантичну інформацію в контекстах різної довжини. Воно спрямоване на подолання нестачі даних для навчання великих мовних моделей (LLM) та сприяння професійній інтеграції та міжкультурній співпраці.

СЕМАНТИКА, СЕМАНТИЧНИЙ АНАЛІЗ ТЕКСТУ, КЛАСТЕРИЗАЦІЯ, МОДЕЛЮВАННЯ ТЕМ, ТЕМАТИЧНИЙ АНАЛІЗ, BERTopic, МІРА ЛЕКСИЧНОЇ РІЗНОМАНІТНОСТІ ТЕКСТУ (MTLD)

**О. М. Yurchenko, Yu. Yu. Povesma. Semantic analysis of contexts of varying lengths using the MTLT lexical diversity metric.** The semantic deployment of sentence meaning in text is achieved mainly by preserving and varying lexical elements that form a common semantic field of meaningful concepts in contexts of varying size. This study, using the example of content between short texts (task questions) and long texts (student answers in the form of essays) from the EFCAMDAT corpus, with the help of the Measure of Textual Lexical Diversity (MTLD), examines lexical methods of representing meaning that are capable of correctly reproducing semantic information in contexts of varying lengths. It aimed to overcome the lack of data for training large language models (LLMs) and to promote professional integration and intercultural cooperation.

SEMANTICS, SEMANTIC TEXT ANALYSIS, CLUSTERING, TOPIC MODELLING, TOPIC ANALYSIS, BERTopic, MEASURE OF TEXT LEXICAL DIVERSITY (MTLD)

### Вступ

У сучасному світі професійні сфери не обмежуються місцевим рівнем, а є міжнародними та міждисциплінарними. Це передбачає інтеграцію іноземців в іншу мовну, культурну та іноді в нову професійну сферу, що, як правило, не відбувається швидко і створює чимало труднощів для людини з певним багажем знань та іншомовною професійною підготовкою. У цьому контексті мета нашого дослідження — знайти метод представлення сенсу, здатний правильно відтворювати семантичну інформацію в текстах різної довжини — є дуже актуальною, оскільки це може вирішити проблему нестачі існуючих даних для навчання великих мовних моделей (LLM), які полегшують вирішення завдань обробки природної мови (NLP).

Оскільки семантичний перенос сенсу речення в текст відбувається головним чином шляхом збереження та варіації лексичних елементів, що утворюють спільне лексико-семантичне поле, вважається, що методи лексико-тематичного аналізу, такі як, наприклад, тематичне моделювання, можуть ефективно виявляти семантичну схожість між реченням і текстом [1]. Представлення сенсу тексту у вигляді «торб слів» (bag-of-words, BoW), своєю чергою, робить актуальним застосування методів лексичного аналізу.

Діапазон лексики, що використовується у живому людському мовленні, відображає розмаїття

словникового запасу, а також рівень володіння мовою, тому матеріалом дослідження було обрано корпус EFCAMDAT, який є відкритим корпусом робіт тих, хто вивчає англійську мову як іноземну [2]. Тож у цьому дослідженні ми вивчаємо, чи може лексична різноманітність (англ. Lexical Diversity (LD)) застосовуватися як для вимірювання розміру лексико-семантичних полів у текстах, написаних на одну ту саму тему, чи у відповідь на одне й те саме питання/завдання, так і для визначення ступеню семантичної близькості двох різних за довжиною контекстів: речення та тексту.

Ми виходимо з гіпотези, що подібні лексико-семантичні елементи у відповідях на одне й те саме питання в межах однієї мови можуть відрізнятися залежно від рівня володіння мовою, але завжди мають відповідати на поставлене питання і бути семантично співвідносними із заданою темою. Іншими словами, якщо 100 осіб відповідають на одне й те саме запитання, їхні відповіді можна звести до суті запитання, тобто до єдиного семантичного елемента, який може бути представленим у значимих поняттях/концептах, у ключових лексемах, а саме запитання можна вважати темою завдання, сформульованою у вигляді фрази.

Таким чином, метою нашої роботи є вирішення проблеми семантичної подібності чи семантичної спорідненості між контекстами різної довжини,

а саме завданням (питанням), сформульованим у вигляді речення, та відповідями на це запитання, у вигляді тексту.

Для досягнення зазначеної мети ми використовуємо два підходи: (1) результати тематичного аналізу та кластеризації та (2) порівняння лексичного представлення отриманих кластерів до кожного з питань корпусу за допомогою міри текстуальної лексичної різноманітності (англ. *Measure of Textual Lexical Diversity (MTLD)* [3]).

Основними науковими питаннями дослідження є:

– Як проявляється процес семантичної схожості/подібності між реченнями та текстом з точки зору лексичних і семантичних відносин?

– Як методи тематичного аналізу дозволяють виявити спільне семантичне поле між реченням і текстом?

– Які критерії можна використовувати для підтвердження семантичної подібності між коротким і довгим текстом?

– Які обмеження існують у текстах щодо переносу семантичного ядра лексико-семантичного поля?

– Чи існує межа, за якою «перенесення сенсу» перестає бути лексичним і стає прагматичним або дискурсивним явищем?

Робота структурована наступним чином: у Вступі представлено гіпотезу та цілі дослідження, у розділі 1 наведено огляд сучасних рішень з лексико-семантичного аналізу текстів, у розділі 2 описано методи та матеріали дослідження, розділ 3 присвячено результатам поточного проекту. У висновках ми робимо підсумки і пропонуємо майбутні перспективи дослідження.

## 1. Огляд існуючих рішень з лексико-семантичного аналізу текстів

Семантична схожість текстів або документів широко вивчається в різних сферах, включаючи обробку природної мови (NLP), порівняння документів, штучний інтелект, семантичну мережу тощо. Сучасні підходи до лексико-семантичного аналізу текстів базуються, з одного боку, на розширенні поняття семантичної подібності, з іншого, на використанні великих мовних моделей, які дають змогу спиратися на контекст.

Головною метою визначення семантичної подібності є вимірювання відстані між семантичними значеннями пари слів, фраз, речень або документів [4]. Семантична подібність (англ. *Semantic similarity*) – це метрика, де поняття відстані між елементами базується на схожості їхнього значення або семантичного змісту, на відміну від лексикографічної подібності [5]. Але оскільки термін «семантична подібність» зазвичай включає лише відношення «є», в нашому дослідженні ми також використовуємо поняття «семантична спорідненість» (англ. *Semantic relatedness*), як синонімічне, але більш широке поняття, оскільки

визначення семантичної спорідненості також включає до себе лексичну ієрархію [6], використання поняття лексико-семантичного поля [7, 8] та методів його вимірювання [9, 10, 11].

Сучасні методи тематичного моделювання текстів за допомогою контекстних LLM [1, 12] мають великий потенціал щодо семантичного аналізу та подальшого вдосконалення міжмовних трансферних моделей [13], що зазвичай використовують лексику для перенесення сенсу. Наразі моделі сімейства BERT вважаються одними з найбільш контекстуалізованих [14] завдяки методу TF-IDF, і порівняно з попередніми методами тематичного моделювання, такими як латентний розподіл Діріхле (LDA), які виявляються недостатніми для вирішення проблеми кореляції між текстом і темою (питанням), оскільки базуються лише на принципі ключових слів [15, 16]. Однак щоб підтвердити ступінь семантичної схожості текстів, необхідно враховувати розмір контексту, щоб ваги були пропорційними [9].

Новизна нашого підходу полягає у тому, що наше завдання є зворотним до анотування/реферування тексту, а запропоноване рішення базується не на суто лексичному аналізі, а на контекстуалізованих векторних представленнях сенсу у вигляді набору лексем, що репрезентують текст. Вимірювання лексичної різноманітності тексту (MTLD), яке детально описано в роботах [17, 18, 19], не обмежується набором слів, відомих автору чи читачеві, а також включає способи використання цих слів у текстах [3, 20, 21].

Таким чином, наш підхід базується на припущенні, що семантичний перенос сенсу речення в текст відбувається головним чином шляхом збереження та варіації лексичних елементів, які утворюють спільне семантичне поле між реченнями та текстами. Отже, методи тематичного аналізу (тематичне моделювання та кластеризація) допомагають виявити семантичну подібність між питанням/завданням і текстом відповіді на нього, а лексичний аналіз допомагає вимірювати міру її близькості.

## 2. Матеріали та методи дослідження

Представлене дослідження<sup>1</sup> є першим кроком у низці досліджень та було проведено на прикладі англійської мови для текстів з корпусу EFCAMDAT<sup>2</sup>, що складається з есе, написаних студентами різних рівнів володіння мовою у відповідь на задані викладачами завдання.

Корпус EFCAMDAT поділений відповідно до загальноприйнятих стандартів Загальноєвропейської системи оцінки мовних знань (CEFR)<sup>3</sup> на 6 рівнів: A1, A2, B1, B2, C1, C2. Він складається з відповідей сту-

<sup>1</sup> <https://gitlab.univ-lille.fr/olena.yurchenko.etu/mtld>

<sup>2</sup> <https://github.com/amichw/EFCAMDAT>

<sup>3</sup> <https://rm.coe.int/1680459f97>

дентів на поставлені запитання: коротких запитань та довших текстів відповідей. На кожному рівні є 24 запитання (теми завдань), за винятком рівня C2, який містить 8 запитань (тем завдань). Загалом корпус містить 128 запитань, які не повторюються з рівня на рівень. Це створює певні труднощі для проведення досліджень на однакові теми на різних мовних рівнях.

Цей корпус текстів також містить інформацію про помилки учнів, частини мови та граматичні відношення; усі завдання оцінені та прокоментовані викладачами [22]. Дані корпусу були нами відфільтровані за позитивною оцінкою від 60 до 100 балів. Загальна кількість даних становить 1 169 298 текстів, розподілених за рівнями: A1 – 621 231, A2 – 304 983, B1 – 166 161, B2 – 60 538, C1 – 14 499, C2 – 1 886.

Порівняння питань (теми завдань у формі речень) та відповідей (відповіді студентів у формі тексту) було здійснено за допомогою наступних бібліотек Python: UMAP, numpry та K-Means у Google Colab для всіх рівнів мови A1–C2 корпусу EFCAMDAT. В ході лексико-семантичного аналізу текстів було використано наступні методи:

1) Тематичний аналіз моделлю BERTopic [14], з виділенням ключових концептів методами KeyBERT, [23], TF-IDF [24, 25] та порівнянням їх за вагою.

2) Кластеризація за допомогою моделі all-MiniLM-L6-v2 (Yin і Zhang, 2024) з алгоритмом HDBSCAN (англ. Hierarchical Density-Based Spatial Clustering of Applications with Noise) [26].

3) Аналіз схожості текстів за методом Жаккара [27].

4) Аналіз лексичного складу текстів за MTLД [19].

Згідно з Treffers-Daller et al. (2018), MTLД обчислюється як середня довжина послідовностей тексту, що мають стабільне співвідношення типу-токену (TTR). Формально:

$$MTLD = \frac{N}{\sum_{i=1}^k S_i} \quad (1)$$

де  $N$  – загальна кількість слів у тексті, а  $S_i$  – довжина кожної послідовності, в якій TTR перевищує заданий поріг (зазвичай 0,72).

MTLD вважається менш чутливим до довжини тексту, що робить його придатним для аналізу коротких есе, написаних студентами. Однак слід зазначити, що різниця в довжині порівнюваних текстів не має перевищувати 50 токенів, тобто цей показник не підходить для порівняння питань і відповідей до них у формі есе [28]. Однак, цей метод дає змогу порівнювати згруповані в ході тематичного аналізу та кластеризації тексти між собою в рамках одного питання та робити висновки щодо міри схожості між ними.

### 3. Експериментальні дослідження

На першому етапі дослідження ми провели тематичний аналіз корпусу EFCAMDAT, результати якого описано в нашій роботі [29]. Далі було проведено кластеризацію, результати якої ми проілюстровали в даній роботі кількома яскравими прикладами, але аналіз був проведений для всіх текстів, що мали позитивну оцінку, і порівняний із 128 питаннями/завданнями, представленими одним реченням. Нарешті, отримані результати аналізу були порівняні за тематичним аналізом та кластеризацією.

На другому етапі було проаналізовано лексичний склад відповідей на кожне зі 128 питань за допомогою розрахунків відсотка перетину кластерів і шуму та коефіцієнта Жаккара в межах одного питання та між питаннями, в межах кожного з 6 рівнів і між рівнями, а також зроблено лексичний аналіз за допомогою MDLT [30].

### 4. Результати кластеризації

**4.1. Низький відсоток «шумності».** Результати кластеризації у порівнянні з тематичним аналізом показують низький відсоток «шумів». Так лише відповіді на 16 із 128 питань мають рівень шуму вище 10%.

Крім того, шум має тенденцію до зростання на вищих рівнях. Наприклад, два питання рівня C1: «Interpreting a prophecy» та «Writing about future lifestyles», а також одне питання рівня C2 «Writing a visualization script» мають шум 100%, тобто вони не були розділені на кластери. Це може свідчити про те, що лексичні та семантичні поля концептів відповідей практично не перетинаються, тобто кожне питання по-особливому сприймається студентом, а текст відповіді має свій сенс, змістовні акценти та тему. Однак усі вони отримали позитивну оцінку від викладача і мають семантично відповідати поставленому питанню. Тим більше, що за результатами лексичного аналізу їх лексична складова не дуже відрізняється за кластерами. Тож, в таких випадках ми маємо визнати, що обрані моделі не справляються з розрізненням семантики.

Також високий рівень шуму (понад 50%) спостерігається в деяких питаннях нижчих рівнів, наприклад, у питаннях рівня A2 «Describing people in photos» та рівня B2 «Giving feedback about a colleague», що вказує на те, що ймовірна причина певної кількості «шуму» залежить не від рівня мови, а від формулювання питань з використанням певних лексичних елементів: загальних або специфічних, що ще раз підкреслює актуальність лексичного аналізу.

**4.2. Кількість кластерів для одного питання.** Метою кластеризації було також показати відсоток перетину кластерів за ключовими словами в рамках одного питання для кожного з 128 випадків. Ми порівняли 50 ключових слів кожного кластера з 100 ключовими

словами питання в цілому, а також співвідношення між кластерами та шумом на основі 50 ключових слів для всіх відповідей, що містять шум (32 питання зі 128).

Наприклад, вже згаданий випадок з високим рівнем шуму, такий як питання рівня A2 «Describing people in photos», яке містить 60% текстів із шумом у відповідях і розподіл інших текстів на 5 кластерів,

є особливо цікавим, коли порівнюються кластери в межах одного питання.

На графіку (рис. 1) відносини між кластерами позначені синім кольором, відносини між кластерами та шумом – рожевим, а відносини між кластерами та 100 спільними ключовими словами у відповідях на питання рівня A2 «Describing people in photos» – зеленим.

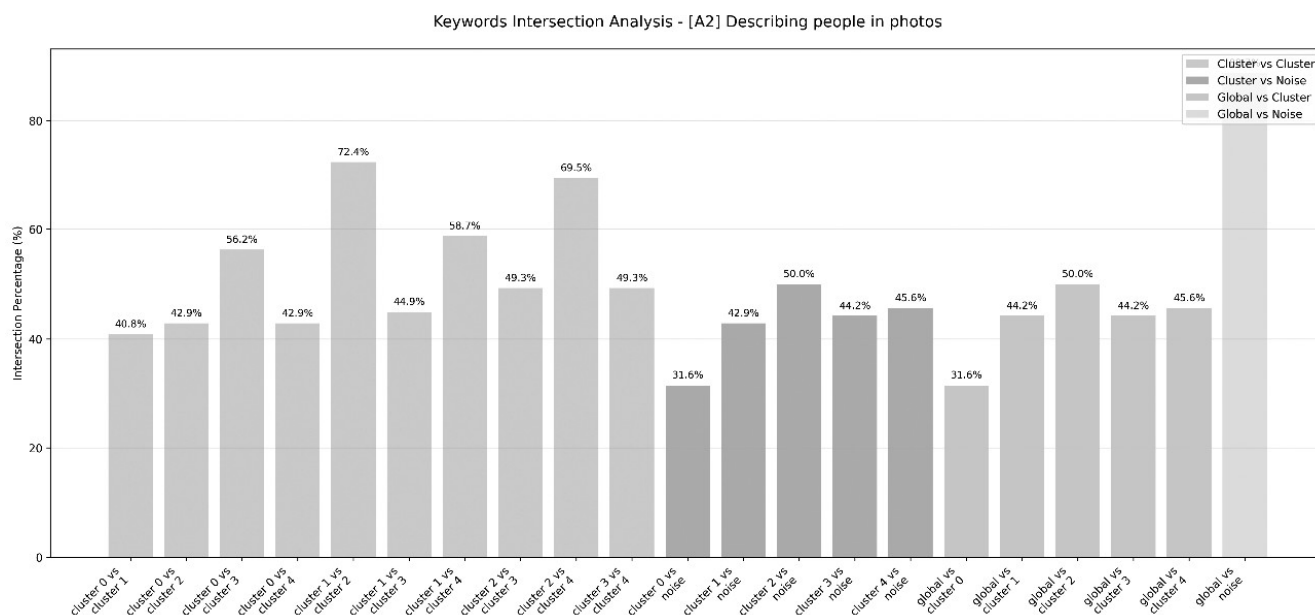


Рис. 1. Відношення кластерів та шуму між собою та до 100 загальних ключових слів питання рівня A2 «Describing people in photos»

Оскільки питання, що містять шуми, представляють найбільшу цікавість з точки зору лексичного складу, ми провели додатковий аналіз шумів, які поводяться як звичайні кластери. Було взято по 20 ключових слів з кожного з 16 питань, шум у яких перевищував 10%, класифіковано їх за TF-IDF-балом та значенням у тексті та порівняно з лексичним складом інших кластерів. Таким чином, ми отримали підтвердження того, що шуми поводяться як окремі кластери і що їх лексика дійсно відповідає поняттям питання.

Для наведеного питання рівня A2 «Describing people in photos», де шум становить 60%, можна бачити, що першим ключовим словом кластерів, і шуму в тому числі, може бути не лише лексема «photo», яка повторює значиме поняття питання, див. рис. 2. Це доводить, що лексичний склад як звичайних, так і шумних кластерів може варіюватися, але він має залишатися в межах одного лексико-семантичного поля, обмеженого основними поняттями питання, оскільки ми взяли всі роботи студентів, які отримали хорошу оцінку.

Однак, як і в природній мові, межі семантичної подібності не зводяться до синонімії, а також включають антонімію, гіперонімію, гіпонімію та інші явища семантичної спорідненості, включаючи асоціації, зміни в порядку слів за рангом у кожному кластері

відповідатимуть індивідуальному представленню тексту автором.

cluster_id	texts size	top_keywords
1	759	wearing, photo, look, hair, black
0	93	photo, wearing, picture, friend, brazil
2	1000	wearing, photo, hair, look, friend
3	915	photo, picture, wearing, friend, look
4	155	photo, wearing, look, picture, shirt
Total	2922	Clusters in total: 5

Рис. 2. Top-5 ключових слів до кластерів питання рівня A2 «Describing people in photos»

Результати співвіднесення кластерів та шумів для кожного зі 128 питань корпусу EFCAMDAT було порівняно також на основі відсотка перетину та коефіцієнта Жаккара, а потім представлені у вигляді матриць перетину для кожного питання, див. рис. 3.

Аналізуючи матриці перетину кластерів і шумів у відсотках та матрицю перетину кластерів і шумів за коефіцієнтом Жаккара, складені для 128 питань і продемонстровані на прикладі відповідей на питання рівня A2 «Describing people in photos», ми дійшли висновку, що:

1) в цілому шуми в питаннях, де вони виділені, поводяться як окремі кластери;

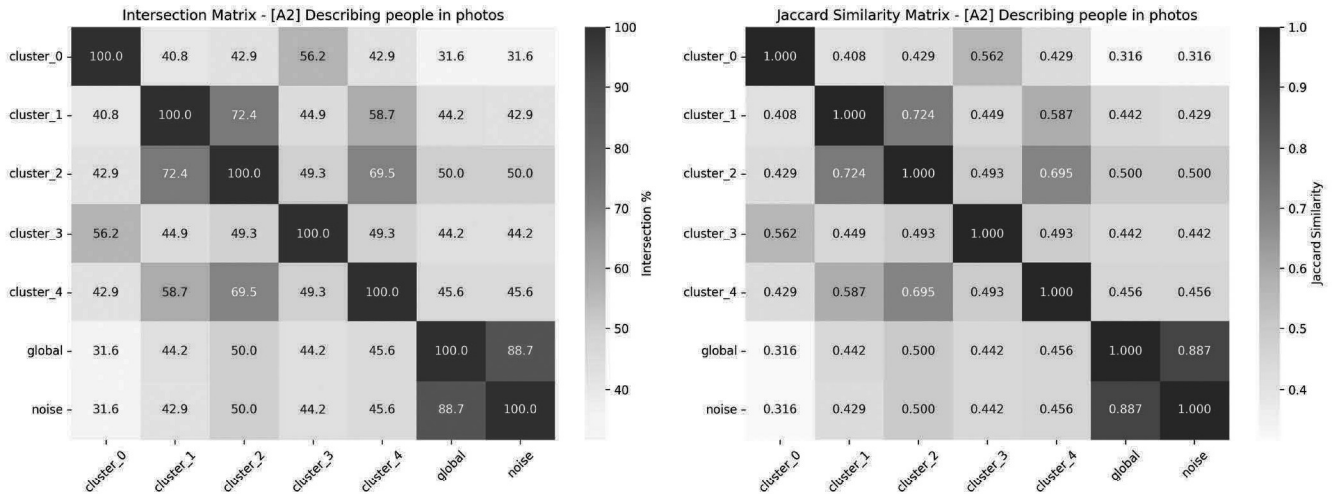


Рис. 3. Матриці перетину текстів 5 кластерів та шумів у порівнянні до 50 загальних ключових слів на основі відсотка перетину та коефіцієнта Джакарда для питання *A2 Describing people in photos*

2) лексичний склад кластерів в цілому дуже схожий і може варіюватися залежно від ваги окремих одиниць, що відображається в рейтингу слів серед 100 спільних ключових слів відповідей на питання;

3) лише відповіді на 10 питань зі 128 містять більше 10 кластерів, а в 30% випадків – більше 5 кластерів на питання.

Таким чином, розподіл кластерів можна вважати досить рівномірним. Він не залежить від рівня мови або кількості шумів, тому, на нашу думку, випадки, що відхиляються від загального розподілу, заслуговують на особливу увагу, зокрема на рівні лексико-семантичного аналізу.

**4.3. Перетин кластерів за лексичним складом та рівнями.** Хоча питання не повторюються на кожному з рівнів мови, нами було проведено порівняльний аналіз лексики за рівнями та групами рівнів, щоб побачити, наскільки лексичний склад повторюється між ними та за питаннями.

Результати показали, що, хоча найвищі показники лексичного перетину, а саме 85% для питань рівня C1 «Making a movie» і 65% для «Applying for sponsorship», стосуються питань високого рівня, лексичний склад відповідей все ж більше залежить від формулювання питання: більш конкретне чи абстрактне – ніж від самого рівня.

Згідно з рейтингом схожості кластерів за рівнем мови, що складає для рівня C2– 0,0484, C1– 0,0420, B1 – 0,0378, B2– 0,0369, A1 – 0,0764, A2 – 0,0151 і порівнюючи лексичний склад питань відповідних рівнів, треба відзначити, що:

- найнижчий рівень, A1, за класифікацією рівнів за схожістю, є вищим за A2, а також B1 порівняно з B2, за винятком останнього випадку групи B, де схожість є більшою;
- за кількістю тем рівні A1 і B2, а також B1 і C1 наближаються, а не ті, що є сусідніми;
- але за середньою схожістю кластерів рівнів

найвищі, C1 і C2, вирізняються.

Також були проведені розрахунки 1) кількості нових слів, що додаються на кожному рівні мови; 2) кількості слів, що залишаються на попередньому рівні; 3) спільної кількості слів між рівнями та групами мови. Найбільш чітко представлена інформація для кожного рівня наведена на діаграмах Вєнна, див. рисунок 4, який ви можете переглянути на наступній сторінці. Діаграми показують певну нерівномірність у додаванні нового словника, особливо між проміжними рівнями, такими як A2 і B1, де до рівня B1 додається на 1500 нових слів менше, ніж до рівнів A2 і B2. Проте приблизно половина лексичних одиниць залишається спільною для рівнів A1-A2, A2-B1, B1-B2, B2-C1 і має тенденцію до збільшення до рівня C1. Значне зменшення кількості нових слів, унікальних для рівня C2, пояснюється тим, що цей рівень містить лише 8 питань, що значно звужує спектр тем, які розглядаються, і, як наслідок, кількість використовуваного словника.

Це пояснюється тим, що, незважаючи на відсутність повторення питань за рівнем вивчення мови, лексика рівномірно переходить з одного рівня на інший, групуючись за лексико-семантичними полями основних понять запропонованих завдань. Але в групі C вона асимілюється з вільним володінням мовою, і лексика цих рівнів є найбагатшою. Однак через обмежену кількість питань і, як наслідок, недостатню кількість текстів на рівні C2, ми не можемо повною мірою проаналізувати все розмаїття тем на цьому рівні.

Таким чином, ми доходимо висновку, що: 1) незалежно від того, чи йдеться про кластери, чи про шумні кластери, вони дуже схожі за лексичним складом; 2) порівняння лексичного складу різних кластерів не дозволяє відповісти на питання, чому в деяких питаннях (завданнях) розподіл відбувається за великою кількістю кластерів, в інших – за малою, в деяких – шум виділяється, в інших – ні.

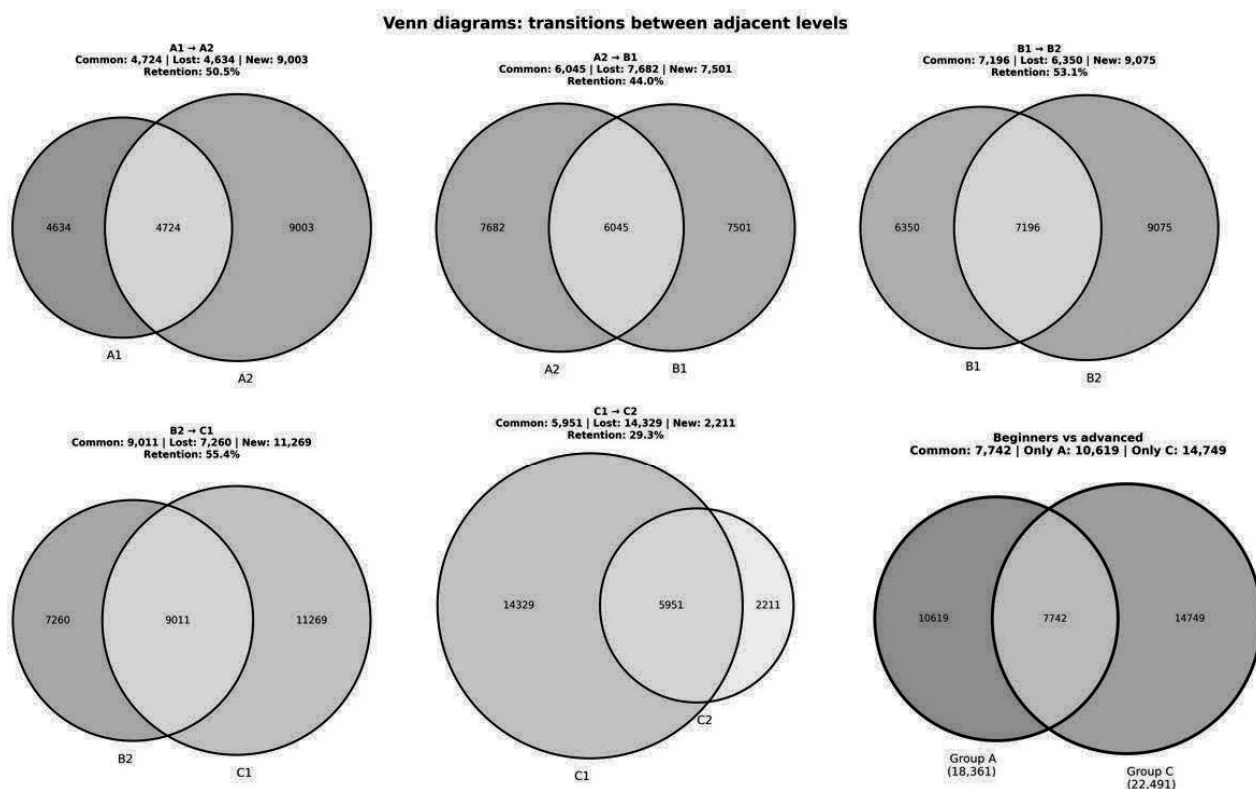


Рис. 4. Діаграми Венна та кількість лексики, що переходить до наступного рівня мови за рівнями та між групами рівнів А та С

## 5. Вимірювання лексичної різноманітності тексту

**5.1. MTLD за рівнями мови.** Для оцінки достовірності отриманих значень було проведено порівняння з типовими значеннями MDLT для різних рівнів мови, як зазначено в роботі Treffers-Daller et al. (2018). Наші значення MDLT для корпусу EFCAMDAT за рівнями A1 – 44,6, A2 – 49,4, B1 – 63,2, B1 – 65,8, C1 – 67,9, C2 – 69,5 відповідають діапазону типових значень, отриманих після лематизації, і показують збільшення лексичної різноманітності до рівнів С, навіть якщо на рівні С2 немає великої кількості текстів (див. рис. 5):

- порівняння загального MTLD рівня та середнього значення за документом показує зростання лексичного запасу до високих рівнів;

- порівняння MTLD за документами всередині рівня показує зменшення лексичного запасу до високих рівнів через зниження кількості документів і питань, особливо на рівні С2.

Крім того, було встановлено, що при застосуванні лематизації кількість неправильно написаних слів, що вважаються унікальними, значно зменшувалася із підвищенням рівня мови. Таким чином, ми змогли мінімізувати їх вплив на розрахунки.

**5.2. MTLD за лексико-семантичними полями.** Як зазначалося вище, питання корпусу EFCAMDAT не повторюються, тому було запропоновано проаналізувати відповіді на питання за лексемами, які можна віднести до одних і тих самих лексичних семантичних полів (ЛСП). Вручну було відібрано питання, що можна віднести до одних і тих самих ЛСП на основі значимих

концептів питання, які ми умовно назвали «звички» і які присутні на всіх рівнях: A1 «Describing Your Family Eating Habits» (концепт «habit»); A2 «Describing Routines» (концепт «routine»); B1 «Writing A Job Advertisement» (концепт «advertisement»); B2 «Setting Rules For Social Networking» (концепт «rule»); C1 «Writing About Future Lifestyles» (концепт «lifestyle»); C2 «Following A Code Of Ethics» (концепт «code»). Лексика цих питань, як і всіх інших, була проаналізована за допомогою кластеризації по 20, 50 і 100 ключових слів на кластер, а також «мішків слів» у 4, 10 і 100 найчастіших словах в темі. Лексеми питань були порівняні з лексемами відповідей за рівнями. Аналіз 4 ключових слів BERTopic показує, що концепти ЛСП «звички» були присутні лише в декількох темах, що відповідають цим питанням, наприклад: тема 5 (working style, working style, habits, problems brought, ask improve work), тема 6 (working style, working style, habits, ask improve work, problems brought).

Цей результат також можна пояснити тим, що наявність певних концептів у запитаннях лише дає орієнтир для вибору у текстах відповідей певних лексем із відповідних лексико-семантичних полів. Самі концепти, взяті із запитань у формі речень, також варіюються залежно від того, як сформульовано запитання.

Таким чином, ми робимо висновок, що збільшення кількості проаналізованих слів з 4–10 до 20–50 підвищує ймовірність знайти поняття у відповідях. Але додаткове збільшення з 50 до 100 ключових слів кластера порівняно зі 100 ключовими словами, спільними для кожного питання, показує, що близькість до лексико-семантичних понять питань зменшується.

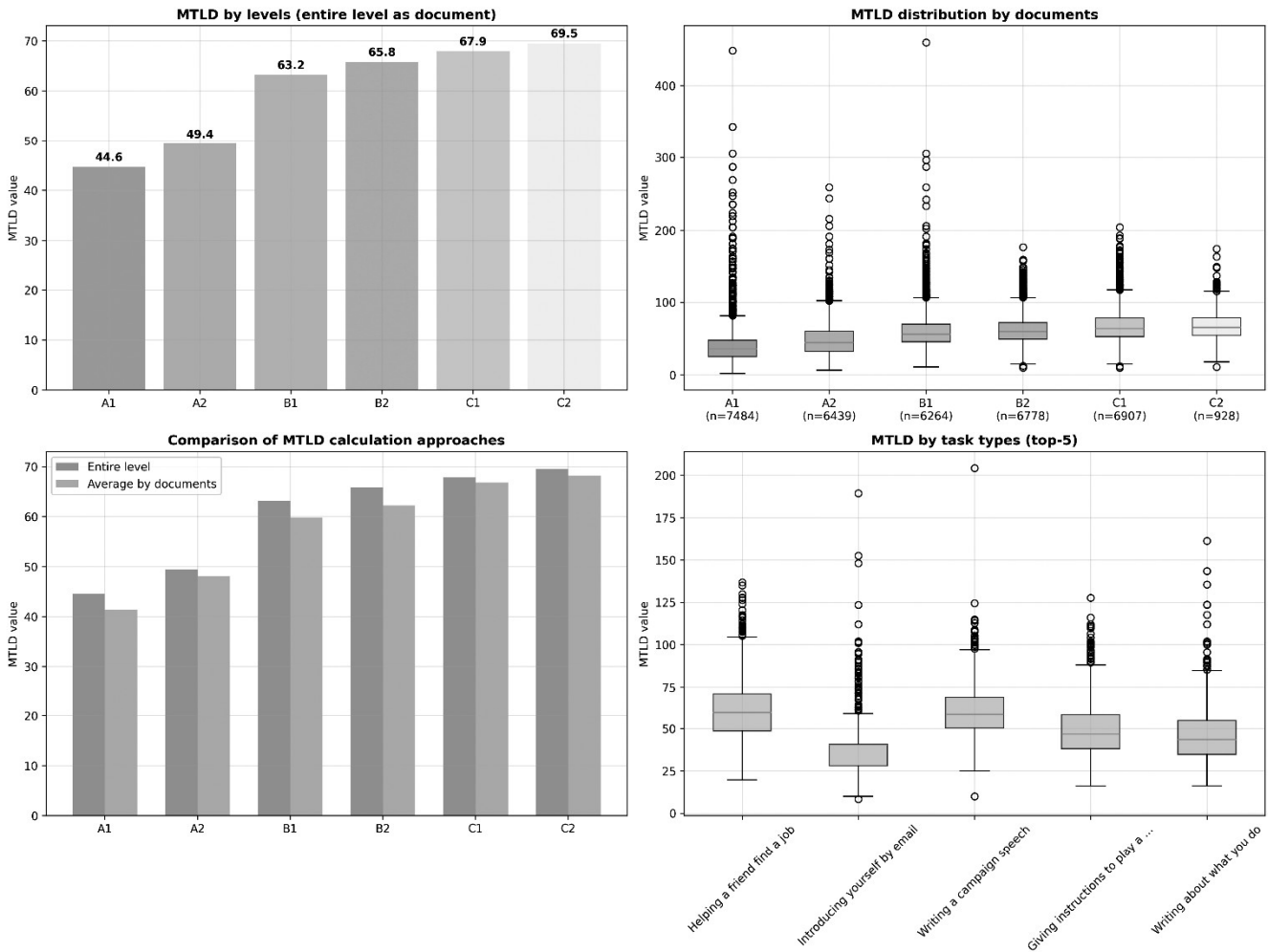


Рис. 5. Результати Measure of Textual Lexical Diversity (MTLD) за рівнями мови

**5.3. MTLD за лексичним складом кластерів.**

Лексичний аналіз за допомогою методу MTLD [30] (Measure of Textual Lexical Diversity) порівняння ключових слів, виконаний додатково на прикладі завдання рівня C1 «Writing a campaign speech» (укр. Написання передвиборчої промови), показує, що алгоритм кластеризації, хоча і розділяє тексти відповідей на завдання на 3 окремі кластери, показані на рис. 6.

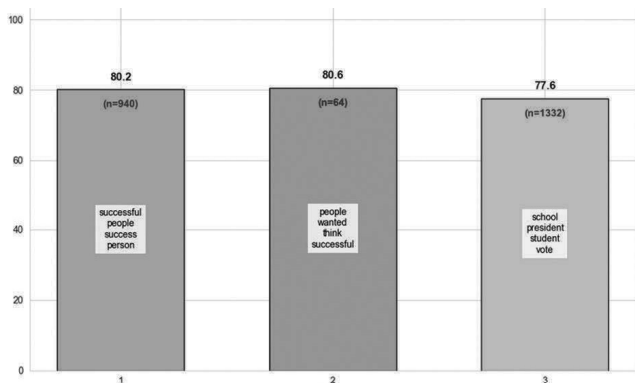


Рис. 6. Порівняння MTLD за кластерами для завдання рівня C1 «Writing a campaign speech»

Але семантично їх можна поділити на 2 теми, які умовно ми назвали:

— «Шкільні вибори»: (лексеми «school», «president», «student», «vote»). Причому цей кластер показав нижчий MTLD (~77,6), оскільки у відповідях використовувався досить обмежений, повторюваний словник.

— «Філософія успіху»: (лексеми «people», «success», «wanted», «successful»). Кластер з текстами, що використовували вказану лексику, показав значно вищий MTLD (80,2–80,6), оскільки абстрактні теми вимагають більш багатого словника.

Таким чином, ми дійшли висновку, що хоча лексичний аналіз складу кластерів допомагає знайти семантичні відмінності між текстами, представлення результатів лексико-семантичного аналізу у вигляді «мішків слів» залишає це на рівні припущень.

**Висновки**

На основі результатів, отриманих за допомогою моделей кластеризації, тематичного аналізу та індексу лексичної різноманітності текстів MDLT, а також з урахуванням обмежень корпусу EFCAMDAT, зокрема відсутності повторюваних завдань на різних мовних рівнях, ми дійшли таких висновків:

1) Аналіз MDLT показав, що між рівнями немає значних лексичних відмінностей, за винятком

орфографічних помилок на нижчих рівнях. Таким чином, лексичний склад відповідей на питання залежить головним чином від формулювання питання.

2) Щодо існування спільного лексичного та семантичного поля між реченням і текстом, слово саме по собі не є комунікативною синтаксичною одиницею, оскільки його сенс походить від речення, що є мінімальною одиницею мовлення. Отже, для семантичного аналізу краще розглядати контексти на прагматичному та дискурсивному рівнях, а не лише на лексичному чи синтаксичному.

3) У процесі семантичного розгортання речення в тексті аналіз лексичного складу кластерів моделлю all-MiniLM-L6-v2 виявився ефективнішим для групування текстів, ніж метод тематичного аналізу, проведений моделлю BERTopic.

4) Незалежно від того, чи йдеться про кластери, чи про кластери з шумом, їх лексичний склад дуже схожий.

5) Порівняння лексичного складу різних кластерів не дозволяє нам відповісти на питання: чому в деяких завданнях розподіл відбувається на велику кількість кластерів, а в інших – на маленьку, і чому в деяких випадках шум виділяється, а в інших – ні. Ми припускаємо, що це залежить від формулювання питання.

6) Хоча лексичний аналіз складу кластерів допомагає виявити семантичні відмінності між текстами, представлення результатів у вигляді мішків слів залишає їх у стані гіпотези.

Тому для вирішення проблеми семантичного аналізу текстів в подальших дослідженнях планується використовувати контекстуалізовану векторну репрезентацію значення тексту у вигляді CLS-токену, отриманого з передостаннього шару нейронної мережі [31, 32, 33], та метрики тематичного аналізу текстів [34].

### Подяки

Дослідження, описане в цій статті, провадилося в рамках написання дисертації, що виконується під керівництвом д.т.н. Чередніченко О.Ю. на кафедрі програмної інженерії та інтелектуальних технологій управління Національного технічного університету «Харківський політехнічний інститут» (Україна), та у межах проекту ANR-17-CE19-0016 CLEAR під керівництвом Наталії Грабар в лабораторії STL – Savoirs, Textes, Langage (UMR 8163 – CNRS) університету Лілля. Ми дякуємо за фінансову підтримку програми PAUSE ANR (Франція), а також студенту магістратури Національного технічного університету «Харківський політехнічний інститут» Арсенію Лукашевському, за допомогу в розрахунках.

### Список літератури:

- [1] Wu, X., Nguyen, T., Zhang, D., Wang, W. Y., & Luu, A. T. (2025). FASTopic: Pretrained Transformer is a Fast, Adaptive, Stable, and Transferable Topic Model. *Advances in Neural Information Processing Systems*, 37, 84447-84481.
- [2] Eertzen G. J., A. T. Lexopoulou A. T., Korhonen A. (2013). Automatic linguistic annotation of large scale L2 databases: The EF-Cambridge open language database (EFCAMDAT), in: 31st Second Language Research Forum (SLRF).
- [3] Jarvis, S. (2013). Capturing the Diversity in Lexical Diversity. *Language Learning*, 63(s1), 87-106.
- [4] Nguyen, M. H., Tran, D. Q. (2021). Estimation in semantic similarity of texts, *Journal of Information Science and Engineering* 37 (2021) 617–633.
- [5] Harispe, S., Ranwez, S., Montmain, J. (2022). *Semantic similarity from natural language and ontology analysis*, Springer Nature.
- [6] Feng, Y., Bagheri, E., Ensan, F., Jovanovic, J. (2017). The state of the art in semantic relatedness: a framework for comparison, *The Knowledge Engineering Review* 32 (2017) 1–30.
- [7] Andersen, P. B. (1990). *A theory of computer semiotics: semi-otic approaches to construction and assessment of computer systems*, Vol. 3, Cambridge University Press.
- [8] Jackson, H., Zé, E. (2000). *Amvela, Words, Meaning, and Vocabulary*, Continuum.
- [9] Vakulenko, M. (2022). Semantic comparison of texts by the metric approach, *Digital Scholarship in the Humanities* 38 (2) (2022) 766–771.
- [10] Lin, Y.-S., Jiang, Y., Lee, S.-J. (2014). A similarity measure for text classification and clustering, *IEEE Trans. on Knowledge and Data Engineering* 26 (2014) 1575–1590. doi:10.1109/TKDE.2013.19.
- [11] Tytgat, J., Wisniewski, G., Betrancourt, A. (2024). Evaluation de la similarité textuelle : Entre sémantique et surface dans les représentations neuronales. In JEPTALNRECITAL.
- [12] Warner, B., Chaffin, A., Clavié, B., Weller, O., Hallström, O., Taghadouini, S., ... Poli, I. (2024). Smarter, better, faster, longer: A modern bidirectional encoder for fast, memory efficient, and long context finetuning and inference, arXiv preprint arXiv:2412.13663.
- [13] Ruder, S., Vulic, I., & Søgaard, A. (2019). A survey of cross-lingual word embedding models. *Journal of Artificial Intelligence Research*, 65(1):569–630.
- [14] Grootendorst, M. R. BERTopic: Neural topic modeling with a class-based TF-IDF procedure, arXiv preprint arXiv:2203.05794 (2022).
- [15] Blei, D. M., McAuliffe, J. D. (2010). Supervised Topic Models, in: *Advances in Neural Information Processing Systems*.
- [16] Blei, D. M., Ng, A. Y., Jordan, M. I. (2003). Latent Dirichlet Allocation, *Journal of Machine Learning Research* 3.993–1022.
- [17] McCarthy, P. M., & Jarvis, S. (2010). MTL, vocd-D, and HD-D: a validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42(2), 381–392.
- [18] Bonvin, A., & Lambelet, A. (2017). Algorithmic and subjective measures of lexical diversity in bilingual written corpora: a discussion. *Corela*.
- [19] Treffers-Daller, J., Parslow, P. & Williams, S. (2018). Back to basics: how measures of lexical diversity can help discriminate between CEFR levels. *Applied Linguistics*, 39 (3). pp. 302-327. ISSN 1477-450X.

- [20] Laufer, B. et Nation, P. (1995). Vocabulary size and use : Lexical richness in L2 written production. *Applied Linguistics*, 16, 307-322.
- [21] Duran, P., Malvern, D., Richards, B. et Chipere, N. (2004). Developmental trends in lexical diversity. *Applied Linguistics*, 25, 220-242.
- [22] Huang, Y., J. Geertzen, R. Baker, A. Korhonen, and T. Alexopoulou (2017). The EF Cambridge Open Language Database (EFCAMDAT) : Information for Users, University of Cambridge and EF Education First.
- [23] Issa, B., Jasser, M.B., Chua, H.N., & Hamzah, M. (2023). A Comparative Study on Embedding Models for Keyword Extraction Using KeyBERT Method. 2023 IEEE 13th International Conference on System Engineering and Technology (ICSET), 40-45.
- [24] Egger, R., Yu, J. (2022). A Topic Modeling Comparison Between LDA, NMF, Top2Vec, and BERTopic to Demystify Twitter Posts, *Frontiers in Sociology* 7 (2022).
- [25] Babalola, O., Ojokoh, B., Boyinbode, O. (2024). Comprehensive Evaluation of LDA, NMF, and BERTopic's Performance on News Headline Topic Modeling, *Journal of Computing Theories and Applications* 2 (2024) 268–289.
- [26] Malzer, C., & Baum, M. (2019). HDBSCAN( $\epsilon$ ): An Alternative Cluster Extraction Method for HDBSCAN. *ArXiv*, abs/1911.02282.
- [27] Travieso, G., Benatti, A., & Costa, L.D. (2024). An Analytical Approach to the Jaccard Similarity Index.
- [28] Koizumi, R. (2012). Relationships Between Text Length and Lexical Diversity Measures : Can We Use Short Texts of Less than 100 Tokens ? *Vocabulary Learning and Instruction*, 1(1), 60-69.
- [29] Grabar, N., Yurchenko, O., Cherednichenko, O., and Lukashovskyi, A. (2025). Exploring Semantic Similarity in English Learners' Texts through Topic Modelling. In *CLW-2025: Computational Linguistics Workshop at 9th International Conference on Computational Linguistics and Intelligent Systems (CoLInS-2025)*, May 15–16, 2025, Kharkiv, Ukraine Vol-3976, p. 92-106.
- [30] Yurchenko, O., Grabar, N., Cherednichenko, O. Analyse du vocabulaire appliquée au transfert sémantique phrase→texte. In: *Actes de la conférence Extraction et Gestion des Connaissances (EGC 2026)*. Vol. RNTI-E-42. Toulouse: Cépaduès-Éditions, 2026. P. 373–374.
- [31] Mollas, I., Bassiliades, N. et Tsoumakas, G. (2019). LioNets : Local Interpretation of Neural Networks through Penultimate Layer Decoding. *arXiv : Learning*.
- [32] Bianchi, F., Terragni, S., et Hovy, D. (2020). Pre-training is a Hot Topic : Contextualized Document Embeddings Improve Topic Coherence. *Annual Meeting of the Association for Computational Linguistics*.
- [33] Bouzina, S., De Rossi, D., Pavlov, V. G. et Moretti, S. (2024). Semantic Latency Mapping of Contextual Vector Embeddings in Transformer-Based Models.
- [34] Terragni, S., Fersini, E., & Messina, E. (2021). Word Embedding-Based Topic Similarity Measures. *International Conference on Applications of Natural Language to Data Bases*.

*Received (Надійшла) 22.01.2026*

*Accepted for publication (Прийнята до друку) 15.02.2026*

*Publication date (Дата публікації) 27.03.2026*



М. А. Гринченко<sup>1</sup>, Д. В. Лісова<sup>2</sup>, Д. О. Куценко<sup>3</sup>

<sup>1</sup>НТУ «ХПІ», м. Харків, Україна, [marinagrunchenko@gmail.com](mailto:marinagrunchenko@gmail.com);  
ORCID ID: 0000-0002-8383-2675

<sup>2</sup>НТУ «ХПІ», м. Харків, Україна, [daryna.lisova@cs.khpi.edu.ua](mailto:daryna.lisova@cs.khpi.edu.ua);  
ORCID ID: 0009-0004-2396-5144

<sup>3</sup>НТУ «ХПІ», м. Харків, Україна, [cureghost17@gmail.com](mailto:cureghost17@gmail.com);  
ORCID ID: 0009-0005-6359-3143

## ПРОЄКТНА МОДЕЛЬ РЕКОМЕНДАЦІЙНОЇ СИСТЕМИ ПІДТРИМКИ КЛІНІЧНИХ РІШЕНЬ ДЛЯ СІМЕЙНОГО ЛІКАРЯ

У статті розглянуто проєктування рекомендаційної системи підтримки прийняття клінічних рішень для сімейного лікаря в умовах зростання обсягів медичних даних та інформаційного перевантаження. Проаналізовано сучасні підходи до побудови Health Recommender Systems і виявлено їхні основні обмеження, зокрема недостатню пояснюваність (тобто здатності алгоритму обґрунтувати, чому було запропоновано певне рішення або рекомендацію), складність інтеграції з електронними медичними записами та ризики конфіденційності даних. Запропоновано проєктну модель рекомендаційної системи, що поєднує методи штучного інтелекту, бази знань доказової медицини та підходи Explainable AI. Розроблено контекстну, функціональну, поведінкову та компонентну діаграми, які формалізують архітектуру та логіку взаємодії системи. Запропонована модель орієнтована на практичне застосування та спрямована на підвищення обґрунтованості клінічних рішень і зменшення кількості необґрунтованих обстежень.

ІНТЕЛЕКТУАЛЬНА СИСТЕМА, РЕКОМЕНДАЦІЙНА СИСТЕМА, ПІДТРИМКА ПРИЙНЯТТЯ РІШЕНЬ, СІМЕЙНИЙ ЛІКАР, ШТУЧНИЙ ІНТЕЛЕКТ, МОДЕЛЬ, ГЛИБОКЕ НАВЧАННЯ

**M. A. Grinchenko, D. V. Lisova, D. O. Kutsenko. Conceptual design of a clinical decision support recommender system for family physicians.** The article is devoted to the design of a recommendation system for clinical decision support for family physicians in the context of increasing volumes of medical data and information overload. Modern approaches to the development of Health Recommender Systems are analyzed, and their key limitations are identified, including insufficient explainability, limited integration with electronic medical records, and data privacy risks. A project-oriented model of a recommendation system that combines artificial intelligence methods, evidence-based medicine knowledge bases, and Explainable AI approaches is proposed. Context, functional, behavioural, and component diagrams are developed to formalise the system architecture and interaction logic. The proposed model is oriented toward practical application and aims to improve the validity of clinical decisions and reduce unnecessary diagnostic tests.

INTELLIGENT SYSTEM, RECOMMENDATION SYSTEM, DECISION SUPPORT, FAMILY DOCTOR, ARTIFICIAL INTELLIGENCE, MODEL, DEEP LEARNING

### Вступ

У сучасних умовах стрімкої цифровізації медицини особливої актуальності набуває питання підвищення ефективності процесу прийняття клінічних рішень на первинному етапі надання медичної допомоги. Сімейний лікар щоденно опрацьовує значні обсяги інформації, що вимагають оперативного аналізу даних пацієнта та формування індивідуального плану лікування відповідно до чинних медичних протоколів і рекомендацій.

Різноманітність джерел клінічних даних, поява нових методів діагностики, велика кількість лікарських засобів і постійне оновлення медичних настанов створюють ефект інформаційного перевантаження, який суттєво ускладнює вибір оптимального лікувального підходу. Згідно з результатами багаторічних досліджень, понад 20% лабораторних тестів та інструментальних обстежень виконуються невиправдано або помилково, що призводить до зростання фінансових витрат, нераціонального використання ресурсів і затримок у постановці діагнозу [1].

Одним з перспективних напрямів вирішення цих проблем є інтелектуальні системи, які допомагають лікарям приймати рішення (Clinical Decision Support Systems, CDSS).

Інтелектуальна система — це штучна система, здатна сприймати навколишнє середовище, аналізувати дані, навчатися на основі досвіду та приймати раціональні рішення для досягнення визначеної мети [2]. Вона поєднує методи штучного інтелекту, машинного навчання та обробки знань, забезпечуючи автономність і адаптивність.

Одним із різновидів інтелектуальних систем є рекомендаційні системи (Recommender Systems), які спеціалізуються на формуванні персоналізованих пропозицій. Згідно з [3], рекомендаційна система — це програмна система, що прогнозує переваги користувача або “рейтинг” об’єкта, використовуючи інформацію про минулу поведінку та схожих користувачів. У медицині такі системи отримали назву Health Recommender Systems (HRS) і використовуються для створення персоналізованих планів лікування, вибору

діагностичних тестів або профілактичних заходів [4]. Інтелектуальні системи у сфері охорони здоров'я активно застосовують методи штучного інтелекту для аналізу клінічних даних пацієнтів, прогнозування перебігу захворювань і формування доказових рекомендацій. Зокрема, медичні рекомендаційні системи розглядаються як один із ключових напрямів розвитку цифрової медицини, оскільки забезпечують персоналізований підхід до кожного пацієнта шляхом інтеграції його індивідуальних характеристик, історії хвороби та сучасних клінічних протоколів. Такі системи виконують не лише інформаційно-консультативну функцію для пацієнтів, надаючи поради щодо лікування, профілактики або способу життя, але й допомагають лікарям приймати обґрунтовані клінічні рішення тобто обирати оптимальні діагностичні тести, терапевтичні стратегії та лікарські засоби з урахуванням індивідуальних особливостей організму, супутніх патологій і можливих протипоказань [5, 6]. Попри значний прогрес у розвитку таких систем, більшість існуючих рішень залишається вузькоспеціалізованими (орієнтованими на кардіологію, онкологію, ендокринологію тощо) або спрямованими переважно на підтримку пацієнтів, а не лікарів-практиків [7, 8]. Попри досягнутий прогрес у галузі інтелектуальних медичних систем, багато існуючих рішень характеризуються відсутністю пояснювальності (explainability), тобто здатності алгоритму обґрунтувати, чому було запропоновано певне рішення або рекомендацію.

Недостатня прозорість у роботі моделей штучного інтелекту створює бар'єри для формування довіри з боку медичних працівників і знижує ефективність впровадження таких технологій у клінічну практику [9, 10, 11, 12]. Як зазначено в дослідженнях [10, 11], пояснюваність є ключовим чинником для прийняття медичних інформаційних систем лікарями, оскільки вона безпосередньо пов'язана з етичними аспектами, відповідальністю та контролем якості рішень. Відповідно до міжнародних стандартів GDPR (General Data Protection Regulation) [13] та HIPAA (Health Insurance Portability and Accountability Act) [14] у процесах обробки медичних даних особливу увагу приділено прозорості, безпеці та етичності, що передбачає можливість відстеження логіки прийняття рішень системою [5, 12].

З огляду на зазначене, виникає потреба у створенні рекомендаційної системи підтримки прийняття рішень, яка поєднувала б точність аналітичних алгоритмів із можливістю пояснення результатів і забезпечувала контроль із боку лікаря. Така система має допомагати сімейному лікарю швидко визначати доцільні обстеження, уникати зайвих тестів, формувати раціональний план лікування та, як наслідок, підвищувати якість медичної допомоги.

## 1. Огляд рекомендаційних систем у медицині

У сучасній медицині рекомендаційні системи, відомі як HRS, розглядаються як один із найперспективніших напрямів розвитку цифрових технологій. Їхнє основне призначення полягає у підтримці прийняття клінічних рішень та персоналізації медичних послуг для кожного пацієнта [5].

Головною функцією таких систем є аналіз клінічних даних, включно з результатами обстежень, симптомами, медичними протоколами та історією хвороби пацієнта, з метою формування обґрунтованих рекомендацій для лікаря. На відміну від традиційних експертних систем, які базуються на жорстко визначених правилах, HRS використовують методи штучного інтелекту, машинного навчання та обробки природної мови, що дає змогу створювати гнучкі, контекстно-залежні поради та прогнози [6]. Результати сучасних досліджень свідчать, що понад 7% усіх запитів у Google пов'язані з питаннями здоров'я, що демонструє високий рівень зацікавленості користувачів у персоналізованій медичній інформації [7]. Проте надлишок даних у відкритих джерелах часто призводить до інформаційного перевантаження як серед пацієнтів, так і серед лікарів. У цьому контексті системи HRS виконують функцію інтелектуального фільтра, який забезпечує користувачам доступ до перевіреної, науково обґрунтованої інформації щодо діагностики, лікування, профілактики або вибору лікарських засобів. Використання таких систем сприяє підвищенню ефективності роботи лікарів, зменшенню кількості помилкових діагнозів і оптимізації процесу лікування [5]. Пандемія COVID-19 істотно прискорила цифрову трансформацію медицини: обмеження фізичного доступу до лікарів та нестача ресурсів зумовили різке зростання попиту на автоматизовані інструменти дистанційної консультації та системи підтримки прийняття клінічних рішень [7]. Саме в цей період HRS продемонстрували свою ефективність у телемедицині, дистанційному моніторингу пацієнтів та підборі індивідуальних стратегій лікування, що підтвердило їхню значущість для сучасної системи охорони здоров'я.

Рекомендаційні системи у медицині реалізуються у низці функціональних доменів, що охоплюють різні аспекти діагностики, лікування, профілактики та реабілітації пацієнтів. Залежно від мети застосування, такі системи можна класифікувати на кілька основних напрямів:

1. Харчування та спосіб життя. HRS, орієнтовані на формування дієтичних рекомендацій, створюють персоналізовані плани харчування, враховуючи індивідуальні характеристики користувача: вік, стать, стан здоров'я, супутні захворювання, а також його харчові вподобання. Для моделювання взаємозв'язків між продуктами харчування, поживними речовинами та обмеженнями часто застосовуються онтологічні

моделі, які забезпечують семантичну сумісність між даними [5].

2. Підбір лікарських засобів. Системи цього типу спрямовані на зменшення кількості помилок при призначенні медикаментів шляхом урахування історії хвороби, результатів аналізів, взаємодій препаратів і супутніх патологій. У практиці використовуються як експертні системи на основі правил (rule-based systems), так і алгоритми колаборативної фільтрації (collaborative filtering), які підбирають терапію за аналогією з подібними клінічними випадками [6]. Наприклад, при лікуванні хронічних захворювань, таких як діабет чи мігрень, ці системи враховують тип патології, рівень ризику та фармакологічну сумісність препаратів, підвищуючи безпечність і ефективність лікування.

3. Геопросторові рекомендаційні системи. Сучасні HRS інтегрують медичні та логістичні параметри, такі як географічна доступність медичних закладів, час прибуття та наявність спеціалістів потрібного профілю. Використання концепції Health Attention Factor (HAF) дає змогу враховувати не лише клінічні характеристики, а й просторово-часовий контекст, формуючи рейтинг оптимальних медичних центрів відповідно до потреб конкретного пацієнта [7].

4. Реабілітація та відновне лікування. У цьому напрямку, HRS застосовуються для створення індивідуальних програм реабілітації, моніторингу прогресу пацієнтів і підвищення їхньої мотивації до виконання рекомендацій. Прикладом є людиноорієнтована система для пацієнтів із ортопедичними захворюваннями, що формує рекомендації на основі клінічних показників і забезпечує пояснюваність результатів [8]. Такі системи підвищують довіру з боку користувачів і сприяють дотриманню призначеного лікувального режиму.

У сучасних наукових дослідженнях спостерігається стійка тенденція до розвитку інтелектуальних медичних систем, у яких поєднуються методи штучного інтелекту, машинного та глибокого навчання для підтримки клінічних рішень. Більшість робіт у цій галузі зосереджена на підвищенні точності діагностики, персоналізації лікування та автоматизації аналізу медичних даних, зокрема за рахунок роботи з неструктурованими джерелами та складними часовими залежностями.

Окремий напрямок досліджень пов'язаний із використанням великих мовних моделей у клінічних інформаційних системах. У роботі [15] запропоновано систему MERA, що інтегрує мовні моделі з підходом Retrieval-Augmented Generation (RAG) для пошуку аналогічних клінічних випадків, формування звітів і підтримки прийняття рішень. Автори наголошують на підвищенні пояснюваності результатів і можливості роботи з неструктурованими даними, водночас

відзначаючи високу обчислювальну складність реалізації. Подібні ідеї простежуються й у дослідженні [17], де великі мовні моделі застосовано в медичних чат-ботах. Поєднання генеративних підходів із механізми доповненого пошуку дозволяє підвищити точність відповідей, однак потребує додаткової адаптації до спеціалізованої медичної термінології.

Поряд із мовними моделями значна кількість робіт присвячена використанню методів машинного навчання для прогнозування клінічних подій. У систематичному огляді [16] проаналізовано застосування алгоритмів Decision Trees, Random Forest та нейронних мереж для прогнозування інфекційних захворювань. Автори показують, що ефективність таких підходів суттєво залежить від повноти та якості даних, а також від можливості стандартизації моделей для різних патологій. Аналогічні обмеження відзначаються й у роботі [19], де ансамблеві методи використовуються для автоматизованого кодування медичних записів і прогнозування повторних госпіталізацій. Хоча отримано покращення точності, модель залишається чутливою до характеристик вхідних даних. У дослідженні [20] акцент зроблено на інтерпретованості рішень за допомогою методів LIME і SHAP, що підвищує прозорість прогнозів, однак водночас обмежує модель через використання вузького набору клінічних ознак.

Окрему групу становлять роботи, спрямовані на аналіз серцево-судинних захворювань. У [18] запропоновано підхід до стратифікації ризику серцевих нападів із використанням методів кластеризації та зменшення розмірності, що дозволяє автоматично виділяти групи пацієнтів за рівнем ризику, але потребує подальшої клінічної валідації. У [22] представлено мультимодальну систему Heart-Net, яка поєднує аналіз МРТ-зображень і сигналів ЕКГ, демонструючи високу точність, проте висуваючи значні вимоги до апаратних ресурсів і обсягу даних. Дослідження [23] доповнює цей напрям, показуючи, що оптимізація гіперпараметрів у моделях SVM, ANN і Random Forest дозволяє підвищити показники якості діагностики, хоча й супроводжується зростанням часу обчислень.

Значну увагу в сучасних роботах приділено графовим і часовим моделям, що дозволяють враховувати складні взаємозв'язки між пацієнтами, захворюваннями та лікуванням. У [24] розроблено графову нейромережу TriGCN, яка моделює взаємодію між хворобами, препаратами та пацієнтами для формування персоналізованих рекомендацій, однак відзначається складністю масштабування. Подальший розвиток цього підходу представлений в [25], де контрастивне графове навчання використовується для підвищення точності рекомендацій лікарських засобів за рахунок кращого узагальнення, що, втім, супроводжується високими обчислювальними витратами. Використання Transformer-архітектур і механізмів часової уваги для

моделювання послідовностей медичних подій у [26] дозволяє підвищити рівень персоналізації рекомендацій, але потребує наявності довгих і якісних часових рядів.

Подібні ідеї простежуються й у роботах, орієнтованих на аналіз електронних медичних записів. У [27] запропоновано модель з механізмом уваги, що поєднує колаборативну фільтрацію та аналіз EHR, демонструючи високу точність рекомендацій за рахунок складної архітектури. У [28] представлено самонавчальну графову модель для прогнозування результатів лікування за умов обмеженої кількості міток, що підвищує практичну цінність підходу, але ускладнює процес навчання. Напівконтрольована GRU-модель, запропонована в [29], враховує часову динаміку клінічних ризиків, однак висуває жорсткі вимоги до точності часових позначок у даних.

Окремо слід відзначити роботи, спрямовані на покращення якості медичних даних і аналіз системних взаємозв'язків. У [30] застосування моделей BERT для виявлення аномалій у медичних записах дозволяє підвищити надійність подальшого аналізу. Подальший розвиток цієї ідеї представлено в [31], де поєднання методів машинного навчання та графових алгоритмів використано для аналізу зв'язків між медичними закладами, що забезпечує кращу масштабованість моделей.

Загалом, проаналізовані дослідження свідчать про поступовий перехід від ізольованих алгоритмічних рішень до комплексних систем, які інтегрують мовні, графові та часові моделі. Особливо перспективним напрямком є поєднання методів глибокого навчання з Explainable AI, що дозволяє не лише підвищити точність прогнозів, а й забезпечити прозорість і довіру до результатів у клінічній практиці.

Глибоке навчання є однією з ключових технологій сучасних рекомендаційних систем, оскільки дозволяє моделювати складні нелінійні залежності між медичними ознаками, анамнестичними даними та клінічними рішеннями. Серед базових архітектур глибокого навчання широко застосовується багатошарова перцептронна мережа, яка належить до класу повнозв'язних нейронних моделей. Її структура забезпечує повну взаємодію між нейронами сусідніх шарів, що створює умови для ефективного виявлення прихованих закономірностей у даних і робить таку модель придатною для задач класифікації, регресії та формування рекомендацій.

Особливою перевагою багатошарових перцептронів є їхня здатність до multi-label класифікації, коли модель формує декілька вихідних рішень одночасно. Це безпосередньо відповідає поставленій задачі визначення переліку лабораторних аналізів на основі сукупності симптомів і факторів ризику пацієнта.

Практична привабливість MLP зумовлена відносно невисокими вимогами до обсягу навчальних даних, простотою процесу навчання та можливістю гнучкої адаптації до різних типів вхідної інформації. Крім того, такі моделі характеризуються високою швидкістю та здатні досягати високої точності за умови коректного підбору архітектури й гіперпараметрів, що робить їх доцільними для використання в прикладних медичних рекомендаційних системах [21].

Аналіз сучасних наукових досліджень показав, що, попри активний розвиток технологій персоналізованої медицини, існуючі HRS мають низку суттєвих обмежень. У таблиці 1 наведено основні недоліки таких систем, виявлені в наукових джерелах, а також можливі напрями їх подолання.

Таблиця 1

Недоліки існуючих медичних рекомендаційних систем (HRS)

Недолік	Медична рекомендаційна система	Можливий напрямок покращення
Обмежена сфера застосування	Більшість HRS орієнтовані на дієтичні або фітнес-рекомендації, але не використовуються у складних клінічних рішеннях (онкологія, кардіологія тощо) [5]	Розробка універсальних HRS для підтримки лікарів загальної практики
Відсутність пояснюваності (Explainable AI)	Алгоритми часто працюють як «чорні скриньки», не надаючи пояснень своїм рішенням, що знижує довіру лікарів [32]	Використання моделей LIME, SHAP, Grad-CAM для пояснення рекомендацій
Складність оцінювання ефективності	Медичні системи потребують перевірки не лише точності, але й безпеки, клінічної доцільності, прийнятності для користувача [33]	Введення клінічних метрик (safety, utility, interpretability) у процес тестування
Проблеми з довірою лікарів і користувачів	Медики не схильні покладатися а рекомендації системи без можливості перевірити логіку її роботи [34]	Включення модуля пояснень та контролю лікаря при затвердженні рішень
Етичні та правові ризики (GDPR, HIPAA)	Недостатній захист персональних медичних даних, ризик порушення конфіденційності [35]	Використання анонімності даних і політик доступу відповідно до GDPR/HIPAA
Відсутність стандартизації даних	Різні формати та відсутність єдиних протоколів взаємодії між HRS [36]	Використання медичних онтологій (SNOMED CT, ICD-10) для узгодження даних

Підсумовуючи результати аналізу, можна зазначити, що основними недоліками сучасних медичних рекомендаційних систем є відсутність пояснюваності алгоритмів, низький рівень інтеграції з клінічними інформаційними системами, недостатній захист персональних даних, а також труднощі з оцінюванням ефективності. Це вказує на необхідність створення нової системи, яка поєднує методи штучного інтелекту, принципи доказової медицини та пояснювані алгоритми (Explainable AI), забезпечуючи при цьому конфіденційність і контроль з боку лікаря.

Саме такі вимоги покладено в основу розроблюваної інтелектуальної системи підтримки прийняття рішень для сімейного лікаря.

## 2. Мета та завдання дослідження

Метою дослідження є підвищення якості формування плану лікування пацієнтів сімейного лікаря шляхом використання запропонованої рекомендаційної системи HRS, яка формує персоналізовані рекомендації на основі симптомів, клінічних даних, анамнезу та доказових медичних протоколів.

Запропонована система передбачає застосування методів штучного інтелекту, машинного навчання та пояснюваної аналітики (Explainable AI), що дає змогу лікарю не лише отримувати обґрунтовані рекомендації, а й інтерпретувати логіку їх формування.

Система також має відповідати міжнародним стандартам GDPR та HIPAA, що гарантує захист і конфіденційність персональних медичних даних пацієнтів.

Для досягнення поставленої мети необхідно вирішити такі задачі:

1. Проаналізувати сучасні підходи до побудови HRS та систем підтримки клінічних рішень (CDSS), визначити їхні переваги та обмеження [6, 7, 9].

2. Спроекувати архітектуру HRS для підтримки прийняття рішень сімейним лікарем із використанням баз знань доказової медицини та електронної медичної картки пацієнта (EMR).

3. Розробити алгоритм формування рекомендацій, який на основі введених лікарем симптомів та історії хвороби автоматично пропонує перелік обстежень і оптимальний план лікування.

4. Забезпечити пояснюваність (Explainable AI) отриманих результатів, щоб лікар міг перевірити логіку прийняття рішень системою.

5. Оцінити ефективність роботи системи за критеріями якості, прозорості, швидкості формування рекомендацій і відповідності вимогам конфіденційності медичних даних.

## 3. Моделювання архітектури та процесів рекомендаційної системи

З метою формалізації процесу проектування рекомендаційної системи підтримки прийняття рішень

для сімейного лікаря на даному етапі було розроблено набір структурно-функціональних діаграм. Вони стають концептуальним описом системи, дозволяючи визначити її архітектуру, основні компоненти, логіку взаємодії між ними, а також послідовність виконання ключових процесів. Діаграми, що розроблені, відображають процес проектування, забезпечуючи узгоджене бачення алгоритмічної логіки, функціональних можливостей і ролей користувачів у межах системи.

Для концептуального опису взаємодії основних компонентів рекомендаційної системи підтримки прийняття рішень сімейного лікаря на етапі проектування побудовано контекстну діаграму, наведену на рис. 1. Вона відображає межі системи, а також її взаємодію з ключовими зовнішніми сутностями, що беруть участь у процесі формування рекомендацій.



Рис. 1. Контекстна діаграма рекомендаційної системи підтримки прийняття рішень

На діаграмі показано, що система отримує вхідні дані від пацієнта у вигляді симптомів і результатів попередніх обстежень, які слугують основою для подальшого аналізу. Джерелами знань для прийняття рішень виступають медичні протоколи та метод формування списку рекомендованих обстежень, що визначають логіку побудови рекомендацій. Додаткову клінічну інформацію система отримує з електронної медичної картки пацієнта, яка містить історію хвороби та попередні дані про стан здоров'я.

Результати роботи рекомендаційної системи передаються до інформаційної системи медичного закладу та системи управління базами даних, де здійснюється збереження сформованих рекомендацій, пояснень до них і звітної інформації. Кінцевим користувачем системи є лікар, який отримує рекомендаційний список обстежень, пояснення до сформованих рішень і підсумковий звіт для підтримки клінічного вибору.

На основі контекстної діаграми розроблено функціональну модель, наведену на рис. 2, яка деталізує внутрішню логіку функціонування рекомендаційної системи та основні етапи обробки клінічної

інформації. Модель відображає послідовність процесів від надходження вхідних даних пацієнта до формування рекомендацій і збереження результатів у медичній інформаційній системі.



Рис. 2. Функціональна модель рекомендаційної системи підтримки прийняття рішень

Центральним механізмом реалізації аналітичних процесів є модулі штучного інтелекту, зокрема AI/LLM-модель, які використовуються на етапах аналізу клінічних даних та формування переліку рекомендованих обстежень. У моделі поєднуються введені симптоми, результати попередніх обстежень і дані з електронної медичної картки з формалізованими медичними протоколами та методами рекомендаційної системи, що дозволяє сформулювати попередній клінічний висновок і обґрунтований список рекомендацій.

Після цього рекомендаційний список, пояснення та супровідна інформація передаються до модуля збереження результатів, де з урахуванням політик зберігання даних здійснюється їх інтеграція в електронну медичну картку пацієнта та інформаційну систему медичного закладу. Підтримку обміну інформацією між функціональними блоками забезпечують база даних і система управління даними.

На рівні управління функціями визначається роль лікаря, який ініціює введення клінічних даних, здійснює перегляд результатів аналізу та приймає рішення щодо затвердження сформованого списку обстежень. Таким чином, функціональна модель рекомендаційної системи підтримки прийняття рішень демонструє структуровану функціональну схему, у якій чітко визначено роль методів штучного інтелекту, інформаційних ресурсів і користувача в процесі підтримки клінічних рішень.

Проектування системи рекомендаційної системи підтримки прийняття рішень

Для опису функціональної взаємодії користувачів із рекомендаційною системою на етапі проектування

побудовано діаграму варіантів використання, наведену на рис. 3. Вона відображає основні сценарії роботи користувачів із системою та їхню роль у процесі формування рекомендацій.

На діаграмі визначено два актори: лікар і адміністратор. Лікар є основним користувачем системи та здійснює авторизацію, введення даних пацієнта і ініціацію формування списку рекомендованих обстежень. У процесі роботи він має можливість перегляду пояснень до рекомендацій, редагування та затвердження сформованого переліку, а також формування підсумкового звіту з можливістю його збереження або інтеграції в електронну медичну картку.

Адміністратор виконує технічні функції, пов'язані з підтримкою роботи системи, зокрема перегляд звітів і управління записами в базі даних. Використані на діаграмі зв'язки відображають залежності між варіантами сценаріїв та послідовність виконання дій, що дозволяє представити узгоджену модель системи з погляду користувачів.

Для відображення послідовності дій під час формування рекомендацій на етапі проектування розроблено діаграму діяльності, наведену на рис. 4. Вона ілюструє взаємодію лікаря, рекомендаційної системи та бази даних у процесі аналізу клінічної інформації та прийняття рішень щодо призначення обстежень.

Процес починається з авторизації лікаря та введення симптомів пацієнта. Отримані дані надходять до рекомендаційної системи, яка доповнює їх інформацією з електронної медичної картки та ініціює запит до бази знань для отримання відповідних клінічних протоколів.

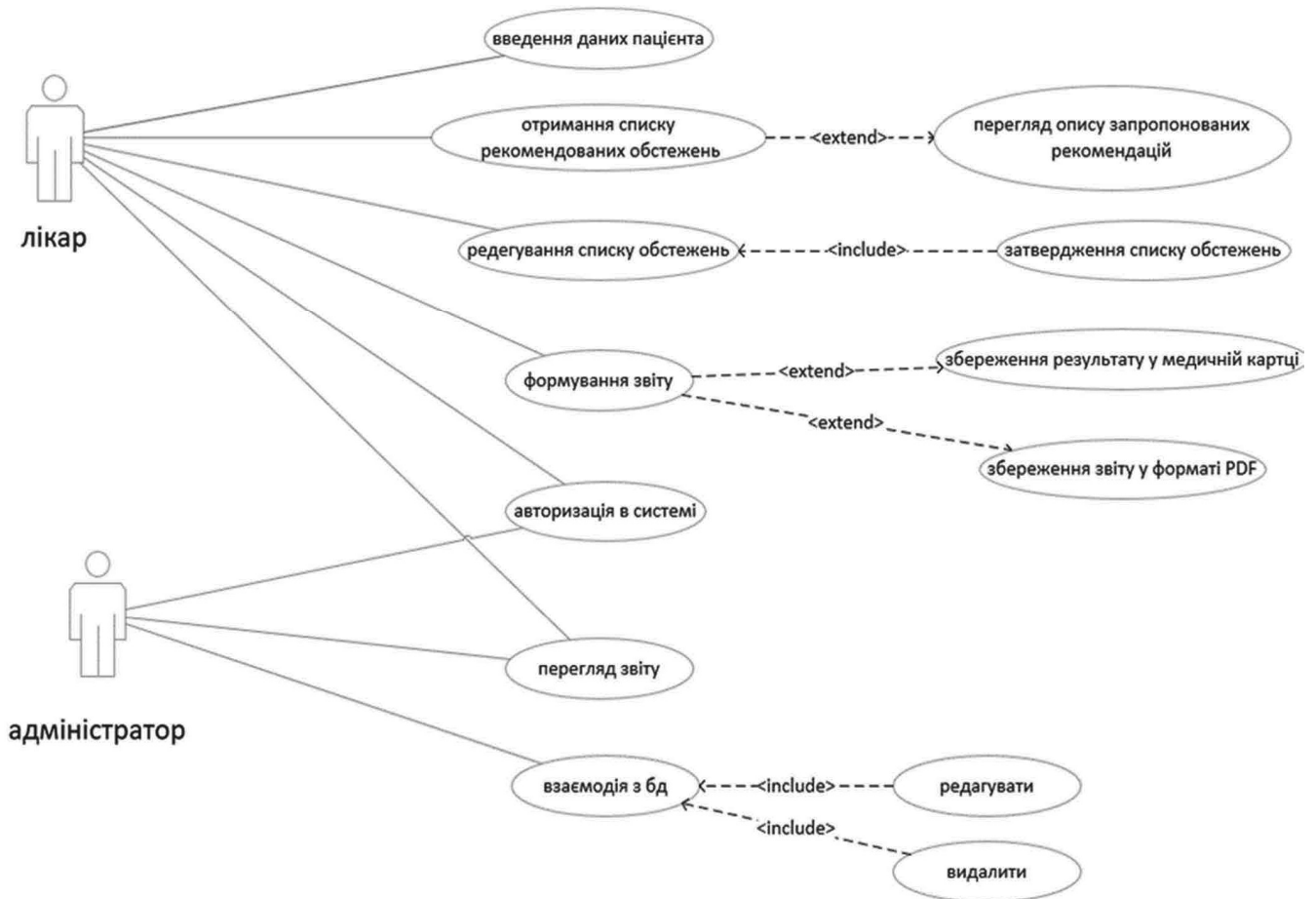


Рис. 3. Діаграма варіантів використання системи підтримки прийняття рішень для сімейного лікаря

На основі зібраної інформації виконується аналітична обробка з використанням методів штучного інтелекту (AI/LLM), у результаті чого формується список рекомендованих обстежень.

Для забезпечення прозорості прийняття рішень система генерує пояснення до сформованих рекомендацій із застосуванням підходів Explainable AI. Лікар здійснює перегляд результатів і, залежно від їхньої відповідності клінічній ситуації, затверджує запропонований перелік або вносить корективи. Після остаточного затвердження результати зберігаються в базі даних, а система формує підсумковий звіт, який може бути інтегрований у електронну медичну картку пацієнта або експортований у зовнішній формат.

Таким чином, діаграма діяльності відображає узгоджену логіку функціонування рекомендаційної системи від введення даних до формування фінального результату та фіксує ключову роль методів штучного інтелекту в підтримці клінічних рішень.

Завершальним етапом проектування є діаграма компонентів, наведена на рис. 5, яка відображає архітектурну структуру рекомендаційної системи та взаємодію її програмних складових. Діаграма фіксує логічний поділ системи на функціональні модулі та показує їхню роль у реалізації процесу формування рекомендацій.

Архітектура системи включає модуль введення клінічних даних, аналітичний модуль на основі методів штучного інтелекту, компонент пояснюваності рішень, модуль генерації звітів, а також засоби зберігання даних і взаємодії з користувачем. Аналітична обробка інформації та формування рекомендацій здійснюються із залученням AI/LLM-моделей, які використовують клінічні протоколи як зовнішнє джерело знань. Модуль пояснюваності забезпечує інтерпретацію результатів, що підвищує прозорість прийняття рішень для лікаря.

Згенеровані рекомендації та звітні матеріали передаються до електронної медичної картки пацієнта та зберігаються в базі даних, що забезпечує цілісність і відтворюваність результатів. Модуль авторизації відповідає за контроль доступу до системи та коректну ідентифікацію користувачів. Таким чином, діаграма компонентів відображає узгоджену взаємодію програмних модулів і логіку обміну даними між ними в межах рекомендаційної системи.

Таким чином, розроблені діаграми формують цілісну проектну модель рекомендаційної системи, що визначає її функціональну логіку, архітектуру та механізми підтримки клінічних рішень із використанням методів штучного інтелекту.

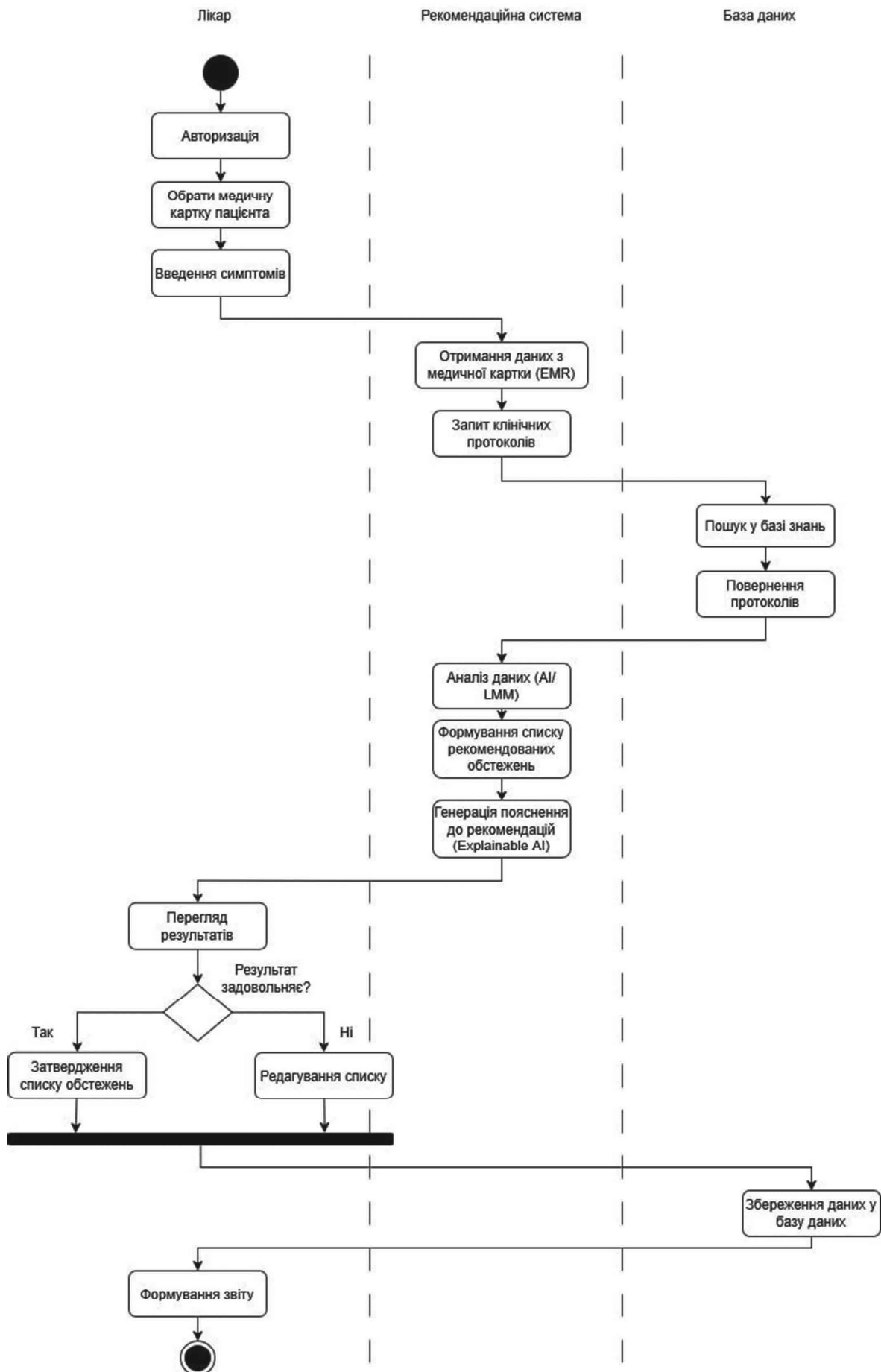


Рис. 4. Діаграма діяльності системи підтримки прийняття рішень для сімейного лікаря

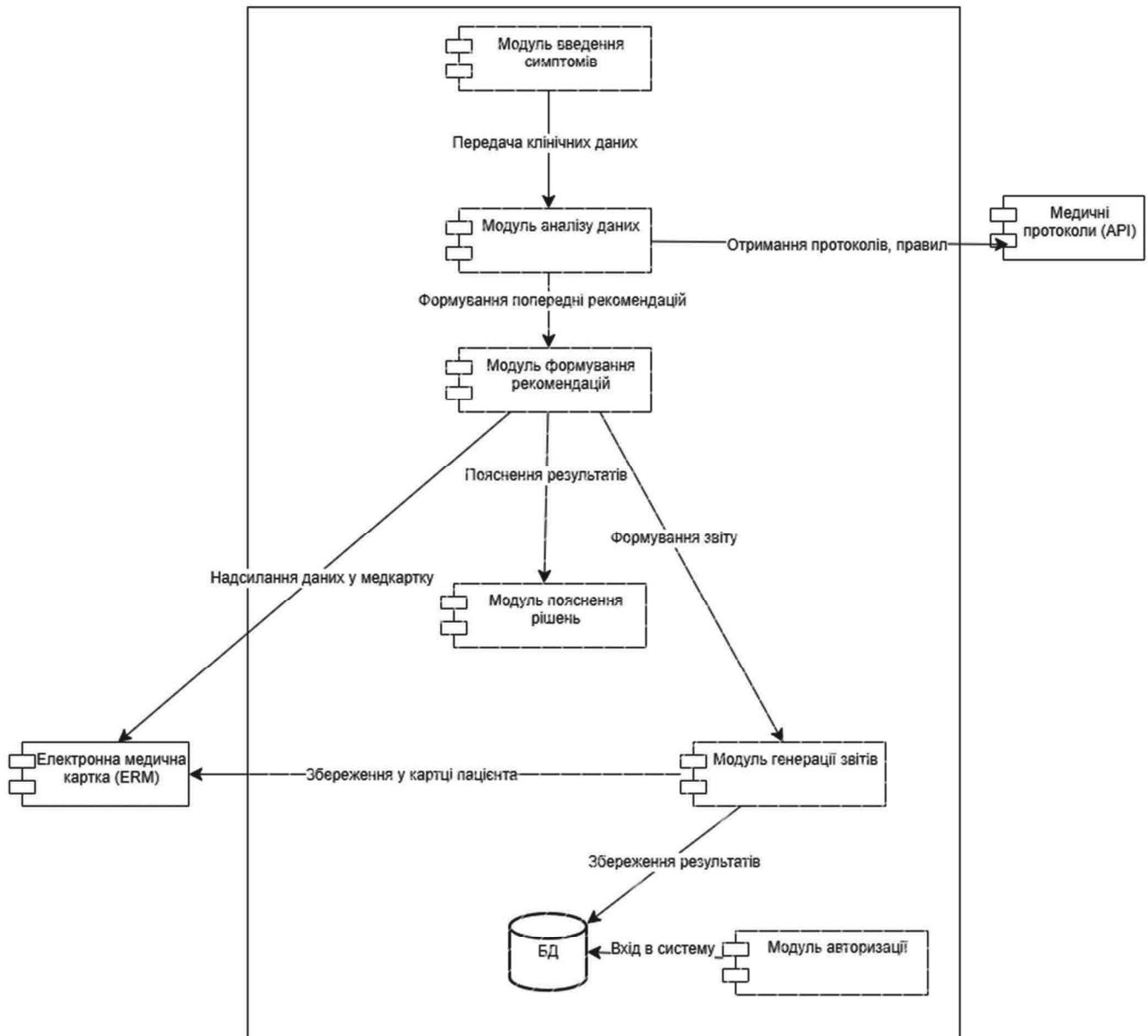


Рис. 5. Діаграма компонентів системи прийняття рішень для сімейного лікаря

### Висновки

У статті проведено аналіз сучасних досліджень у галузі рекомендаційних систем для медицини, зокрема систем підтримки прийняття клінічних рішень на основі методів штучного інтелекту. Показано, що більшість наявних HRS орієнтовані на окремі медичні напрями або спеціалізовані клінічні задачі та недостатньо враховують потреби лікарів первинної ланки, зокрема в частині інтеграції з повсякденною клінічною практикою та забезпечення прозорості рекомендацій.

На основі проведеного аналізу сформовано концепцію рекомендаційної системи підтримки прийняття рішень для сімейного лікаря, орієнтовану на формування переліку лабораторних обстежень з урахуванням симптомів, анамнезу та клінічних протоколів. У дослідженні виконано структурно-функціональне

проекування системи, що включає розробку контекстної діаграми, функціональної моделі, діаграми варіантів використання, діаграми діяльності та діаграми компонентів. Побудовані моделі дозволили формалізувати архітектуру системи, логіку взаємодії її компонентів і дії користувачів, а також визначити роль методів штучного інтелекту та Explainable AI у процесі формування рекомендацій.

Запропонована проектна модель створює основу для подальшої реалізації рекомендаційної системи, орієнтованої на підтримку роботи лікаря, підвищення обґрунтованості клінічних рішень і зменшення кількості необґрунтованих обстежень. Отримані результати можуть бути використані як основа для подальших досліджень, експериментальної реалізації та впровадження інтелектуальної системи підтримки прийняття рішень у практику сімейної медицини.

## Список літератури

- [1] Zhi M., Ding E. L., et al. (2013). The landscape of inappropriate laboratory testing: a 15-year meta-analysis / *PLoS ONE*. — Vol. 8, No. 11. — e78962. — URL: <https://doi.org/10.1371/journal.pone.0078962> (дата звернення: 11.11.2025).
- [2] Dobrowolski T. (2020). What is an Intelligent System? // *ResearchGate*. URL: DOI: 10.13140/RG.2.2.25833.01129 (дата звернення: 11.11.2025).
- [3] Bobadilla J., Ortega F., Hernando A., Gutierrez A. (2013). Recommender Systems Survey // *Knowledge-Based Systems*. — Vol. 46. — P. 109–132. URL: DOI: 10.1016/j.knsys.2013.03.012 (дата звернення: 11.11.2025).
- [4] Zhou H., Xiong F., Chen H. (2023). A Comprehensive Survey of Recommender Systems Based on Deep Learning // *Applied Sciences*. — Vol. 13, No. 20. — Article ID 11378. URL: DOI: 10.3390/app132011378 (дата звернення: 11.11.2025).
- [5] Tran T. N. T., Nguyen P. T. T., Nguyen N. T. et al. (2020). *Recommender systems in the healthcare domain: state-of-the-art and research issues*. *Journal of Intelligent Information Systems*. URL: DOI: 10.1007/s10844-020-00633-6 (дата звернення: 11.11.2025).
- [6] Cai Y., et al. (2022). *Health Recommender Systems Development, Usage, and Evaluation from 2010 to 2022: A Scoping Review*. *International Journal of Environmental Research and Public Health*, 19(22), 15115. URL: DOI: 10.3390/ijerph192215115 (дата звернення: 11.11.2025).
- [7] Torres-Ruiz M., Quintero R., Guzman M., Chui M. (2023). *Healthcare Recommender System Based on Medical Specialties, Patient Profiles, and Geospatial Information*. *Sustainability*, 15(1), 499. URL: DOI: 10.3390/su15010499 (дата звернення: 11.11.2025).
- [8] Singh A., Schooley B., Mobley J., et al. (2025). *Human-Centered Design of a Health Recommender System for Orthopaedic Shoulder Treatment*. *BMC Medical Informatics and Decision Making*. URL: <https://bmcmmedinformdecismak.biomedcentral.com> (дата звернення: 11.11.2025).
- [9] Ammeling J., Aubreville M., et al. (2025). *An Interdisciplinary Perspective on AI-Supported Decision Making in Medicine* // *Technology in Society*. URL: DOI: 10.1016/j.techsoc.2024.102791 (дата звернення: 11.11.2025).
- [10] Frasca M., La Torre D., Pravettoni G., Cutica I. (2024). *Explainable and Interpretable Artificial Intelligence in Medicine: A Systematic Bibliometric Review* // *SpringerLink*. URL: DOI: 10.1007/s44163-024-00114-7 (дата звернення: 11.11.2025).
- [11] Gambetti A., Han Q., Soares C. (2025) *A Survey on Human-Centered Evaluation of Explainable AI Methods in Clinical Decision Support Systems* // *arXiv preprint*. URL: DOI: 10.48550/arXiv.2502.09849 (дата звернення: 11.11.2025).
- [12] Ali O., Abdelbaki W., et al. (2023). *A Systematic Literature Review of Artificial Intelligence in the Healthcare Sector: Benefits, Challenges, Methodologies, and Functionalities* // *Journal of Innovation & Knowledge*. URL: DOI: 10.1016/j.jik.2023.100333 (дата звернення: 11.11.2025).
- [13] Regulation (EU) 2016/679 of the European Parliament and of the Council (General Data Protection Regulation), OJ L 119, 04.05.2016, as amended by OJ L 127, 23.5.2018.
- [14] HIPAA Guide. *HIPAA Compliance Guide*. URL: <https://www.hipaaguide.net/hipaa-compliance-guide/> (дата звернення: 11.11.2025).
- [15] Ibrahim, A.; Khalili, A.; et al. (2025). MERA: Medical Electronic Records Assistant. *Mach. Learn. Knowl. Extr.* 7, 73. URL: <https://doi.org/10.3390/make7030073> (дата звернення: 11.11.2025).
- [16] Santangelo, O.E.; Gentile, V.; et al. (2023). Machine Learning and Prediction of Infectious Diseases: A Systematic Review. *Mach. Learn. Knowl. Extr.* 5, 175–198. URL: <https://doi.org/10.3390/make5010013> (дата звернення: 11.11.2025).
- [17] Bora, A.; Cuayáhuatl, H. (2024). Systematic Analysis of Retrieval-Augmented Generation-Based LLMs for Medical Chatbot Applications. *Mach. Learn. Knowl. Extr.* 6, 2355–2374. URL: <https://doi.org/10.3390/make6040116> (дата звернення: 11.11.2025).
- [18] Gonzalez-Franco, J.D. et al. (2025) Revolutionizing Cardiac Risk Assessment: AI-Powered Patient Segmentation Using Advanced Machine Learning Techniques. *Mach. Learn. Knowl. Extr.* 7, 46. URL: <https://doi.org/10.3390/make7020046> (дата звернення: 11.11.2025).
- [19] Yin, Y.; Shao, Y.; et al. (2025) Machine-Learned Codes from EHR Data Predict Hard Outcomes Better than Human-Assigned ICD Codes. *Mach. Learn. Knowl. Extr.* 7, 36. URL: <https://doi.org/10.3390/make7020036> (дата звернення: 11.11.2025).
- [20] Gao, X., Alam, S., Shi, P. et al. (2023) Interpretable machine learning models for hospital readmission prediction: a two-step extracted regression tree approach. *BMC Med Inform Decis Mak* 23, 104. URL: <https://doi.org/10.1186/s12911-023-02193-5> (дата звернення: 11.11.2025).
- [21] Rawas, S., Tafran, C., ALSaeed, D., Al-Ghreimil, N. (2024). Transforming Healthcare: AI-NLP Fusion Framework for Precision Decision-Making and Personalized Care Optimization in the Era of IoMT. *Computers, Materials & Continua*, 81(3), 4575–4601. URL: <https://doi.org/10.32604/cmcc.2024.055307> (дата звернення: 11.11.2025).
- [22] Aleskait, D. M., et al. (2024). Heart-Net: A Multi-Modal Deep Learning Approach for Diagnosing Cardiovascular Diseases. *Computers, Materials & Continua*, 80(3), 3967–3990. URL: <https://doi.org/10.32604/cmcc.2024.054591> (дата звернення: 11.11.2025).
- [23] Akhund TMNU, Al-Nuwaier WM. (2024), Improving Prediction Efficiency of Machine Learning Models for Cardiovascular Disease in IoST-Based Systems through Hyperparameter Optimization. *Computers, Materials & Continua*, 80(3): 3485–3506. URL: <https://doi.org/10.32604/cmcc.2024.054222> (дата звернення: 11.11.2025).
- [24] Zhou, H., Liao, S., & Guo, F. (2024). TriGCN: Graph Convolution Network Based on Tripartite Graph for Personalized Medicine Recommendation System. *Systems*, 12(10), 398. URL: <https://doi.org/10.3390/systems12100398> (дата звернення: 11.11.2025).
- [25] Li, X., Zhang, Y., Li, X., Wei, H., & Lu, M. (2023). DGCL: Distance-wise and Graph Contrastive Learning for medication recommendation. *Journal of Biomedical Informatics*, 139, 104301. URL: <https://doi.org/10.1016/j.jbi.2023.104301> (дата звернення: 11.11.2025).

- [26] Su Y, Shi Y, Lee W, Cheng L, Guo H. (2022) TAHNet: Time-aware hierarchical dependency network for medication recommendation. *J Biomed Inform.* May; 129: 104069. URL: doi: 10.1016/j.jbi.2022.104069 (дата звернення: 11.11.2025).
- [27] Mi, J., Wang, D., et al. (2024). *ACDNet: Attention-guided Collaborative Decision Network for effective medication recommendation.* / *Journal of Biomedical Informatics, Volume 149, 104570* URL: <https://doi.org/10.1016/j.jbi.2023.104570> (дата звернення: 11.11.2025).
- [28] Liu, Y., Zhang, Z., et al. (2025). GatorCLR: Personalized predictions of patient outcomes on electronic health records using self-supervised contrastive graph representation. *Journal of biomedical informatics*, 168, 104851. URL: <https://doi.org/10.1016/j.jbi.2025.104851> (дата звернення: 11.11.2025).
- [29] Nogues, I. E., Wen, J., et al. (2024). Semi-supervised Double Deep Learning Temporal Risk Prediction (SeDDLer) with Electronic Health Records. *Journal of biomedical informatics*, 157, 104685. URL: <https://doi.org/10.1016/j.jbi.2024.104685> (дата звернення: 11.11.2025).
- [30] Niu, H., Omitaomu, O. A., et al. (2024). EHR-BERT: A BERT-based model for effective anomaly detection in electronic health records. *Journal of biomedical informatics*, 150, 104605. URL: <https://doi.org/10.1016/j.jbi.2024.104605> (дата звернення: 11.11.2025).
- [31] Niu, H., Omitaomu, O. A., et al. (2025). Anomaly Detection in Electronic Health Records Across Hospital Networks: Integrating Machine Learning With Graph Algorithms. *IEEE journal of biomedical and health informatics*, 29(5), 3723–3735. URL: <https://doi.org/10.1109/JBHI.2025.3527752> (дата звернення: 11.11.2025).
- [32] Frasca, M., La Torre, D., Pravettoni, G., & Cutica, I. (2024). *Explainable and Interpretable Artificial Intelligence in Medicine: A Systematic Bibliometric Review. AI and Ethics*, Springer. URL: DOI: 10.1007/s44163-024-00114-7 (дата звернення: 11.11.2025).
- [33] Cai, Y., Liu, J., Xu, J., & Xie, J. (2022). *Health Recommender Systems Development, Usage, and Evaluation from 2010 to 2022: A Scoping Review. International Journal of Environmental Research and Public Health*, 19(22), 15115. URL: DOI: 10.3390/ijerph192215115 (дата звернення: 11.11.2025).
- [34] Ammeling, J., Aubreville, M., et al. (2025). *An Interdisciplinary Perspective on AI-Supported Decision Making in Medicine. Technology in Society*, 75, 102791. URL: DOI: 10.1016/j.techsoc.2024.102791 (дата звернення: 11.11.2025).
- [35] Torres-Ruiz, M., Quintero, R., et al. (2023). *Healthcare Recommender System Based on Medical Specialties, Patient Profiles, and Geospatial Information. Sustainability*, 15(1), 499. URL: DOI: 10.3390/su15010499 (дата звернення: 11.11.2025).
- [36] Sharma, V., Samant, S. S., et al. (2024). An Integrative Framework for Healthcare Recommendation Systems: Leveraging the Linear Discriminant Wolf–Convolutional Neural Network (LDW-CNN) Model. *Diagnostics*, 14(22), 2511. URL: DOI: 10.3390/diagnostics14222511 (дата звернення: 11.11.2025).

Received (Надійшла) 20.01.2026

Accepted for publication (Прийнята до друку) 15.02.2026

Publication date (Дата публікації) 27.03.2026

УДК: 004.021:004.89:004.41

DOI: [https://doi.org/10.30837/bi.2026.1\(104\).04](https://doi.org/10.30837/bi.2026.1(104).04)

М. І. Шапошніков

НТУ «ХПІ», м. Харків, Україна, [nikshaposhnikov01@gmail.com](mailto:nikshaposhnikov01@gmail.com),  
ORCID ID: 0009-0007-8737-2083

## ІНФОРМАЦІЙНА СИСТЕМА ПЛАНУВАННЯ РОЗПОДІЛУ РЕСУРСІВ ДЛЯ ПІДВИЩЕННЯ РЕЙТИНГУ ЗАКЛАДУ ВИЩОЇ ОСВІТИ

У статті розглянуто питання стратегічного управління закладом вищої освіти (ЗВО) в умовах глобальної конкуренції та впливу міжнародних університетських рейтингів, зокрема QS World University Rankings. Метою дослідження є розробка інформаційної системи підтримки управлінських рішень щодо планування ресурсів ЗВО на основі ключових показників ефективності (КПІ), що безпосередньо впливають на рейтингові позиції університету. Запропоновано модель бізнес-процесу у нотатції BPMN, яка інтегрує зовнішні рейтингові механізми QS WUR з внутрішньою системою КПІ структурних підрозділів університету та управлінським циклом PDCA. Розроблено багаторівневу архітектуру інформаційної системи формування рекомендацій, що поєднує оптимізаційні алгоритми та модулі аналітики. Впровадження запропонованої інформаційної системи забезпечує підвищення обґрунтованості управлінських рішень, прозорість планування ресурсів і сприяє зростанню конкурентоспроможності ЗВО в міжнародному освітньому просторі.

ЗАКЛАД ВИЩОЇ ОСВІТИ, КЛЮЧОВІ ПОКАЗНИКИ ЕФЕКТИВНОСТІ, ПЛАНУВАННЯ РЕСУРСІВ, ІНФОРМАЦІЙНА СИСТЕМА, БІЗНЕС-ПРОЦЕС, РОЗПОДІЛ РЕСУРСІВ, УПРАВЛІНСЬКІ РІШЕННЯ

**M. I. Shaposhnikov. An information system for resource allocation planning to increase the ranking of a higher education institution.** The article addresses the strategic management of higher education institutions (HEIs) in the context of global competition and the influence of international university rankings, particularly the QS World University Rankings. The aim of the study is to develop an information system for decision support in resource planning for HEIs, based on key performance indicators (KPIs) that directly affect the university's ranking. A BPMN-based business process model is proposed that integrates external QS WUR ranking mechanisms with the internal KPI system of university structural units and the PDCA management cycle. A multi-layer architecture of a recommendation-oriented information system combining optimization algorithms and analytical modules is developed. The implementation of the proposed information system enhances the justification of managerial decisions, ensures transparency of resource planning, and contributes to increasing the competitiveness of HEIs in the international educational environment.

HIGHER EDUCATION INSTITUTION, KEY PERFORMANCE INDICATORS, RESOURCE PLANNING, INFORMATION SYSTEM, BUSINESS PROCESS, RESOURCE ALLOCATION, MANAGEMENT DECISIONS

### Вступ

Вплив світових рейтингів на стратегічний розвиток закладу вищої освіти (ЗВО) є значним та багатограним. З одного боку, високі позиції в таких рейтингах, як Quacquarelli Symonds World University Rankings (QS WUR), Times Higher Education (THE), Academic Ranking of World Universities (ARWU) зміцнюють репутацію ЗВО, сприяючи залученню талановитої молоді та визнаних вчених з усього світу. З іншого боку, прагнення покращити свої позиції в рейтингах спонукає ЗВО до перегляду та оптимізації своїх внутрішніх бізнес процесів, від дослідницької діяльності до міжнародної співпраці. Це, в свою чергу, вимагає від керівництва ЗВО глибокого стратегічного планування та адаптації до умов глобального освітнього середовища, щоб ефективно планувати розподіл ресурсів та визначити пріоритетні напрямки розвитку для підвищення його рейтингу [1, 2]. Високі рейтингові показники сприяють зміцненню репутації університету, залученню талановитих студентів і науковців та розширенню міжнародного партнерства. Попри те, орієнтація на покращення рейтингових позицій потребує від керівництва ЗВО перегляду внутрішніх бізнес-процесів, удосконалення системи управління ресурсами та впровадження сучасних інформаційних технологій.

Особливої актуальності набуває інтеграція зовнішніх вимог рейтингових агентств із внутрішньою системою КПІ структурних підрозділів університету. Такий підхід дозволяє забезпечити узгодженість стратегічних і операційних цілей та підвищити ефективність управлінських рішень, спрямованих на покращення рейтингових позицій ЗВО на міжнародному рівні, забезпечуючи його сталий розвиток та конкурентоспроможність.

### 1. Огляд літератури

Сучасні процеси глобалізації та цифровізації освіти обумовлюють необхідність удосконалення системи управління ЗВО, що зумовлює актуальність відповідних наукових досліджень. У статті [3] запропоновано програмне забезпечення WURS, розроблене для моделювання рейтингів університетів. WURS можна використовувати для дослідження показників систем рейтингу та розробки стратегії для підвищення в рейтингах. Проте, у програмному забезпеченні не враховуються взаємозв'язки між показниками рейтингу, що впливає на якість моделювання. У роботі [4] розглянуто гібридну систему TRACE-CS, яка поєднує символічну логіку з великою мовною моделлю (LLM) для створення прозорих пояснень щодо рішень

планування розкладу проведення університетських дисциплін.

У роботах [5, 6] представлено порівняння продуктивності та поточних обмежень, що формує цілісне бачення екосистеми рекомендацій на основі LLM. Описано виклики, пов'язані з часом відгуку, достовірністю та справедливістю рекомендацій, що робить огляд важливим для подальших досліджень у цій сфері. Розглянуто систему рекомендацій для вибору дисциплін студентами, яка використовує фреймворк Retrieval-Augmented Generation із GPT-3.5 і GPT-4. Вона аналізує запити природною мовою, генерує ідеальний опис курсу, а потім зіставляє його з каталогом. Система демонструє високу точність, релевантність і прозорість рекомендацій.

Управління університетом потребує сучасної та інтегрованої інформаційної системи для планування доступних ресурсів. Це [7] дослідження показало важливість впровадження ERP-систем у вищій освіті. ERP-системи та технології відіграють дуже активну роль у покращенні ефективності університетів та дозволяють їм ефективно та прозоро керувати нею, а також покращують комунікацію та координацію між організаційними підрозділами всередині організації та з організаціями у зовнішньому середовищі. Важливою рекомендацією для навчальних закладів є перегляд своєї організаційної структури та реінжиніринг.

У дослідженні [8] запропоновано схему планування та розподілу ресурсів PSAS (Peer-dependent Scheduling and Allocation Scheme) з використанням прогностичного навчання для оптимізації розкладу завдань і розподілу ресурсів, що мінімізує застої та максимізує пропускну здатність системи.

У цій статті [9] запропонована система, яка сприятиме підвищенню ефективності управління науковими результатами, покращенню доступу до інформації та підтримці процесу прийняття рішень на основі реальних даних. У роботі розглянуто ключові питання і проблеми, які виникають на шляху до розробки інтегрованої інформаційної системи управління університетом. Проте, ця система не пов'язана з оцінкою впливу діяльності університету на позиції університету у міжнародних рейтингах.

У статті [10] визначені фактори зовнішнього та внутрішнього середовища, які впливають на функціонування університету та якість освіти в екстремальних умовах функціонування, представлені ключові контексти прискореної цифрової трансформації академічного процесу та компенсаційні інструменти антикризового освітнього менеджменту; розкриті ключові складові трансформації освітнього процесу і забезпечення його якості в умовах війни; продемонстровано приклади пошуку балансу ресурсів на основі аналізу індикаторів та показників

організаційної структури освітнього процесу.

У цій статті [11] представлено інноваційний підхід до оптимізації розподілу ресурсів університетської інноваційної та підприємницької освіти через застосування алгоритму колаборативної фільтрації, вдосконаленого Flemingo Optimized Collaborative Filtering Classification (FOCFC).

Дослідження [12] містить практичні рекомендації для керівників ЗВО щодо покращення адміністративних результатів за допомогою ШІ з акцентом на навчання, етичних нормах та розбудові інфраструктури високоякісних даних. У роботі використано методи описової статистики для аналізу центральних тенденцій, а також кореляційний, регресійний аналіз та тести ANOVA для оцінки взаємозв'язків і прогнозного впливу змінних. Результати демонструють, що вищий рівень інтеграції ШІ суттєво корелює з покращенням ефективності процесів прийняття рішень.

Аналіз показав відсутність комплексних інформаційних систем, що поєднують методологію QS WUR, внутрішню систему KPI ЗВО та оптимізаційні моделі планування ресурсів. Системний аналіз показників діяльності та планування ресурсів із використанням сучасних інформаційних технологій здатні суттєво підвищити ефективність управління ЗВО та сприяти досягненню його стратегічних цілей спрямованих на підвищення позицій університету в міжнародному освітньому просторі. Це, своєю чергою, обґрунтовує актуальність розробки інформаційної системи, призначеної для підтримки управлінських рішень керівництва ЗВО.

## 2. Мета дослідження

Метою дослідження є підвищення якості планування ресурсів за допомогою запропонованої інформаційної системи, яка призначена для підтримки управлінських рішень керівництва ЗВО спрямованих на досягнення ключових показників ефективності (KPI), що впливають на позиції університету в QS WUR рейтингу.

## 3. Методологія проведення дослідження

Дослідження та моделювання на основі ретроспективних даних університету та моделей [13], стали основою для створення інформаційної системи управління ресурсами. Запровадження такої системи дозволить реалізувати підтримку прийняття управлінських рішень у сфері стратегічного розвитку ЗВО з орієнтацією на покращення його рейтингових позицій.

Розробка інформаційної системи, яка спрямована на досягнення зазначених цільових показників та базується на моделі оптимізації розподілу ресурсів [14, 15], методів нелінійного програмування та генетичному алгоритмі для формування обґрунтованих

рекомендацій щодо покращення індикаторів QS WUR, моделі бізнес-процесу формування рекомендацій для підвищення показників ЗВО у рейтингу QS WUR, аналітичну оцінку ефективності запропонованих рішень.

Система планується інтегрувати із корпоративною інформаційною системою університету, базами даних і LLM-модулем. Такий підхід забезпечує адаптивність та самонавчання системи в процесі накопичення історичних даних. Модель формування рекомендацій для підвищення показників ЗВО у рейтингу QS WUR є системною основою рекомендаційної системи управління рейтингом університету.

Запропонована модель реалізує замкнений управлінський цикл PDCA та забезпечує ієрархічну агрегацію KPI від рівня кафедр до рівня керівництва університету. Це дозволяє здійснювати моніторинг відхилень фактичних значень показників від цільових і формувати рекомендації щодо оптимального розподілу ресурсів.

Процес оцінювання результативності діяльності закладу вищої освіти відповідно до методології QS WUR представлено у вигляді BPMN-моделі [16], яка відображає взаємодію зовнішніх джерел даних, адміністративних підрозділів університету. На рис. 1 представлено модель бізнес процесу збору, обробки, аналізу показників ефективності діяльності ЗВО, яка складається з кількох пулів і доріжок, що демонструють розподіл ролей між структурними підрозділами університету та зовнішніми джерелами даних. Модель відображає взаємодію зовнішніх і внутрішніх джерел даних, етапи аналітичної обробки, нормалізації показників і прийняття управлінських рішень. У зовнішньому пулі «QS WUR» здійснюється збір та очищення даних із різних джерел (Scopus, Elsevier, глобальні опитування академічної спільноти, роботодавців, випускників тощо). Дані від університету формуються у необхідній формі та завантажуються через портал QS WUR.

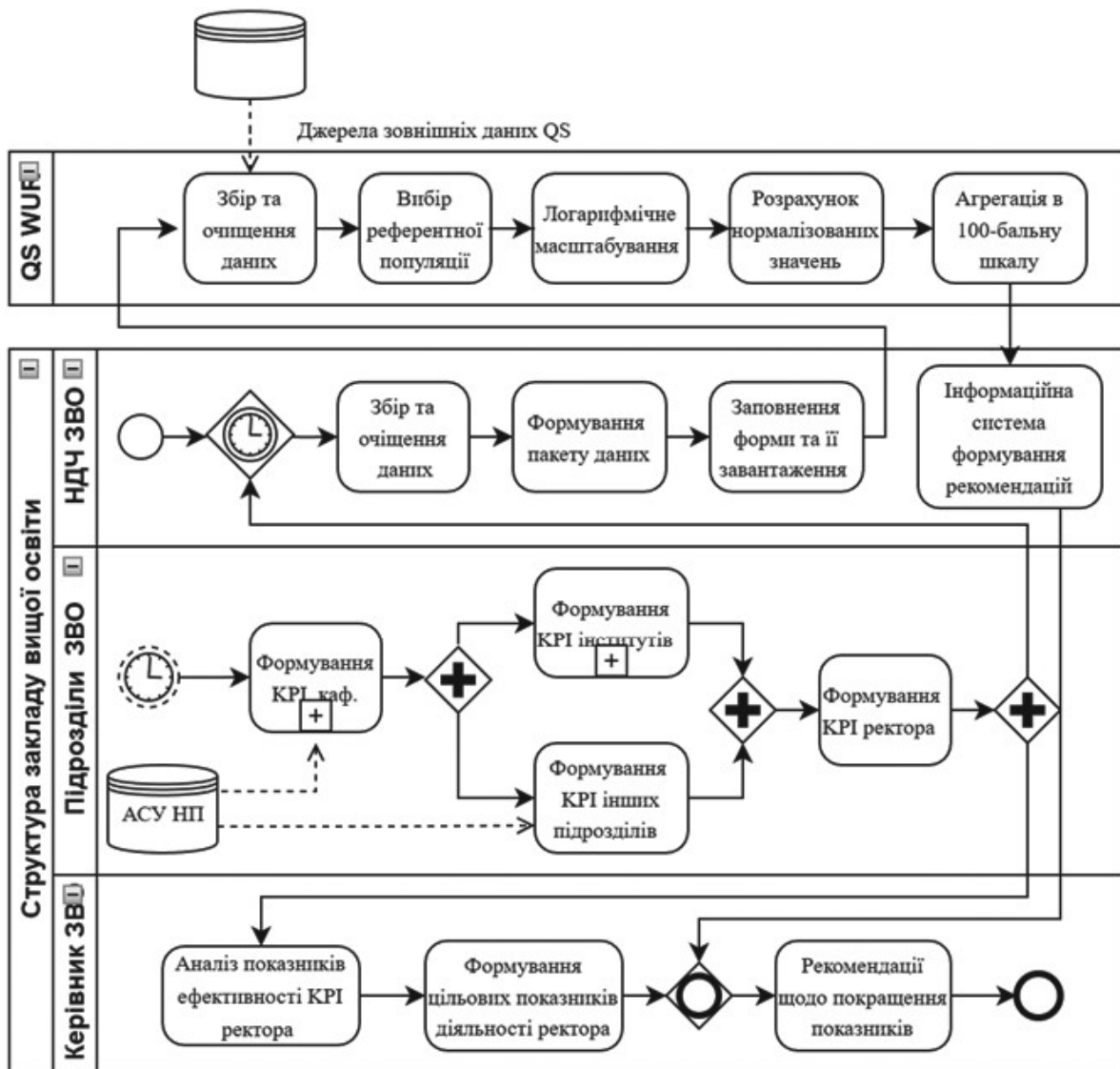


Рис. 1. Модель бізнес процесу збору, обробки, аналізу та вдосконалення показників ефективності діяльності ЗВО

Цей пул моделі відповідає зовнішньому середовищу та відображає етапи обробки інформації рейтинговою агенцією QS.

На цьому рівні реалізується збір та очищення вихідних даних із різних джерел (Scopus, Elsevier, глобальні опитування академічної спільноти, роботодавців, випускників тощо). Очищення передбачає усунення помилок, дублікатів і невідповідностей. Подальші дії включають вибір референтної популяції, що дозволяє порівнювати університети зі схожим профілем і масштабом діяльності. Далі виконується логарифмічне масштабування для показників із високою варіативністю – це забезпечує більшу стабільність і прозорість. Після цього проводиться нормалізація даних, що переводить усі індикатори у шкалу 0–100, а потім здійснюється агрегація у 100-бальну шкалу з урахуванням вагових коефіцієнтів та розраховується загальний бал ЗВО. Отримані результати передаються до інформаційної системи формування рекомендацій, яка надає сценарії та рекомендації щодо покращення показників університету.

Внутрішній пул «Структура закладу вищої освіти» складається з трьох підпулів: науково-дослідна частина (НДЧ), підрозділи ЗВО та керівництво. Вони пов'язані з діяльністю університету. НДЧ ЗВО виконує функції внутрішнього збору даних, необхідних для розрахунку ключових показників. До основних завдань цього етапу належать: отримання статистичних даних від структурних підрозділів, їх попереднє структурування, формування пакетів інформації за категоріями індикаторів QS WUR та внесення відомостей до інформаційної системи формування рекомендацій. У межах цього етапу інформаційна система автоматично формує первинні рекомендації щодо можливих напрямів підвищення показників університету. У цьому пулі виконується збір і очищення внутрішніх даних, зокрема відомостей про публікаційну активність, участь у міжнародних проектах, академічну мобільність, кадровий склад тощо.

Отримані результати завантажуються та зберігаються в інформаційній системі формування рекомендацій у НДЧ, де вони використовуються для подальшого моніторингу і вдосконалення діяльності університету. Сформовані рекомендації надаються керівнику ЗВО для аналізу та прийняття управлінських рішень щодо покращення показників.

У внутрішньому пулі моделі представлені дії НДЧ університету, які забезпечують інформаційну інтеграцію між QS WUR і структурними підрозділами ЗВО. Пул «Підрозділи ЗВО» демонструє, як на рівні окремих структурних одиниць формується внутрішня система KPI. Кожна кафедра надає набір показників (кількість публікацій, цитованість, участь у грантах, академічні зв'язки, впровадження інновацій, тощо). Дані зберігаються в АСУ НП (автоматизованій

системі управління навчальним процесом) і передаються на рівень навчально-наукових інститутів (ННІ). Формування KPI ННІ здійснюється шляхом агрегування показників кафедр, що дозволяє визначити середні значення ефективності та відхилення. Для інших підрозділів (адміністративних, інформаційних, фінансових) визначаються допоміжні показники результативності. На основі узагальнених результатів формуються інтегральні KPI ректора, які відображають стратегічну ефективність управління університетом, узгоджену з вимогами QS WUR та цільовими показниками керівника ЗВО.

Завершальний етап процесу належить керівництву університету, яке здійснює аналітичну інтерпретацію отриманих результатів. На цьому рівні виконується аналіз ефективності KPI ректора та порівняння досягнутих результатів із цільовими орієнтирами та показниками провідних університетів. Далі здійснюється формування цільових показників діяльності ректора на наступний звітний період з урахуванням інформації, сформованої інформаційною системою, щодо покращення показників університету. Рекомендації можуть охоплювати дії, пов'язані із розвитком наукової комунікації, підвищенням публікаційної активності, активізацією міжнародного партнерства, підготовкою кадрів, розвитком цифрової інфраструктури тощо.

Модель має багаторівневу структуру з вертикальними та горизонтальними зв'язками: вертикальні зв'язки забезпечують ієрархічну підпорядкованість, горизонтальні зв'язки відображають дії, які пов'язані із інформаційним обміном між підрозділами та НДЧ, а також між ЗВО та QS WUR. Процес має циклічний характер і реалізує принцип безперервного удосконалення (PDCA) [17], де Plan пов'язано з плануванням KPI, Do відображає збір і завантаження даних, Check реалізує аналіз і порівняння показників, Act реалізує формування рекомендацій і впровадження змін.

Таким чином, модель створює зв'язок між зовнішньою системою QS WUR та внутрішньою структурою ЗВО, що дозволяє адаптувати стратегію розвитку університету до міжнародних стандартів оцінювання. Модель демонструє систему управління ефективністю, яка поєднує зовнішню оцінку (QS) з внутрішньою системою KPI (підрозділи) та управлінським циклом (керівник ЗВО), використовуючи НДЧ. Інформаційні потоки мають ітераційний характер: результати аналізу повертаються на попередні рівні для корекції стратегічних показників, що забезпечує замкнений цикл управління ефективністю (feedback loop).

Таким чином, модель бізнес-процесу збору, обробки, аналізу та вдосконалення показників ефективності діяльності ЗВО відображає цілісний процес управління KPI університету, орієнтований на підвищення його конкурентоспроможності в міжнародних рейтингах.

Важливу роль в цьому процесі відіграє інформаційна система формування рекомендацій керівництву ЗВО для покращення позиції у QS WUR.

Для реалізації цієї системи були розроблені програмні компоненти, представлені на рис. 2. Ці компоненти враховують вимоги до сучасних інструментів розробки інформаційних систем.

Архітектура розробленої системи побудована за принципами багаторівневого розподілення та включає чотири основні рівні: UI Layer, Services Layer, Data Layer та External Systems. Така структура забезпечує модульність, масштабованість і можливість незалежного розвитку окремих компонентів. Розглянемо кожен рівень детальніше.

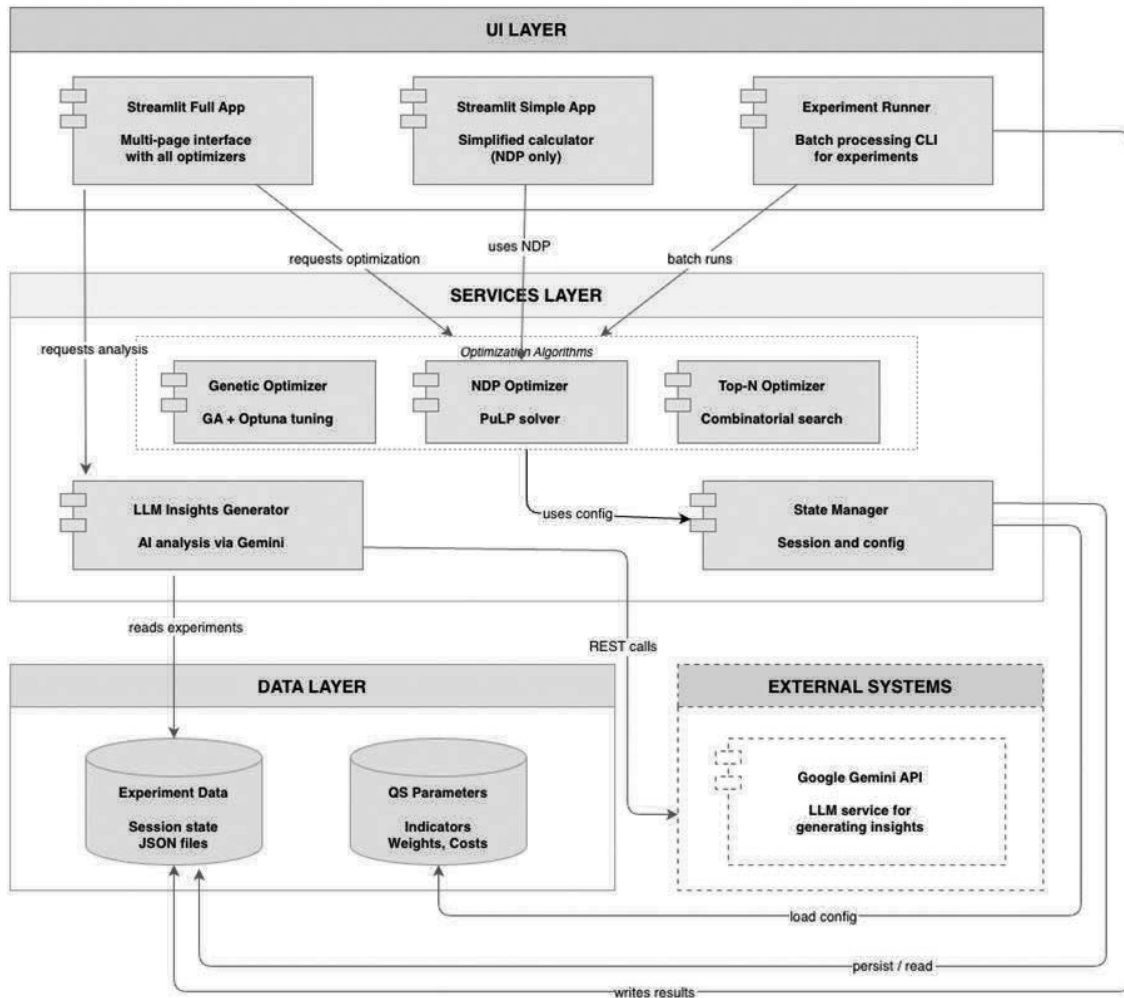


Рис. 2. Архітектура інформаційної системи формування рекомендацій

Рівень UI Layer містить набір інтерфейсних модулів, які забезпечують взаємодію користувача з системою та ініціюють виклики оптимізаційних та аналітичних сервісів. До цього рівня входять: Streamlit Full App це багатосторінковий веб-інтерфейс, який надає доступ до всіх доступних оптимізаторів, можливостей конфігурації та інструментів аналізу. Streamlit Simple App це спрощений калькулятор, орієнтований на використання лише нелінійної оптимізації (NLP). Experiment Runner це CLI-інструмент для проведення пакетних (batch) експериментів без участі користувача. UI Layer надсилає запити оптимізації до сервісного рівня, відображає результати та забезпечує вибір параметрів експериментів.

Рівень Services Layer реалізує основну бізнес-логіку системи та є центральним ядром архітектури. Цей рівень включає дві головні функціональні підсистеми

і складається з декількох незалежних алгоритмічних модулів: Genetic Optimizer – це модуль генетичної оптимізації, що підтримує тонке налаштування параметрів через Optuna. NLP Optimizer / PuLP Solver – це реалізація нелінійної оптимізації на основі пакета PuLP. Top-N Optimizer – це модуль комбінаторного пошуку, який орієнтований на знаходження оптимальних наборів розмірності N.

Ці алгоритмічні модулі отримують конфігурацію завдання від UI Layer або сервісів управління станом і повертають оптимізаційні результати у Data Layer. LLM Insights Generator – це модуль генерації аналітичних висновків за допомогою моделей Gemini. Після завершення оптимізації він може отримувати дані експерименту для глибокого аналізу та текстових інтерпретацій. State Manager – це модуль управління конфігураціями, сесіями та параметрами. Він

зчитує та оновлює налаштування через Data Layer, а також забезпечує узгоджений доступ до конфігурацій для всіх оптимізаторів. Services Layer також здійснює REST-взаємодію з Data Layer для зчитування вхідних параметрів, зберігання результатів і отримання конфігурацій.

Рівень Data Layer також відповідає за зберігання та об'єднує: Experiment Data та QS Parameters. У Experiment Data зберігаються результати оптимізації, проміжні сесійні стани та JSON-файли експериментів. У QS Parameters зберігаються параметри індикаторів, ваги, обмеження, вартісні коефіцієнти та інші дані, що використовуються оптимізаторами. Цей рівень може реалізовуватися як локальні JSON-сховища або як окремий REST-сервіс залежно від конкретної інфраструктури. Дані зберігаються в стандартизованому форматі, що дозволяє виконувати пакетні обчислення та повторювати експерименти.

До рівня External Systems (зовнішні сервіси) належать зовнішні інтелектуальні сервіси, які використовуються для генерації текстових інсайтів та пояснень. Google Gemini API – це LLM, який забезпечує створення пояснень результатів, текстових звітів та аналітичних коментарів. Інші потенційні LLM-сервіси використовуються для розширення функціональності або дублювання моделей. Взаємодія з LLM здійснюється через модуль LLM Insights Generator, який передає дані оптимізації та отримує сформовані аналітичні звіти.

Загальний процес схеми взаємодії виконується наступним чином: UI Layer формує запит оптимізації або запуск експерименту. Services Layer отримує запит, завантажує конфігурацію із State Manager, надсилає параметри до відповідного оптимізатора. Оптимізатор виконує розрахунки і передає результати в Data Layer.

За потреби, LLM Insights Generator зчитує результати оптимізації, відправляє їх у Gemini API та отримує текстові інсайти. UI Layer забирає результати та формує користувацьку візуалізацію.

Запропонована архітектура забезпечує наступні характеристики: модульність – це можливість незалежної заміни або розвитку оптимізаторів. Масштабованість – це підтримка нових UI-рішень або LLM-сервісів. Прозорість експериментів завдяки Data Layer з чіткою структурою JSON-зберігання. гнучку конфігурацію через State Manager. Розширюваність – це додавання нових алгоритмів, API чи аналітичних модулів без зміни ядра системи.

#### 4. Етапи формування рекомендацій

Для створення та формування рекомендацій при плануванні ресурсів ЗВО розроблено інформаційну систему UniRank Optimizer, яка орієнтована на підвищення позицій університету в рейтингу QS WUR. За допомогою інформаційної системи створюються сценарії оптимального розподілу ресурсів між показниками QS WUR для максимізації загального бала університету з урахуванням обмежень ресурсів.

Розглянемо основні етапи створення та формування рекомендацій в системі UniRank Optimizer.

Етап 1. На першому етапі здійснюється налаштування параметрів оптимізації (рис. 3). Співробітники НДЧ через веб-інтерфейс Streamlit вводять вхідні параметри, так як: поточні значення дев'яти показників QS WUR (AR, ER, FSR, CPF, IFR, ISR, IRN, EO, SUS), коефіцієнти важливості у формулі QS Score (сума = 1.0), обмеження на приріст кожного показника за рік, обсяг ресурсів RES для покращення показника на одиницю. Загальний обсяг ресурсів (MAX\_RES) – доступні ресурси для покращення рейтингу.

Бюджет: 100,00 | Поточний бюджет: 100 Resource units

Бюджет встановлено

Поточні значення показників (2025 рік) | Ваги показників (формула QS score) | Максимальні покращення на один рік (Delta) | Вартість покращення (Cost)

Введіть поточні значення показників QS рейтингу

AR - Academic Reputation - Репутація в академічному середовищі	6,50	ER - Employer Reputation - Репутація серед роботодавців	10,60
FSR - Faculty Student Ratio - Співвідношення викладачів до студентів	54,30	CPF - Citations per Faculty - Цитування на викладача	1,30
IFR - International Faculty Ratio - Частка іноземних викладачів	1,70	ISR - International Student Ratio - Частка іноземних студентів	20,10
IRN - International Research Network - Міжнародна дослідницька мережа	11,40	EO - Employment Outcomes - Результати працевлаштування	4,00
SUS - Sustainability - Сталість розвитку	1,60		

Рис. 3. Налаштування параметрів оптимізації

Етап 2. Потрібно обрати режим та метод оптимізації. Користувач обирає режим роботи: (1) повна оптимізація всіх показників, (2) цільова оптимізація обраних показників, N стратегії – це автоматичний пошук найкращих комбінацій. Доступні методи оптимізації: генетичний алгоритм (GA) – це еволюційна оптимізація з Optuna для автоналаштування гіперпараметрів. Підходить для нелінійних задач. Нелінійне програмування (NDP) – це цілочисельне програмування з солвером CBC. Гарантує глобальний оптимум, швидший за GA.

Етап 3: Виконати процес оптимізації. При запуску: State Manager зберігає параметри в QS Parameters

Store, оптимізатор (GA/NDP) формулює задачу: з обмеженнями ресурсів, приросту та максимуму. Результати зберігаються в Experiment Data Store.

Етап 4: Користувач може переглянути отримані результати (рис. 4). Інформаційна система формує звіт з метриками (QS Score, приріст, витрати RES, ефективність, час) та візуалізаціями (порівняльні діаграми, графіки приросту, scatter plot ефективності, розподіл ресурсів, heatmap стратегій, динаміка GA). Детальна інформація представлена у таблиці, яка показує для кожного показника поточне та оптимізоване значення, приріст і витрати.

## Результати топ-N оптимізації (GA)

### Найкраща стратегія

QS Score	Витрати Res	Ефективність	Покращення
<b>11.745</b>	<b>98.5</b>	<b>0.006</b>	<b>5.1%</b>
↑ +0.575	↓ -1.5		

Покращені показники: AR (Academic Reputation), CPF (Citations per Faculty), IRN (International Research Network)

Алгоритм: GA

▼ Детальні значення найкращої стратегії

Показник	Поточне значення	Нове значення	Зміна	Внесок у QS
AR - Academic Reputation - Репутація в академічному середовищі	6.5	7.4	0.9	2.22
ER - Employer Reputation - Репутація серед роботодавців	10.6	10.6	0	1.59
FSR - Faculty Student Ratio - Співвідношення викладачів до студентів	54.3	54.3	0	5.43
CPF - Citations per Faculty - Цитування на викладача	1.3	1.6	0.3	0.32
IFR - International Faculty Ratio - Частка іноземних викладачів	1.7	1.7	0	0.085
ISR - International Student Ratio - Частка іноземних студентів	20.1	20.1	0	1.005
IRN - International Research Network - Міжнародна дослідницька мережа	11.4	16.3	4.9	0.815
EO - Employment Outcomes - Результати працевлаштування	4	4	0	0.2
SUS - Sustainability - Сталість розвитку	1.6	1.6	0	0.08

### Топ-3 стратегії

#	Показники	QS Score	Витрати Res	Алгоритм
1	AR (Academic Reputation), CPF (Citations per Faculty), IRN (International Research Netv	11.745	98.5	GA
2	AR (Academic Reputation), IRN (International Research Network), SUS (Sustainability)	11.73	98.5	GA
3	AR (Academic Reputation), FSR (Faculty Student Ratio), IRN (International Research Net	11.715	100	GA

▼ Статистика

Комбінацій	Максимум	Середнє	Час
<b>84</b>	<b>11.745</b>	<b>11.515</b>	<b>27.2c</b>

Рис. 4. Результати пошуку комбінацій топ-N оптимізації

Етап 5: Сформувані рекомендації за допомогою LLM. LLM Insights Generator трансформує числові результати у практичні рекомендації. Після натискання кнопки «Генерувати AI інсайт», система формує промпт з даними оптимізації та відправляє до Google Gemini API. LLM генерує українською мовою

пріоритетні напрямки покращення, конкретні дії для показників, оцінку ризиків та стратегічні висновки.

Етап 6. Проведення моніторингу та формування звітів. Система забезпечує моніторинг KPI QS через дашборди метрик, порівняння експериментів та експорт звітів у форматах PDF, Excel та JSON (рис. 5).

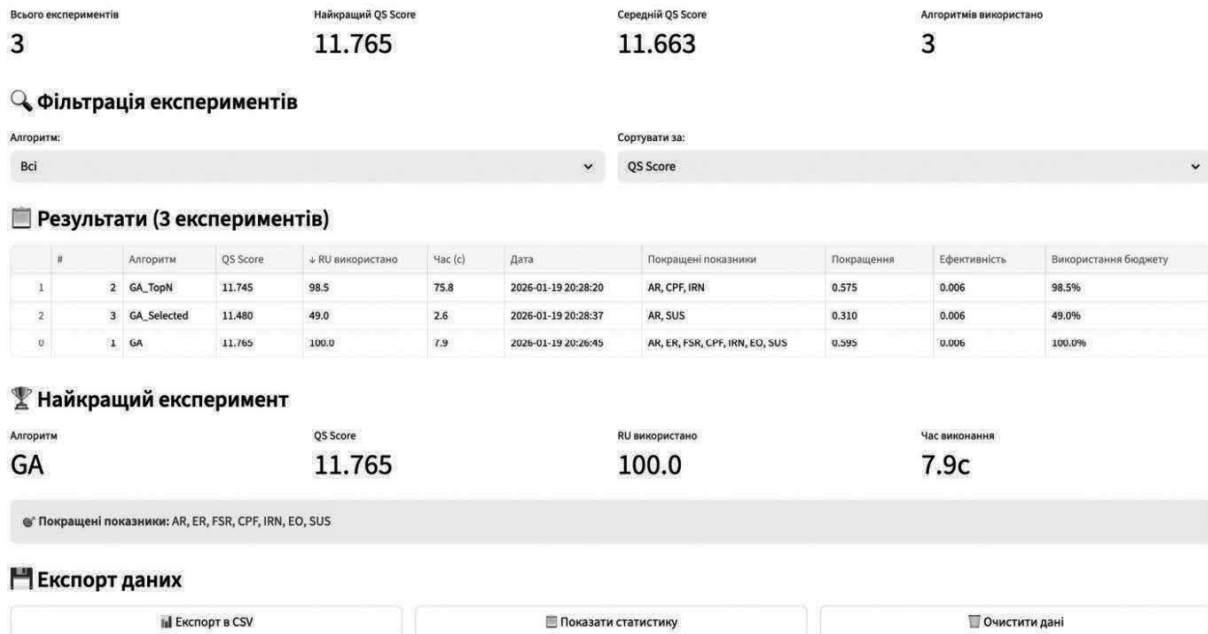


Рис. 5. Результати експериментів

Етап 7. Затвердження стратегій та пріоритетів.

Керівник університету на основі отриманих результатів, AI-рекомендацій та порівняння альтернативних сценаріїв приймає рішення щодо затвердження плану покращення показників, визначає пріоритети інвестування та формує план дій.

Рекомендаційна система забезпечує повний цикл підтримки прийняття рішень від введення даних до рекомендацій. Головні можливості інформаційної системи UniRank Optimizer є реалізація методів оптимізації (GA/NDP), інтеграція з модулем LLM, візуалізація результатів, експорт звітів та модульна архітектура. Система призначена для оперативного планування та управління ЗВО.

### Висновки

У роботі розроблено інформаційну систему формування рекомендацій для планування ресурсів ЗВО, яка орієнтована на підвищення позицій університету в рейтингу QS WUR. Запропонована BPMN-модель забезпечує інтеграцію зовнішніх рейтингових механізмів із внутрішньою системою KPI та створює основу для обґрунтованої підтримки управлінських рішень.

Запропонована інформаційна система ґрунтується на комплексному аналізі даних щодо ринку освітніх послуг, результатів наукової діяльності, рівня

міжнародної активності ЗВО та інших важливих чинників. Розроблена модель бізнес процесу забезпечує інтеграцію механізмів рейтингування QS WUR із внутрішньою системою управління ЗВО, що дозволяє створити ефективний механізм прийняття управлінських рішень на основі KPI. Реалізація інформаційної системи надає керівництву ЗВО рекомендації щодо оптимального планування розподілу ресурсів з метою підвищення рейтингових позицій, з урахуванням особливостей організаційної структури ЗВО, що забезпечує цілеспрямоване та узгоджене управління ресурсами. Отримані результати підтверджують доцільність впровадження сучасних IT-рішень у процеси стратегічного планування та ресурсного менеджменту університету. Запропонована інформаційна система може виступати ефективним інструментом підтримки прийняття управлінських рішень, спрямованих на підвищення міжнародної конкурентоспроможності ЗВО та забезпечення його сталого розвитку.

Практична значущість системи полягає у можливостях: стандартизувати процес формування KPI на всіх рівнях управління університетом; підвищити точність і прозорість внутрішнього моніторингу ефективності; формувати рекомендації для підвищення позицій університету у світових рейтингах.

Таким чином, розроблено інформаційну систему, яка, застосовуючи методи обробки природної

мови на базі LLM, забезпечує інтеграцію механізмів рейтингування QS WUR із внутрішньою системою управління ЗВО. Це дозволяє створити ефективний механізм прийняття управлінських рішень на основі KPI. Впровадження розробленої системи сприяє підвищенню конкурентоспроможності університету, оптимізації планування ресурсів та підвищенню прозорості оцінювання діяльності університету і формуванню стратегій сталого розвитку у глобальному освітньому просторі.

Подальші дослідження доцільно спрямувати на розширення набору аналітичних моделей та інтеграцію прогнозних інструментів.

#### Список літератури

- [1] QS Quacquarelli Symonds. (n.d.). QS World University Rankings. URL: <https://support.qs.com/hc/en-gb/articles/4405955370898-QS-World-University-Rankings> (дата звернення 14.01.2026).
- [2] Times Higher Education. (2025). World University Rankings 2025 Methodology. URL: <https://www.timeshighereducation.com/content/impact-participation> (дата звернення 14.01.2026).
- [3] Siniksaran, E., Satman, M.H. (2020). WURS: a simulation software for university rankings—software review. *Scientometrics* 122, С. 701–717 URL: <https://doi.org/10.1007/s11192-019-03269-8> (дата звернення 14.01.2026).
- [4] Vasileiou, S. L., & Yeoh, W. (2025). TRACE-CS: A Synergistic Approach to Explainable Course Scheduling Using LLMs and Logic. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(28), С. 29706–29708. URL: <https://doi.org/10.1609/aaai.v39i28.35374>
- [5] Z. Zhao et al. (2024) “Recommender Systems in the Era of Large Language Models (LLMs),” in *IEEE Transactions on Knowledge and Data Engineering*, vol. 36, no. 11, pp. 6889–6907, Nov. URL: doi: 10.1109/TKDE.2024.3392335
- [6] Likang Wu, Zhi Zheng, et al. (2024). A Survey on Large Language Models for Recommendation. pp. 1–35. URL: <https://doi.org/10.48550/arXiv.2305.19860> (дата звернення 14.01.2026).
- [7] Alloush & Mahendrawathi (2020). ERP Systems in Higher Education: A Systematic Literature Review. URL: <https://doi.org/10.24089/j.sisfo.2020.01.003><https://doi.org/10.1080/15228053.2020.1846480> (дата звернення 14.01.2026).
- [8] Alfahid, A., Lhioui, C. et al. (2025). Peer-driven task scheduling and resource allocation for enhanced performance in industrial IoT systems. *Sci Rep* 15, URL: <https://doi.org/10.1038/s41598-025-98910-3> (дата звернення 14.01.2026).
- [9] Рябчун, Ю., Гоц, В., Серета Д. (2024). Інформаційна система обліку результатів наукової діяльності закладу вищої освіти. Управління розвитком складних систем, (59), 174–181. URL: <https://doi.org/10.32347/2412-9933.2024.59.174-181> (дата звернення 14.01.2026).
- [10] Spivakovsky, O., Malchykova, D., Yatsenko, V. (2024). Digitalisation of Crisis Management and Resource Balancing for Educational Quality Assurance at a Displaced University *Information Technologies and Learning Tools*, Vol 102, №4. URL: DOI: 10.33407/itlt.v102i4.5761 (дата звернення 14.01.2026).
- [11] Zhang, X., Chen, L., & Li, J. (2023). A Survey on Large Language Models for Recommendation. *arXiv preprint*. URL: <https://doi.org/10.48550/arXiv.2305.19860> (дата звернення 14.01.2026).
- [12] Jiayang, Z., & Ali, H. B. Y. (2025). AI-Enhanced Decision Support Systems for Strategic Higher Education Management: A Framework for Improving Decision-Making Efficiency and Stakeholder Trust. *International Journal of Academic Research in Business and Social Sciences*, 15(10), 1153–1168. URL: <http://dx.doi.org/10.6007/IJARBS/v15-i10/26819>
- [13] Grinchenko M., Shaposhnikov M. (2024). Analysis of higher education institutions’ performance indicators based on QS WORLD university rankings assessment // *Вісник Національного технічного університету «ХПІ»*. Серія: Стратегічне управління, управління портфелями, програмами та проектами. № 2(9). С. 16–26. URL: <https://doi.org/10.20998/2413-3000.2024.9.3> (дата звернення 14.01.2026).
- [14] M. Shaposhnikov, M. Grinchenko. (2025). Mathematical modeling for university resource optimization based on QS WUR indicator // *Bulletin of the National Technical University “KhPI”*. Series: System analysis, control and information technology. — № 2 (14) С. 54–61. URL: DOI: 10.20998/2079-0023.2025.02.07 (дата звернення 14.01.2026).
- [15] M. Shaposhnikov, M. Grinchenko (2025.) Towards information technology – driven nonlinear optimization models for enhancing QS university rankings. *Proceedings of the XV International Scientific and Practical Conference*. Lyon, France. Pp. 39–40 URL: <https://isg-konf.com/current-trends-in-the-development-of-existing-research-and-theories/> (дата звернення 14.01.2026).
- [16] BPMN for research. URL: <https://github.com/camunda/bpmn-for-research> (дата звернення 26.11.2025).
- [17] Li Ning, & Chin Nyuk Sang. (2024). Research on Applying the Deming Cycle (PDCA) Management Method to Promote the Innovative Practical Ability of Art and Design Education Students. *International Journal of Sciences: Basic and Applied Research (IJSBAR)*, 72(1), 303–321. URL: <https://www.gssrr.org/JournalOfBasicAndApplied/article/view/16627> (дата звернення 23.11.2025).

*Received (Надійшла) 06.02.2026*

*Accepted for publication (Прийнята до друку) 01.03.2026*

*Publication date (Дата публікації) 27.03.2026*



V. Bukhanovskiy<sup>1</sup>, N. Ryabova<sup>2</sup>

<sup>1</sup>Kharkiv National University of Radioelectronics, Kharkiv, Ukraine,  
volodymyr.bukhanovskiy@nure.ua, ORCID iD: 0009-0004-9647-4194

<sup>2</sup>Kharkiv National University of Radioelectronics, Kharkiv, Ukraine,  
nataliya.ryabova@nure.ua, ORCID iD: 0000-0002-3608-6163

## FUZZY-ADAPTERS: INTEGRATION OF FUZZY MEMBERSHIP FUNCTIONS FOR VISION TRANSFORMERS, EFFECTIVE TRANSFER LEARNING

This paper introduces Fuzzy-Adapter, a PEFT method for efficiently adapting Vision Transformers (ViT) under resource constraints. The approach replaces standard ReLU activations with trainable Gaussian membership functions to improve uncertainty modelling. Experimental results on the CIFAR-10 dataset using `deit_tiny` demonstrate that Fuzzy-Adapter achieves 92.6% accuracy, significantly outperforming the ReLU-based adapter (90.2%) and the baseline (85.7%). The findings suggest that integrating neuro-fuzzy logic into bottleneck architecture enhances model adaptability and performance in data-limited environments.

VISION TRANSFORMER; PARAMETER-EFFICIENT FINE-TUNING; FUZZY LOGIC; ADAPTER; IMAGE CLASSIFICATION

**В.О. Бухановський, Н.В. Рябова. Fuzzy-adapters: Інтеграція нечітких функцій належності для ефективного трансферного навчання Vision Transformers.** У цій роботі представлено Fuzzy-Adapter — це метод PEFT, розроблений для ефективного адаптації Vision Transformers (ViT) в умовах обмежених ресурсів. Підхід передбачає заміну стандартних активацій ReLU на навчальні Гауссові функції належності для покращення моделювання невизначеності. Експериментальні результати на наборі даних CIFAR-10 з використанням моделі `deit_tiny` демонструють, що Fuzzy-Adapter досягає точності 92,6%, що значно перевищує показники адаптера на основі ReLU (90,2%) та базового методу (85,7%). Отримані висновки свідчать про те, що інтеграція нейро-нечіткої логіки в архітектуру bottleneck підвищує адаптивність та продуктивність моделі в умовах дефіциту даних.

VISION TRANSFORMER; PARAMETER-EFFICIENT FINE-TUNING; НЕЧІТКА ЛОГІКА; АДАПТЕР; КЛАСИФІКАЦІЯ ЗОБРАЖЕНЬ

### Introduction

Vision Transformer (ViT) architectures have become the standard in many computer vision tasks. It demonstrated exceptional performance by modelling global data dependencies. However, this success comes at a cost: modern ViT models are massive and training them from scratch requires large-scale datasets and significant computational resources [1].

The field of Parameter-Efficient Fine-Tuning (PEFT) emerged as a branch of transfer learning. One of the most successful approaches in this domain is the use of adapters. Adapters are small modules inserted into the layers of a pre-trained model. During adaptation to a new task, the weights of the backbone model remain “frozen”, while only the parameters of these compact modules are trained. This approach reduces the number of trainable parameters and computational costs significantly [2].

A key element of popular Bottleneck adapters is the non-linear activation function applied between the down-projection and up-projection layers. Traditionally, standard functions such as ReLU are chosen [3]. However, these functions operate with crisp, binary, or semi-linear transformations. We hypothesise that such logic limits the adapter’s ability to model the uncertainty and imprecision inherent in new datasets.

In this work, we explore a hybrid neuro-fuzzy approach by replacing standard activations with fuzzy membership functions. We introduce Fuzzy-Adapters, where the activation is a trainable Gaussian membership function [4].

This allows the adapter not merely to apply fixed non-linearity, but to learn a flexible, adaptive transformation that better suits the specifics of the target task.

Our contribution lies in the practical implementation and validation of this idea. We integrated Fuzzy-Adapters into the “`deit_tiny`” ViT model and conducted experiments on a reduced CIFAR-10 dataset to simulate a resource-constrained scenario. We demonstrate that our approach provides a statistically significant improvement in accuracy over standard ReLU adapters and the baseline of training only the classifier.

### 1. Analysis of last achievements and publications

Classification is a fundamental task in computer vision, assigning an image to a specific category from a predefined set. However, modern state-of-the-art models, such as Vision Transformers, require massive datasets to achieve high accuracy from scratch. To mitigate this, transfer learning techniques have emerged. It allows the transfer of knowledge obtained from one set of tasks or domains to another.

To understand the proposed approach, it is necessary to consider three key pillars underpinning our research: the Vision Transformers architecture, the adapter-based training methodology, and the fundamental principles of neuro-fuzzy systems.

Vision Transformer architecture is a successor to the successful Transformer architecture (initially developed for natural language processing) in the field of computer

vision. Unlike Convolutional Neural Networks (CNNs), which process information hierarchically through local receptive fields [5], ViT models an image as a sequence of tokens. The operating principle involves dividing the input image into fixed, non-overlapping patches. Each patch is linearly projected into a vector space (embedding), forming a sequence of “tokens”. Positional information is added to these tokens via positional encoding. In classification tasks, a special [CLS] token is typically added to the sequence, and its final representation is used to predict the image class. The core of architecture is the self-attention mechanism [6]. This mechanism enables the model to assess the importance of each token relative to all others in the sequence and effectively capture global relationships between different (even spatially distant) regions of the image. The capability for global contextual modelling ensures high ViT performance, especially when training on large datasets.

Transfer learning is a technique that enables the adaptation of knowledge acquired from one set of tasks or domains to another [7]. This implies that a model trained using vast resources, including data, computational power, time, and cost, which have been made open-source, can be fine-tuned and reused in new settings by the broader engineering community at a significantly lower price than the initial resource requirements. This represents a significant step forward toward the democratisation of large models and, more broadly, artificial intelligence.

Traditionally, three primary approaches define the transfer learning landscape: pre-trained models as classifiers, fine-tuning, and feature extraction [8, p. 254].

The fine-tuning approach involves taking a pre-trained model (e.g., one trained on ImageNet), replacing the final classification layers, and retraining the entire network (or a portion of it) on the new dataset.

The Feature Extraction approach also utilises a pre-trained network but keeps the backbone weights “frozen”. Only the new classification layers are trained on the target dataset.

While full fine-tuning often yields better accuracy, the immense size of modern models makes it computationally expensive and data-intensive. This challenge has driven the development of Parameter-Efficient Fine-Tuning (PEFT) methods that adapt pre-trained models with a minimal number of updated parameters. Among these methods, adapters have emerged as a fundamental and efficient solution.

Adapters are among the most fundamental and efficient PEFT methods. These are small, supplementary modules integrated into the layers of a frozen backbone model. During training, only the weights of these adapters and, usually, the final classification layer are updated, while millions (or billions) of parameters of the main model remain unchanged.

In this work, we focus on the typical Bottleneck Adapter architecture (Figure 1). Such a module is embedded inside a Transformer block (after the feed-forward or attention layer). It includes a fully connected layer that compresses the input representation into a lower-dimensional space, a non-linear function application layer, and a fully connected layer that projects the representation back to its initial dimension.

The entire module is connected via a residual connection. The key advantage is that the number of trainable parameters is significantly smaller than the total number of parameters in a layer. Although there are other popular PEFT methods, such as Low-Rank Adaptation (LoRA) [9], IA3 [10], Prefix Tuning [11], or Prompt Tuning [12], Bottleneck adapters provide a reliable and well-researched basis for integrating new mechanisms.

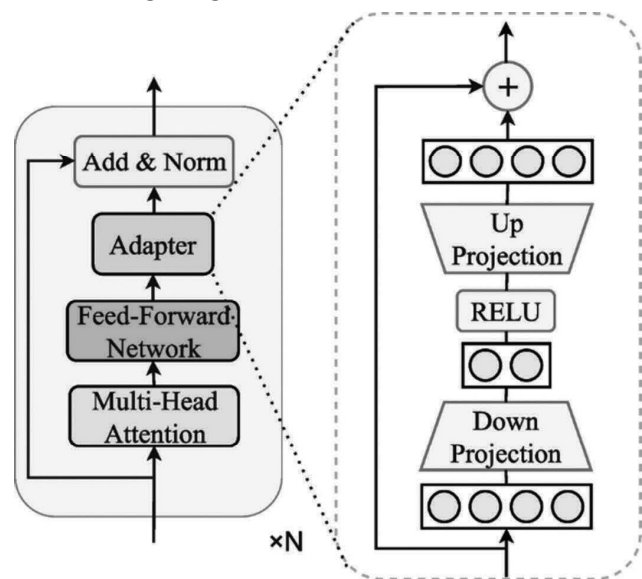


Fig. 1. Bottleneck Adapter architecture

The research gap we address lies in non-linear activation within the adapter. Standard approaches use “crisp” functions that apply complex, semi-linear transformations. We propose turning to neuro-fuzzy systems.

Neuro-fuzzy logic represents a hybrid intelligent system architecture that synergises the learning capabilities of artificial neural networks with the reasoning power of fuzzy logic. While artificial neural networks excel at pattern recognition, data adaptation, and learning from vast datasets, they often function as “black boxes” with crisp internal representations. Conversely, fuzzy logic provides a framework for modelling uncertainty, imprecision, and ambiguity, allowing for reasoning that approximates human cognition. The integration of the two paradigms results in systems capable of learning from data while operating with interpretable, rule-based knowledge structures

The central concept of fuzzy logic is the membership function. Unlike classical Boolean logic, where an element either belongs to a set (value 1) or not (value 0), fuzzy logic allows an element to belong to a set with a certain “degree of membership” – a real number in the range [0, 1].

In the context of hybrid neuro-fuzzy systems, these functions play a role analogous to activation functions in deep learning (such as ReLU or Sigmoid) but offer the additional benefit of interpretability. They transform crisp input values (such as pixel intensities or feature vectors) into fuzzy degrees of membership, enabling the system to process inherently incomplete or noisy information.

There are many forms of such functions: triangular, trapezoidal, Gaussian, etc (Figure 2). Among the various available membership functions, we specifically select the Gaussian Membership Function (GMF) for integration into the adapter architecture. This choice is driven by several factors that make GMF particularly suitable for gradient-based optimisation in neural networks.

First and foremost is the property of smoothness and differentiability. Unlike triangular or trapezoidal functions, which have sharp corners and non-differentiable points, the Gaussian function is smooth and continuously differentiable across its entire domain. This is a mandatory requirement for modern deep learning, as it ensures gradients can flow uninterrupted during backpropagation, enabling stable weight updates.

Secondly, the Gaussian function offers specific trainable parameters: centre and width. They directly correspond to interpretable properties of the data distribution. By treating these as learnable parameters rather than fixed hyperparameters, the neural network gains additional degrees of freedom. This allows the adapter to independently tune the activation shape to match the specific uncertainty profile of the new dataset.

Finally, the Gaussian function naturally models uncertainty. While a standard ReLU activation applies a hard threshold (zeroing out all negative values), a Gaussian activation provides a soft window of attention around a specific centre. This enables the model to assign high activation only to features that fall within a learned range of relevance, effectively filtering out noise and irrelevant data.

While integrating fuzzy membership functions as activation mechanisms offers significant advantages in terms of flexibility and uncertainty modelling, it introduces specific challenges compared to standard activations.

The first thing to mention is computational Complexity. Standard activation functions are computationally inexpensive and require only simple thresholding. In contrast, fuzzy functions, particularly Gaussian MFs, involve exponential calculations. This increases the floating-point operations (FLOPs) needed for both the forward pass and gradient calculation during backpropagation.

The second trade-off is optimisation difficulty. In a standard adapter, the activation function is usually fixed. In the proposed neuro-fuzzy approach, the parameters of the membership function (e.g., centre and width) are trainable. This increases the complexity of the optimisation landscape. If not carefully initialised, parameters may vanish or explode, leading to numerical instability or gradients that hinder convergence. By making the activation

function parametric, the model's total number of trainable parameters increases slightly. But in the context of Bottleneck adapters, this minimal increase requires the optimiser to find a delicate balance between the weights of the linear projections and the shape of the activation function.

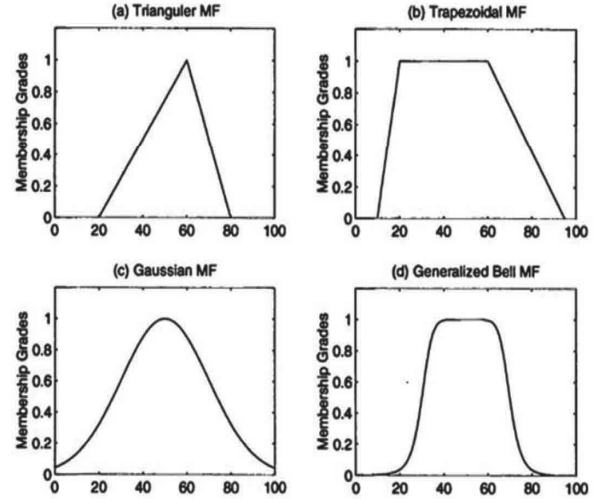


Fig. 2. Membership function types: a) triangular, b) trapezoidal, c) gaussian, d) generalised bell [4]

Despite these limitations, we hypothesise that replacing the standard ReLU activation in the Bottleneck adapter with a trainable Gaussian membership function will enable the adapter module to learn a more flexible, smoother transformation. Instead of a fixed non-linearity, the parameters of the activation function itself (its centre and width) become part of the optimisation process. This provides the model with additional degrees of freedom for a more refined adaptation of internal representations to the specifics of the new dataset.

Our work builds on the standard Bottleneck adapter architecture. The standard module consists of a dimensionality-reducing layer, a non-linear activation function, and a dimensionality-restoring layer with a residual connection.

Our key contribution is to replace the standard “crisp” activation with a trainable fuzzy membership function. In this work, we focused on the Gaussian Membership Function (GMF). We implemented this function as a separate module. Mathematically, it is described by the formula (1).

$$f(x, m, \sigma) = e^{-\frac{1}{2} \left( \frac{x-m}{\sigma+\varepsilon} \right)^2}, \quad (1)$$

where  $\varepsilon$  – a small constant for numerical stability;  $m$  – centre;  $\sigma$  – width.

$m$  and  $\sigma$  are not fixed hyperparameters, but trainable parameters. This enables the neural network to independently optimise the shape of the activation function for each adapter during training.

In practice, in addition to the main Bottleneck layers described above, the input tensor undergoes normalisation

before the down-projection layer, followed by GMF, drop-out, the up-projection layer, and finally the residual connection [13].

## 2. Materials and methods

Selecting an appropriate backbone architecture involves a trade-off between representational power and computational efficiency. Standard Vision Transformers (ViT-Base), comprising approximately 86 million parameters, use global self-attention with quadratic complexity and typically require massive datasets (e.g., JFT-300M) to surpass Convolutional Neural Networks (CNNs) due to their lack of inherent inductive biases.

An alternative modern architecture is the Swin Transformer [14], which introduces hierarchical feature representations and shifted windowing mechanisms to achieve linear complexity with respect to image size. While Swin Transformers excel in dense prediction tasks, even their lightest variant, Swin-T (Tiny), contains approximately 29 million parameters.

In contrast, the Data-efficient Image Transformer (DeiT) utilises a specialised distillation token to learn effectively from teacher ConvNets on smaller datasets like ImageNet-1k [15]. The “deit\_tiny” variant selected for our experiments contains only about 5 million parameters – roughly 17% of the size of Swin-T and 6% of ViT-Base. This extreme compactness makes it uniquely suited to the resource-constrained environment of our research, enabling rapid iteration on the reduced CIFAR-10 subset without the significant computational overhead of Swin or standard ViT architectures.

For the base architecture, we selected the “deit\_tiny\_patch16\_224” model. This characteristic makes it particularly suitable for scenarios where access to large-scale data for fine-tuning is limited, aligning with our experimental constraint of using a reduced CIFAR-10 [16] subset (10% of the original data). Additionally, the “tiny” variant is highly compact, reducing computational resource requirements and accelerating the experimentation process while maintaining competitive performance.

The integration process was architected through a custom wrapper class “ViTWithAdapter”, designed to encapsulate the pre-trained “deit\_tiny” backbone. First, to preserve the learned feature representations, all parameters of the base ViT model were “frozen” by disabling gradient computation for all parameters. Second, a single instance of the Fuzzy-Adapter was sequentially injected into the architecture, positioned to process the output of each Transformer block. Next, the original classification head was discarded and replaced with a new linear layer mapping the model’s embedding dimension of 192 to an output dimension of 10, corresponding to the CIFAR-10 classes. Consequently, the optimisation process was strictly isolated to the adapter modules and the new head. This configuration resulted in a highly efficient setup with 156 718 trainable parameters against a total of 5 681 134, representing merely 2.76% of the model’s capacity.

Data preprocessing and augmentation pipelines were implemented utilising the “torchvision.transforms” module from the PyTorch ecosystem – a critical preprocessing step involved addressing the resolution disparity between the input data and the model architecture. Since the CIFAR-10 dataset consists of 32 32-pixel images, and the pre-trained “deit\_tiny\_patch16\_224” model was designed for 224 224 inputs, the pipeline included upsampling transformations to match the required dimensions.

To mitigate the risk of overfitting, we integrated standard data augmentation techniques alongside statistical normalisation. Furthermore, to strictly enforce the experimental constraint of limited data availability, the training process did not utilise the full dataset. Instead, we selected indices at random to create a reduced training subset containing only 10% of the original CIFAR-10 dataset (5,000 images). The test dataset was used in its entirety.

To evaluate the effectiveness of our approach, we compared three model configurations:

1. ViT Head-Only – the baseline model, where only the new classification head is trained;
2. ViT with ReLU Adapter – the control model using an identical Bottleneck Adapter architecture but with standard ReLU activation;
3. ViT with Fuzzy Adapter (Gaussian) – our proposed model with activation based on the Gaussian membership function.

The implementation utilised the PyTorch framework as the primary engine for tensor computations, model construction (torch.nn), and optimisation (torch.optim). The torchvision library was employed to manage the CIFAR-10 dataset and to execute image processing pipelines, specifically utilising torchvision.transforms for data augmentation and normalisation. To access the pre-trained “deit\_tiny\_patch16\_224” model, we integrated the timm (PyTorch Image Models) library. NumPy was used for high-performance numerical operations and array handling, particularly for managing data indices when creating the reduced dataset subsets. Matplotlib was used to visualise the training dynamics, including loss and accuracy curves, and to display prediction samples. Finally, to ensure the reliability and reproducibility of our results, the random module was used to set a fixed random seed across the environment, and the os module managed file system interactions for saving model checkpoints. There are snippets of GMF implementation in Figure 3 and the Fuzzy Adapter implementation in Figure 4.

```
class GaussianMembershipFunction(nn.Module):
    def __init__(self, mu=0.0, sigma=1.0):
        super(GaussianMembershipFunction, self).__init__()
        self.mu = nn.Parameter(torch.tensor(mu, dtype=torch.float))
        self.sigma = nn.Parameter(torch.tensor(sigma, dtype=torch.float))

    def forward(self, x):
        eps = 1e-6
        return torch.exp(-((x - self.mu) ** 2) / (2 * self.sigma ** 2 + eps))
```

Fig. 3. Snippet of GMF implementation

```

class BottleneckAdapter(nn.Module):
    def __init__(self, input_dim, bottleneck_dim, dropout=0.1, init_scale=1e-3, mu=0.0, sigma=1.0, activation_type='gaussian'):
        super(BottleneckAdapter, self).__init__()

        self.down_proj = nn.Linear(input_dim, bottleneck_dim)

        if activation_type == 'gaussian':
            self.activation = GaussianMembershipFunction(mu=mu, sigma=sigma)
        elif activation_type == 'relu':
            self.activation = nn.ReLU()

        self.up_proj = nn.Linear(bottleneck_dim, input_dim)
        self.dropout = nn.Dropout(dropout)
        self.layer_norm = nn.LayerNorm(input_dim)

        self.alpha = nn.Parameter(torch.tensor(0.1, dtype=torch.float))
        with torch.no_grad():
            self.down_proj.weight.data.normal_(mean=0.0, std=init_scale)
            self.down_proj.bias.data.zero_()
            self.up_proj.weight.data.normal_(mean=0.0, std=init_scale)
            self.up_proj.bias.data.zero_()

    def forward(self, x):
        residual = x
        x = self.layer_norm(x)
        x = self.down_proj(x)
        x = self.activation(x)
        x = self.dropout(x)
        x = self.up_proj(x)

        return residual + self.alpha * x

```

Fig. 4. Snippet of Fuzzy Adapter implementation

The training process was identical for all models, containing adapters to ensure a fair comparison. Standard Cross-Entropy Loss was used, given that this is a multi-class classification task [17]. The Adam optimiser was selected [18]. The model was trained for five epochs. At each epoch, loss and accuracy were calculated on the training set, followed by full validation on the test set. The key performance metric was test-set accuracy, as the data distribution was balanced [19].

The key result of the study is the accuracy comparison of the three approaches: (1) ViT Head-Only, (2) ViT with

ReLU Adapter, and (3) ViT with Fuzzy Adapter. The final accuracy scores on the test dataset are summarised in Table 1. A visualisation of the predictions for our proposed variant is shown in Figure 5.

Table 1

Comparison of model accuracy

Model	Accuracy, %
ViT Head-Only	85,7
ViT with ReLu Adapter	90.2
<b>ViT with Fuzzy Adapter</b>	<b>92.6</b>

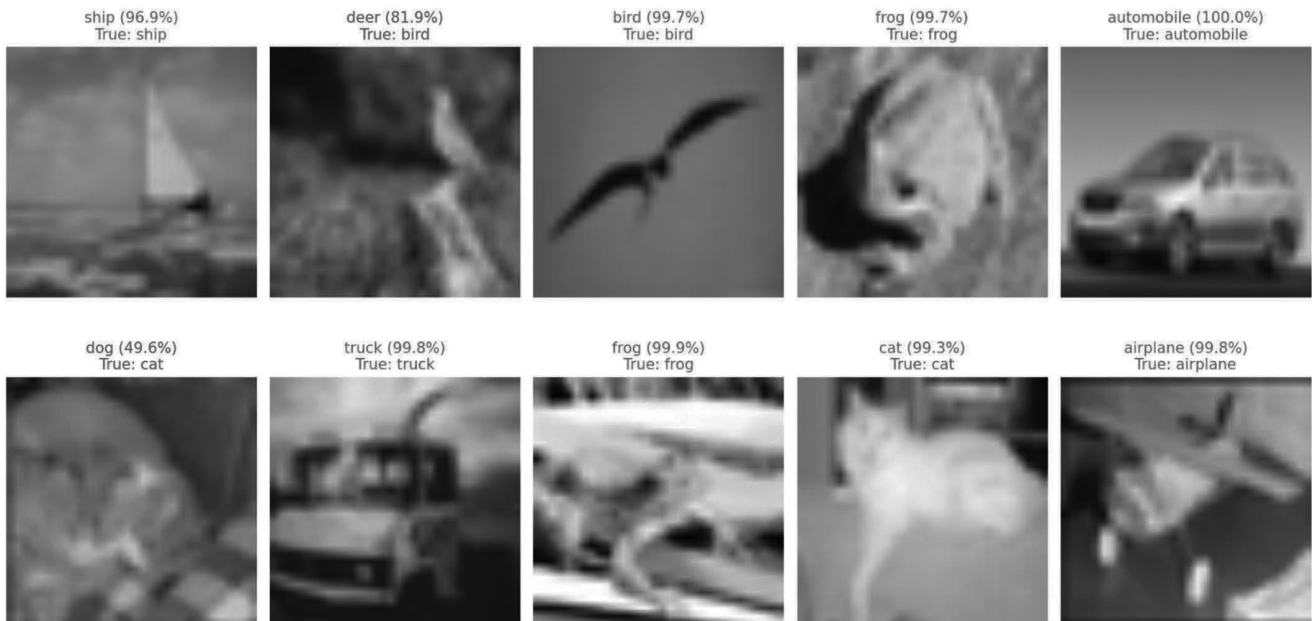


Fig. 5. Prediction visualisation

The results clearly demonstrate the significant advantages of using adapters for transfer learning compared to training only the classification head. The ViT Head-Only model yielded the lowest accuracy (85.7%). Integrating standard adapters with ReLU activation increased accuracy by 4.5%, to 90.2%. Our proposed model, which utilises a trainable Gaussian membership function, achieved an accuracy of 92.6%. This outperforms the standard ReLU adapter by 2.4%, confirming the effectiveness of integrating fuzzy activation functions for more flexible model adaptation.

### 3. Conclusion and perspectives of further development

The obtained results confirm the effectiveness of parameter-efficient fine-tuning (PEFT) but, more importantly, highlight the significant impact of the choice of activation function within adapters.

First, the significant gap of 4.5% between ViT Head-Only (85.7%) and ViT with ReLU Adapter (90.2%) confirms that training only the classification head is insufficient, whereas adding adapters allows the model to adapt significantly better to a new data domain.

Second, the key result of this study is that ViT with Fuzzy Adapter (92.6%) outperforms ViT with ReLU Adapter (90.2%). We hypothesise that this is due to differences in activation functions. This provides the model with significantly greater flexibility. The ability to model uncertainty and learn smoother transformations is likely the reason for the superior adaptation to the specifics of the new dataset. Despite the promising results, this work has limitations. The study was conducted on a single reduced dataset (CIFAR-10), using one base model (“deit\_tiny”) and a single fuzzy function, and evaluating only one metric: accuracy.

The obtained results show several promising directions for further research:

- investigation of other membership functions;
- integration into other PEFT architectures;
- analysis of hyperparameters and other ViT models (e.g. LaViT [20]);
- utilisation of other data sources and more specific domains.

In this paper, we investigated the effectiveness of integrating fuzzy logic into parameter-efficient fine-tuning methods for ViT. We introduced Fuzzy-Adapters — a modification of Bottleneck adapters where a trainable Gaussian membership function replaces the standard ReLU activation function.

Experiments conducted on a reduced CIFAR-10 dataset using the “deit\_tiny” model demonstrated the superiority of the proposed approach. Our ViT with Fuzzy Adapter model achieved the highest accuracy of 92.6%. This outperforms both the baseline approach, which trains only the classification head (85.7%), and the standard ViT with ReLU Adapter (90.2%).

The results suggest that providing adapters with greater flexibility through parameterised, smooth activation functions derived from fuzzy logic enables the model to adapt more effectively to the features of a new domain. So, this

study confirms that hybrid neuro-fuzzy approaches are a promising direction for the further development of parameter-efficient fine-tuning methods.

### References

- [1] An image is worth 16x16 words: transformers for image recognition at scale / A. Dosovitskiy et al. 2021. URL: <https://doi.org/10.48550/arXiv.2010.11929>.
- [2] Parameter-Efficient fine-tuning methods for pretrained language models: a critical review and assessment / L. Xu et al. 2023. URL: <https://doi.org/10.48550/arXiv.2312.12148>.
- [3] Fuzzy logic membership function. URL: <https://research-hubs.com/post/engineering/fuzzy-system/fuzzy-membership-function.html>.
- [4] Parameter-Efficient transfer learning for NLP / N. Houlsby et al. 2019. URL: <https://doi.org/10.48550/arXiv.1902.00751>.
- [5] O’Shea K., Nash R. An introduction to convolutional neural networks. 2015. URL: <https://doi.org/10.48550/arXiv.1511.08458>.
- [6] Attention is all you need / A. Vaswani et al. 2023. URL: <https://doi.org/10.48550/arXiv.1706.03762>.
- [7] Bukhanovskiy V., Ryabova N. Transfer learning methods in computer vision. Сучасні проблеми і досягнення в галузі радіотехніки, телекомунікацій та інформаційних технологій. Запоріжжя, Україна, 2024. P. 323–324.
- [8] Azunre P. Transfer learning for natural language processing. Manning Publications Co. LLC, 2021.
- [9] LoRA: low-rank adaptation of large language models / E. J. Hu et al. 2021. URL: <https://doi.org/10.48550/arXiv.2106.09685>.
- [10] Few-Shot Parameter-Efficient Fine-Tuning is Better and Cheaper than In-Context Learning / H. Liu et al. 2022. URL: <https://doi.org/10.48550/arXiv.2205.05638>.
- [11] Li X. L., Liang P. Prefix-Tuning: optimizing continuous prompts for generation. 2021. URL: <https://doi.org/10.48550/arXiv.2101.00190>.
- [12] Lester B., Al-Rfou R., Constant N. The power of scale for parameter-efficient prompt tuning. 2021. URL: <https://doi.org/10.48550/arXiv.2104.08691>.
- [13] Deep residual learning for image recognition / K. He et al. 2015. URL: <https://doi.org/10.48550/arXiv.1512.03385>.
- [14] Swin transformer: hierarchical vision transformer using shifted windows / Z. Liu et al. 2021. URL: <https://doi.org/10.48550/arXiv.2103.14030>.
- [15] Training data-efficient image transformers & distillation through attention / H. Touvron et al. 2021. URL: <https://doi.org/10.48550/arXiv.2012.12877>.
- [16] Krizhevsky A., Nair V., Hinton G. CIFAR-10 and CIFAR-100 Datasets. URL: <https://www.cs.toronto.edu/~kriz/cifar.html>.
- [17] Mao A., Mohri M., Zhong Y. Cross-Entropy loss functions: theoretical analysis and applications. 2023. URL: <https://doi.org/10.48550/arXiv.2304.07288>.
- [18] Kingma D. P., Ba J. Adam: a method for stochastic optimization. 2017. URL: <https://doi.org/10.48550/arXiv.1412.6980>.
- [19] Google for Developers. Classification: accuracy, recall, precision, and related metrics. URL: <https://developers.google.com/machine-learning/crash-course/classification/accuracy-precision-recall>.
- [20] You only need less attention at each stage in vision transformers / S. Zhang et al. 2024. URL: <https://doi.org/10.48550/arXiv.2406.00427>.

Received (Надійшла) 09.02.2026

Accepted for publication (Прийнята до друку) 01.03.2026

Publication date (Дата публікації) 27.03.2026

УДК 004.9

DOI: [https://doi.org/10.30837/bi.2026.1\(104\).06](https://doi.org/10.30837/bi.2026.1(104).06)А. С. Федоричко<sup>1</sup>, В. А. Висоцька<sup>2</sup>, Р. О. Линник<sup>3</sup><sup>1</sup>НУ «Львівська політехніка», м. Львів, Україна, andrii.fedorychko.sa.2022@lpnu.ua,  
ORCID iD: 0009-0005-1593-3481<sup>2</sup>НУ «Львівська політехніка», м. Львів, Україна, victoria.a.vysotska@lpnu.ua,  
ORCID iD: 0000-0001-6417-3689<sup>3</sup>НУ «Львівська політехніка», м. Львів, Україна, roman.o.lynnik@lpnu.ua  
ORCID iD: 0009-0007-0948-4338

## ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ КЛАСИФІКАЦІЇ ТЕКСТІВ ДЛЯ ВИЯВЛЕННЯ ОЗНАК ПТСР НА ОСНОВІ ТЕХНОЛОГІЙ ГЛИБИННОГО НАВЧАННЯ ТА NLP

Стаття присвячена дослідженню ефективності сучасних архітектур обробки природної мови у завданні автоматизованого виявлення ознак посттравматичного стресового розладу (ПТСР) за текстовими даними. В основі роботи лежить порівняльний аналіз моделі distilbert-base-uncased для оригінальних англійських текстів та xlm-roberta-base для даних, адаптованих українською мовою за допомогою машинного перекладу через API Google Translate. Для проведення експериментів сформовано репрезентативний набір даних обсягом 4822 записи, що включає тексти з верифікованими ознаками ПТСР, контрольні тексти інших психологічних діагнозів та нейтральний контент, що дозволило оцінити здатність моделей до диференціальної діагностики. Експериментальні результати продемонстрували високу ефективність класифікації: англійська модель досягла точності 0.90 при F1-score 0.88–0.91, тоді як україномовна версія показала результат 0.85 Accuracy, зберігши при цьому високий рівень повноти (Recall) значенням 0.88 для цільового класу. Доведено, що незважаючи на незначну втрату точності, зумовлену морфологічною складністю української мови та похибками автоматичного перекладу, використання мультимовних трансформерів є життєздатним підходом для створення систем попереднього скринінгу психічного стану в українському цифровому середовищі.

ПОСТТРАВМАТИЧНИЙ СТРЕСОВИЙ РОЗЛАД, ГЛИБИННЕ НАВЧАННЯ, ОБРОБКА ПРИРОДНОЇ  
МОВИ, МАШИННИЙ ПЕРЕКЛАД, КЛАСИФІКАЦІЯ ТЕКСТІВ, ТРАНСФОРМЕРНІ АРХІТЕКТУРИ,  
МУЛЬТИМОВНІ МОДЕЛІ

**A. S. Fedorychko, V. A. Vysotska, R. O. Lynnyk. Information technology for text classification of PTSD signs detection based on deep learning and NLP.** The article is devoted to the study of the effectiveness of modern natural language processing architectures in the task of automated detection of signs of post-traumatic stress disorder (PTSD) from text data. The work is based on a comparative analysis of the distilbert-base-uncased model for original English-language texts and the xlm-roberta-base model for data adapted into Ukrainian via machine translation via the Google Translate API. To conduct the experiments, a representative dataset of 4822 records was formed, including texts with verified signs of PTSD, control texts for other psychological diagnoses, and neutral content, which allowed evaluation of the models' ability to perform differential diagnosis. Experimental results demonstrated high classification efficiency: the English-language model reached an accuracy of 0.90 with an F1-score of 0.88–0.91, while the Ukrainian-language version showed an accuracy of 0.85, maintaining a high level of Recall at 0.88 for the target class. It has been shown that, despite a slight loss in accuracy due to the morphological complexity of the Ukrainian language and automatic translation errors, the use of multilingual transformers is a viable approach for creating preliminary mental state screening systems in the Ukrainian digital environment.

POST-TRAUMATIC STRESS DISORDER, DEEP LEARNING, NATURAL LANGUAGE PROCESSING,  
MACHINE TRANSLATION, TEXT CLASSIFICATION, TRANSFORMER ARCHITECTURES, MULTILINGUAL  
MODELS

### Вступ

Посттравматичний стресовий розлад (ПТСР) є важким психічним станом, що виникає внаслідок переживання або спостереження травматичних подій, таких як загроза життю чи здоров'ю [1]. Останніми роками спостерігається стрімке зростання кількості досліджень, присвячених застосуванню методів машинного навчання (ML) для прогнозування та визначення цього розладу [2]. Така тенденція значною мірою зумовлена впливом пандемії COVID-19, яка стала глобальною травмою та спричинила сплеск стресових станів у всьому світі [1]. Попри технічний прогрес, у науковій спільноті залишаються занепокоєння щодо клінічної достовірності та здатності отриманих ML-моделей до генералізації, що гальмує їхнє впровадження практикуючими фахівцями [2].

Традиційна діагностика ПТСР через клінічні інтерв'ю часто супроводжується суб'єктивністю та ризиком помилкового визначення розладу як депресії, що суттєво погіршує прогнози лікування [1]. Стигматизація психічних проблем та небажання пацієнтів відкрито обговорювати симптоми також перешкоджають своєчасній допомозі [4]. У цьому контексті обробка природної мови (NLP) пропонує потужні інструменти для аналізу цифрових наративів: постів у соціальних мережах, стенограм терапії та онлайн-форумів з метою виявлення лінгвістичних патернів, характерних для ПТСР [4]. Текстові дані, як цифрові біомаркери, мають суттєві переваги, адже вони безпечні для пацієнтів, потребують мінімальних витрат на збір і безпосередньо відображають емоційний стан людини через зміст мовлення [1].

Експериментальні дослідження з використанням аналізу настрою та глибинного навчання демонструють високу ефективність у скринінгу, досягаючи точності понад 80% на даних, отриманих з інтерв'ю [1]. Оповіді пацієнтів, зібрані за допомогою опитувальників, що симулюють клінічні протоколи DSM-5, виявилися надзвичайно ефективними для швидкої ідентифікації симптомів, особливо у вразливих групах населення [3]. Проте систематичні огляди підкреслюють, що для успішного впровадження таких систем виявлені ML-прогнозування повинні чітко узгоджуватися з існуючим теоретичним розумінням природи розладу [2].

Моделі, які побудовані на основі структури Transformer, зокрема модель BERT, демонструють значну перевагу у завданні класифікації симптомів та поведінкових індикаторів порівняно з класичними алгоритмами машинного навчання, такими як SVM або KNN [3]. Поєднання передових технік NLP, включаючи синтаксичний парсинг, тематичне моделювання та глибинне навчання, дозволяє створювати перспективні системи скринінгу, що забезпечують об'єктивність процесу діагностики в цифровому середовищі [4]. Таким чином, розробка інтегрованих моделей, що об'єднують емпіричні результати ML із фундаментальними теоретичними знаннями, є ключовим кроком до покращення доступності та точності психіатричної допомоги [2].

### 1. Постановка проблеми

Сучасні нейромережеві моделі, що базуються на архітектурі Transformer, стали ключовим етапом еволюції систем машинного перекладу (МП), практично повністю витіснивши застарілі статистичні та правило-орієнтовані підходи [5]. Аналіз роботи провідних систем, таких як DeepL, Google Translate та GPT, демонструє їхню високу ефективність у технічних та публіцистичних матеріалах, проте виявляє значні труднощі при перекладі художніх творів, де часто втрачаються стилістичні відтінки та іронія [5]. Експериментальні дослідження підкреслюють, що попри значну економію часу, якість автоматизованого перекладу все ще потребує контролю та обов'язкового пост-редагування людиною [5], [6]. Недостатній рівень якості адаптації спеціалізованих текстів є серйозним бар'єром, оскільки системи МП часто демонструють кращий результат у звичному для них домені, ніж у вузьких професійних сферах [5].

Технічна реалізація систем діагностики ПТСР ускладнюється проблемою незбалансованості даних, яка виникає через складність збору реальних клінічних інтерв'ю [7]. Для подолання цього дефіциту перспективним є використання великих мовних моделей (LLM) для збільшення обсягу тренувальних даних шляхом синтезації текстових транскрипцій та

перефразування наявних тренувальних зразків [7]. Водночас порівняльний аналіз архітектур свідчить, що спеціалізовані моделі, адаптовані до домену ментального здоров'я, значно перевершують загальні трансформерні моделі у завданні детекції симптомів, особливо у випадках із супутньою депресією [8].

Додатковими викликами при впровадженні таких систем залишаються питання мовної складності, конфіденційності даних та вартості інтеграції рішень у існуючі екосистеми [9]. Таким чином, актуальним завданням є розробка україномовної системи скринінгу, яка б поєднувала швидкість сучасного МП із високою точністю спеціалізованих нейромережевих архітектур при збереженні професійного редакторського контролю [6].

### 2. Векторизація тексту та трансформерні моделі

Сучасний етап розвитку технологій обробки природної мови базується на переході від класичних методів вкладання слів, таких як Word2Vec, GloVe чи FastText, до складніших контекстно-залежних архітектур [10]. Революційним проривом у цій галузі стала поява моделі BERT (Bidirectional Encoder Representations from Transformers), яка за рахунок двонаправленого аналізу тексту дозволила значно глибше розуміти семантику та контекст повідомлень [11]. Завдяки здатності до крос-лінгвального перенесення знань та ефективного виконання широкого спектру завдань, починаючи від токенизації до аналізу тональності, архітектура BERT стала стандартом для побудови інтелектуальних систем розуміння мови [11].

Ефективність застосування трансформерних моделей, зокрема RoBERTa та GPT, у задачах емоційного та семантичного аналізу сьогодні підтверджується через порівняльний аналіз стандартних метрик класифікації: точності, повноти та F1-міри [13]. Такий підхід дозволяє визначити найбільш доцільні сценарії використання нейромережевих архітектур для автоматизованої обробки складних лінгвістичних патернів [13]. Водночас для практичного впровадження у високонавантажені системи критично важливою є оптимізація обчислювальних ресурсів. Модель DistilBERT пропонує ефективне вирішення цього питання, зберігаючи високу точність оригінального BERT при значному підвищенні швидкості обробки тексту [12]. Її здатність до розуміння контексту в поєднанні з можливістю тонкого налаштування (fine-tuning) робить її ідеальним інструментом для класифікації та вилучення ключової інформації [12].

Окремим пріоритетним напрямом є розвиток моделей для української мови. Використання зібраних спільноту корпусів текстів дозволяє не лише застосовувати готові рішення, а й тренувати власні моделі, адаптовані до конкретних предметних областей [10]. Синергія ембедінг моделей, можливостей швидких

архітектур енкодерів, систем та якісних мовних даних створює необхідний технологічний фундамент для аналізу специфічного контенту, зокрема у сферах, де важлива емоційна точність та глибоке розуміння намірів автора [12].

### 3. Огляд сучасних методів аналізу текстів у сфері ментального здоров'я

Використання лінгвістичних особливостей у описах власного досвіду пацієнтів відкриває широкі можливості для автоматизованого скринінгу ПТСР, оскільки вибір слів безпосередньо відображає стан психічного здоров'я людини [13]. Традиційні методи кількісного аналізу часто стикаються з труднощами при обробці неструктурованого тексту, проте застосування алгоритмів вилучення ключових слів (наприклад, Chi-square) дозволяє будувати моделі з високим рівнем узгодженості між комп'ютерною діагностикою та оцінками психіатрів [13]. Дослідження показують, що поєднання n-gram моделей із різними алгоритмами машинного навчання, такими як Naive Bayes чи метод опорних векторів (SVM), допомагає ідентифікувати специфічні патерни вербальної поведінки [14]. При цьому використання уніграм часто забезпечує найвищу точність прогнозування, тоді як мультиграми сприяють балансуванню метрик чутливості та специфічності моделі [14].

На даний момент робочі процеси NLP вже дозволяють аналізувати не лише короткі тексти, а й величезні масиви електронних медичних карток (EMR), що налічують мільйони клінічних нотаток [15]. Використання наперед навчених трансформерних моделей для вилучення критеріїв дослідницьких доменів (RDoC) дозволяє візуалізувати траєкторії захворювання та виявляти гендерні відмінності у проявах розладів. Наприклад, аналіз клінічних даних свідчить про вищий рівень сенсомоторних порушень та відхилень у системах позитивної і негативної валентності у жінок порівняно з чоловіками, а також у ветеранів порівняно з цивільними особами. Такий контекстно-залежний аналіз у реальному часі є критично важливим для моніторингу ефективності психотерапії та виявлення ризиків суїцидальної поведінки [15].

Незважаючи на значні успіхи, виклик залишається у диференціації ПТСР від супутніх станів, зокрема, депресії [16]. Порівняльний аналіз різних класифікаторів на спеціалізованих наборах даних, таких як DAIC-WOZ, демонструє, що ансамблеві методи, а саме Random Forest та XGBoost, здатні досягати точності близько 84%, що суттєво перевищує результати класичних моделей SVM [16]. Це підтверджує, що вдосконалення налаштувань параметрів та вибір відповідних архітектур є ключовими для прогресу в ранньому виявленні ментальних розладів [13]. Таким чином, інтеграція методів автоматизованого аналізу

самовираження пацієнтів у клінічну практику дозволяє не лише скоротити витрати, а й забезпечити своєчасну допомогу на ранніх етапах розвитку розладу [14].

### 4. Опис даних та їхня обробка

Основою проведеного дослідження став репрезентативний корпус текстових даних, сформований шляхом агрегації кількох тематичних датасетів із платформи Kaggle. Сумарний обсяг вибірки склав 4800+ записів, які були розподілені на два основні класи. Цільовий клас, що містить ознаки ПТСР, включає 2000+ текстів, зібраних з різних джерел (зокрема, сабсетів обсягом від 80 до 1300+ записів), що дозволило охопити широку варіативність лінгвістичного вираження травматичного досвіду. Для забезпечення високої селективності моделі та мінімізації помилок диференціальної діагностики, контрольна група розміром 2780 текстів була сформована за комбінованим принципом. До її складу увійшли 1500 записів, що відповідають іншим психологічним станам та діагнозам, а також 1280 текстів нейтрального характеру, включаючи коментарі, форумні обговорення та публіцистичні статті. Такий підхід до формування датасету дозволяє моделі виявляти саме специфічні маркери ПТСР, а не просто ідентифікувати загальну емоційну забарвленість або депресивні настрої.

Для посилення обґрунтованості обраного підходу до формування вибірки та візуалізації якісного складу даних використано метод побудови хмар слів. Цей інструмент відіграє критичну роль на етапі розвідувального аналізу даних (EDA), оскільки дозволяє миттєво ідентифікувати найбільш частотні лексеми, що домінують у кожному з класів. У контексті диференціальної діагностики ПТСР, така візуалізація дає змогу визначити наявність у текстах специфічної клінічної термінології та описів симптомів, таких як «trauma», «flashback» або «nightmare», які складають семантичне ядро розладу. Це підтверджує, що модель буде навчатися на змістовних маркерах травматичного досвіду, а не на загальноживаній лексиці, притаманній негативному сентименту в цілому.

На рис. 1 наведена хмара слів дозволяє провести глибокий семантичний аналіз лексичного складу цільового класу, що містить ознаки ПТСР. Центральне місце та найбільший розмір слова «feel» свідчать про те, що в основі досліджуваного корпусу лежать тексти, насичені суб'єктивними емоційними переживаннями та описом внутрішніх станів. Це вказує на високу концентрацію самонаративів, де автори фокусуються на власних відчуттях. Слово «ptsd», будучи одним із ключових візуальних центрів, підтверджує тематичну цілісність вибірки та пряму самоідентифікацію респондентів із даним станом у процесі опитування або ведення записів. Висока частотність дієслів «know»,



Оскільки хмара чітко відображає симптоматику (слова «симптоми», «біль», «страх»), модель отримує можливість вибудувати стійкі асоціації між мовними парами на рівні глибоких вкладань. Це дозволяє нівелювати проблему незбалансованості та дефіциту локальних даних, використовуючи потужний лінгвістичний фундамент, закладений під час претренінгу на великих корпусах. Таким чином, україномовна хмара слів слугує не просто ілюстрацією перекладу, а виступає засобом валідації якості підготовки даних, гарантуючи, що модель буде оперувати релевантними концептами, а не шумовими артефактами машинного перекладу.

Порівняльний аналіз англійської та україномовної хмар слів виявляє високий ступінь ізоморфізму семантичних структур, що підтверджує валідність використання машинного перекладу для підготовки корпусу даних. Найбільш вираженою особливістю є повне збереження доміанти суб'єктивного переживання: англійське семантичне ядро «feel», «know» та «want» трансформувалося у точні українські відповідники «відчуваю», «знаю» та «хочу», що зберігає фокус на першій особі, характерний для самонаративів у сфері ментального здоров'я. Ця лексична симетрія є критично важливою для багатомовних архітектур, оскільки дозволяє моделі вибудувати ідентичні векторні вкладання для емоційних станів незалежно від мови вираження.

У термінологічному аспекті спостерігається цікава трансформація: якщо в оригінальному датасеті центральне місце займає абревіатура «ptsd», то в українській версії вона доповнюється та частково заміщується прикметником «стресовий». Це свідчить про те, що Google Translate API успішно розпізнав контекст розладу, адаптувавши його під природні для української мови синтаксичні конструкції, водночас зберігши специфічні маркери, такі як «птер» та «симптоми». Клінічний кластер, що включає поняття «trauma»/«травми», «nightmare»/«кошмари» та «flashback»/«спогади», демонструє майже ідентичну частотність у обох корпусах, що гарантує збереження діагностичних ознак розладу після перекладу.

Реляційний та соціальний контексти також демонструють стійкість: частотне вживання слів «people», «friend», «family» в англійському сегменті синхронізоване із появою лексем «людей», «друзів», «сім'ї» в українському, що вказує на збереження описів міжособистісних конфліктів та соціальної ізоляції, притаманних ПТСР. Поява в українській хмарі слова «важко» як прямого семантичного еквівалента англійському «hard» додатково підсилює емоційну тональність вибірки. Така глибока семантична відповідність дозволяє зробити висновок, що сформований україномовний корпус не втратив своєї діагностичної цінності, а автоматизована адаптація

професійної лексики пройшла без суттєвих змістових викривлень, створюючи надійну базу для подальшої класифікації.

## 5. Архітектура моделі

Архітектурна основа цього дослідження базується на фундаментальному переході від статичних методів вкладання слів до глибоких контекстно-залежних представлень, що стало можливим завдяки впровадженню архітектури Transformer. Відповідно до останніх оглядів нейромережових технологій, саме ця архітектура дозволила якісно відтворювати складні синтаксичні структури та враховувати контекстуальні зв'язки, які раніше були недоступні для правило-орієнтованих чи статистичних підходів [9]. В основі обраних моделей лежить механізм самоуваги (self-attention), який дозволяє системі динамічно визначати вагу кожного слова в реченні відносно інших, формуючи цілісне семантичне поле. Для опрацювання англійського корпусу застосовано модель bert-base-uncased, яка, маючи 12 послідовних шарів кодувальників та 110 мільйонів параметрів, забезпечує глибоке двонаправлене розуміння тексту [11]. Використання саме "uncased" версії дозволяє нівелювати вплив регістру, що є критично важливим для аналізу неформальних текстів із соціальних мереж та форумів, де користувачі часто ігнорують правила пунктуації та капіталізації. Процес підготовки даних для цієї архітектури реалізується через токенизатор WordPiece, який розбиває слова на субодиниці, ефективно вирішуючи проблему обмеженого словника та дозволяючи моделі коректно інтерпретувати складні медичні терміни навіть у разі їх фрагментарного написання.

Переходячи до аналізу україномовного сегмента, вибір мультимовної моделі xlm-roberta-base був зумовлений її винятковою здатністю до крос-лінгвального перенесення знань, що дозволяє використовувати досвід, накопичений на гігантських англійських корпусах, для покращення якості класифікації українською мовою. На відміну від стандартного BERT, ця архітектура була навчена за методологією RoBERTa з використанням динамічного маскування токенів та виключенням завдання прогнозування наступного речення, що, як підтверджують останні дослідження, веде до більш стабільних та точних вкладань. Особливої уваги заслуговує використання токенизації SentencePiece на рівні байтів, яке ідеально підходить для української мови з її багатоморфологією та флективністю, оскільки дозволяє моделі вловлювати семантику коренів слів незалежно від їхніх префіксів чи закінчень. Це безпосередньо корелює із завданням детекції ПТСР, де емоційний стан автора часто закодований у специфічних дієслівних формах та прикметникових конструкціях.

Важливим елементом архітектури є фінальний шар класифікації, який впроваджується безпосередньо над вихідним станом спеціального токена [CLS]. Цей токен акумулює в собі агреговане представлення всієї послідовності, що подається на вхід, і саме його вектор використовується лінійним шаром для прийняття рішення про приналежність тексту до класу «ПТСР» або «Не ПТСР». Такий підхід дозволяє використовувати потужний потенціал наперед навчених трансформерів, проводячи лише тонке налаштування (fine-tuning) на специфічному наборі даних, що значно підвищує ефективність порівняно з навчанням моделей з нуля. Кожен із 12 шарів моделі містить 12 голів уваги, які паралельно аналізують різні аспекти лінгвістичної структури, від базової граматики на нижніх рівнях до складних семантичних концептів та емоційних станів на верхніх. Це створює надійний технологічний фундамент, який дозволяє виявляти неявні маркери травматичного стресу, що часто залишаються поза увагою класичних алгоритмів машинного навчання. Весь процес реалізовано за допомогою екосистеми Hugging Face та бібліотеки transformers, що забезпечує високу відтворюваність експериментів та можливість гнучкого налаштування гіперпараметрів навчання для досягнення максимальної точності у 84% та вище, як це зафіксовано у порівняльних дослідженнях на аналогічних датасетах.

### 6. Методологія досліджень

На рис. 3 подана діаграма активності відображає архітектурну логіку та послідовність технологічних операцій, що складають основу розробленої системи для автоматизованого скринінгу ознак ПТСР.

Робочий процес ініціюється етапом прийому вхідних даних, після чого архітектура передбачає розгалуження на паралельні потоки для забезпечення крос-лінгвальної валідності дослідження. Одна з гілок спрямована на роботу з автентичним англомовним корпусом, тоді як паралельний сегмент відповідає за генерацію та підготовку україномовної вибірки. Важливим аспектом цього етапу є впровадження багаторівневої стратегії мовної адаптації, де після первинної обробки тексту система застосовує логіку умовного переходу. Зокрема, для текстових одиниць, обсяг яких не перевищує технічного ліміту в 5000 символів, використовується програмний інтерфейс Google Translate API, що є оптимальним з точки зору швидкості та стабільності обробки. У випадках, коли структура повідомлення виявляється складнішою або перевищує вказане обмеження, задіюються потужності великих мовних моделей (LLM) для забезпечення глибшої семантичної точності та збереження контекстуальних відтінків, критичних для психіатричного дискурсу.

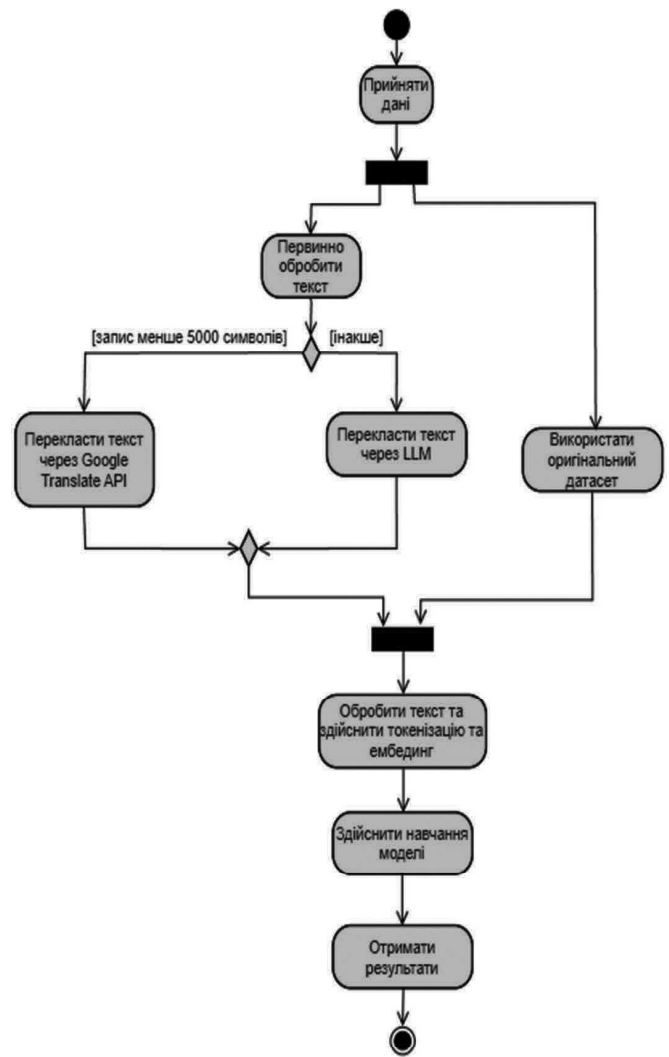


Рис. 3. UML-Діаграма активності

Після синхронізації оригінального та адаптованого потоків даних конвеєр переходить до стадії поглибленої лінгвістичної обробки, де відбувається інтеграція токенизації та ембедінгу. На цьому рівні система трансформує неструктуровані текстові послідовності у багатовимірні векторні представлення, застосовуючи алгоритми ембедінгу, які дозволяють правильно векторизувати навіть специфічну морфологію та термінологію. Центральним етапом діаграми є безпосереднє навчання моделі, що базується на архітектурі трансформерів із використанням механізму самоуваги для ідентифікації прихованих маркерів травматичного стресу. Завершальна фаза активності фокусується на отриманні та інтерпретації результатів, де агреговані контекстуальні представлення проходять крізь шар класифікації для формування об'єктивного висновку щодо наявності ознак розладу. Така структурована послідовність дій гарантує високу відтворюваність експерименту та дозволяє ефективно поєднувати можливості машинного перекладу з передовими методами глибинного навчання.

Навчання англомовної моделі bert-base-uncased здійснювалося за допомогою оптимізатора AdamW із

початковим темпом навчання  $5 \times 10^{-5}$  та коефіцієнтом розпаду wag 0,01. Процес тривав 3 епохи з використанням лінійного планувальника та встановленим рівнем дропауту 0,3 для запобігання перенавчання. Для україномовного сегмента навчання моделі xlm-roberta-base тривало 5 епох із темпом навчання  $3 \times 10^{-5}$ , розміром пакета 16 та обмеженням довжини послідовності у 128 токенів.

## 7. Результати дослідження

Використання матриці заплутаності (Confusion Matrix) є критично важливим для оцінки даного дослідження, оскільки в задачах медичного та психологічного скринінгу загальний показник точності (Accuracy) не завжди відображає реальну ефективність системи. Матриця дозволяє детально проаналізувати структуру помилок, що особливо важливо для ідентифікації ПТСР, де ціна пропущеного випадку (хибного негативу) може бути значно вищою за помилковий діагноз (хибний позитив). Порівняння матриць для англійської та української моделей дає змогу наочно побачити, як саме машинний переклад впливає на здатність нейронної мережі розрізнити специфічні патерни психічного розладу.

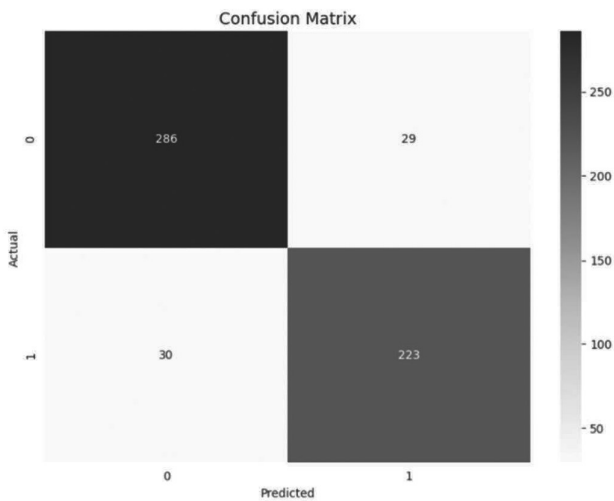


Рис. 4. Матриця заплутаності моделі для англійської мови

Матриця помилок для оригінальної моделі на основі архітектури BERT демонструє високу стабільність і збалансованість прогнозів. Алгоритм успішно ідентифікував 223 випадки ПТСР (True Positives) та 286 контрольних текстів (True Negatives), що свідчить про глибоке розуміння семантичних маркерів у рідному для моделі англійськомовному середовищі. Показники помилок першого та другого роду є майже ідентичними: 29 хибнопозитивних та 30 хибнонегативних результатів. Така симетрія вказує на відсутність зміщення моделі в бік певної категорії та підтверджує надійність обраних гіперпараметрів навчання.

У випадку з україномовною моделлю на основі XLM-RoBERTa матриця помилок відображає

збереження високої чутливості до цільового класу, проте зі зміщенням у бік гіпердіагностики. Модель коректно класифікувала 181 випадок ПТСР та 230 нейтральних записів. Позитивним аспектом є низька кількість хибнонегативних результатів у 24 випадки, що свідчить про здатність архітектури вловлювати ознаки ПТСР навіть після автоматичного перекладу. Водночас зростання кількості хибнопозитивних спрацювань до 47 вказує на те, що машинна конвертація тексту вносить певний лексичний шум, який модель може помилково інтерпретувати як ознаки травматичного досвіду.

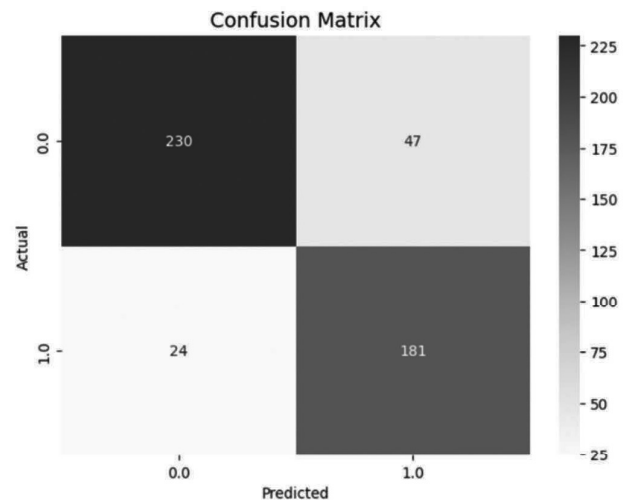


Рис. 5. Матриця заплутаності моделі для української мови

Зіставлення матриць підтверджує, що основний негативний вплив машинного перекладу зосереджений у зоні специфічності, а саме здатності відсіювати негативні випадки. В англійській моделі кількість помилок розподілена рівномірно, тоді як україномовна версія схильна до помилок типу «хибна тривога». Проте критично важливим є те, що кількість пропущених випадків (хибнонегативних результатів) в україномовній версії не зросла, а навіть дещо зменшилася у відсотковому співвідношенні, що робить цей підхід цілком придатним для завдань первинного скринінгу, де пріоритетом є максимальне охоплення потенційних хворих.

Аналіз ROC-кривої (Receiver Operating Characteristic) є фундаментальним для оцінки здатності бінарного класифікатора розрізнити стани ПТСР та контрольної групи незалежно від обраного порогу ймовірності.

ROC-крива англійськомовної моделі на основі архітектури BERT демонструє винятковий результат із показником  $AUC = 0,962$ . Графік характеризується стрімким вертикальним підйомом у лівій частині координатної площини, що вказує на здатність моделі досягати високого рівня справжніх позитивних результатів (True Positive Rate) при мінімальній кількості

хибних спрацювань. Такий результат підтверджує, що в оригінальному мовному середовищі нейронна мережа ідеально ідентифікує специфічні лінгвістичні маркери, характерні для ПТСР.

Зіставлення двох графіків показує мінімальну деградацію метрики AUC (лише на 2,3%) при переході від оригінальних текстів до україномовних. Це є ключовим доказом того, що машинний переклад не руйнує глибокі семантичні структури та емоційні патерни, які є маркерами розладу. Таким чином, обидва варіанти моделей демонструють високу прогностичну цінність, проте англійська версія залишається дещо стабільнішою за рахунок відсутності лексичних викривлень, притаманних автоматичній конвертації мови.

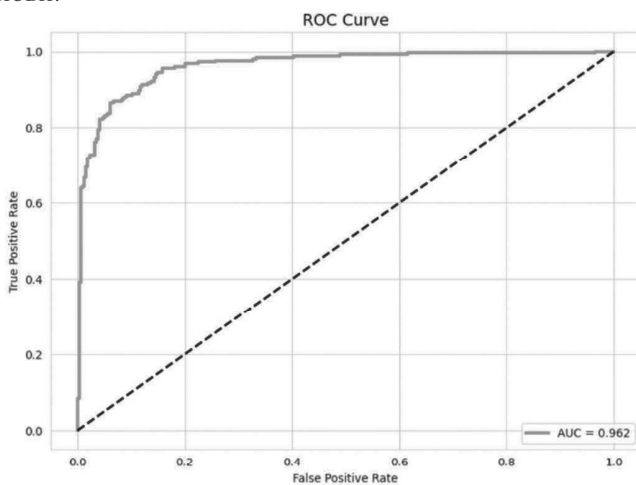


Рис. 6. Графік ROC-кривої моделі для англійської мови

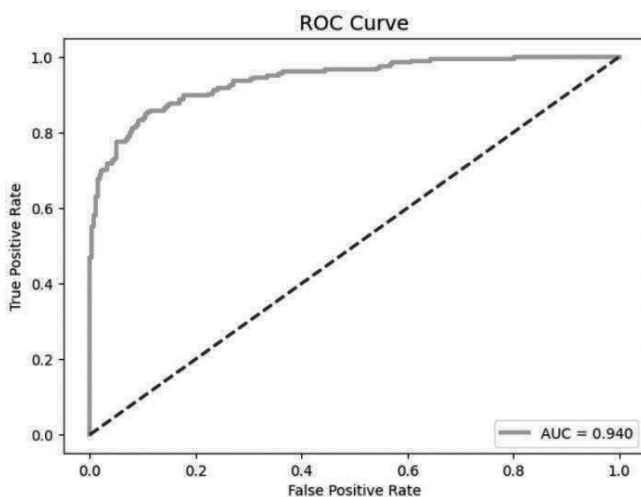


Рис. 7. Графік ROC-кривої моделі для української мови

Для україномовної моделі показник площі під кривою становить  $AUC = 0,940$ , що згідно з загальноприйнятими стандартами оцінки медичних тестів вважається відмінним показником. Хоча траєкторія кривої має дещо плавніший вигин порівняно з англійською версією, вона стабільно утримується у верхньому лівому куті графіка. Це свідчить про те, що

модель XLM-RoBERTa успішно справляється з класифікацією, зберігаючи високу роздільну здатність класифікатора

Крива точності-повноти (Precision-Recall) дозволяє оцінити якість класифікації саме для цільового класу (ПТСР), що є критично важливим за умови можливого дисбалансу даних. Вона наочно демонструє здатність моделі знаходити релевантні випадки розладу, не допускаючи при цьому надмірної кількості хибних значень.

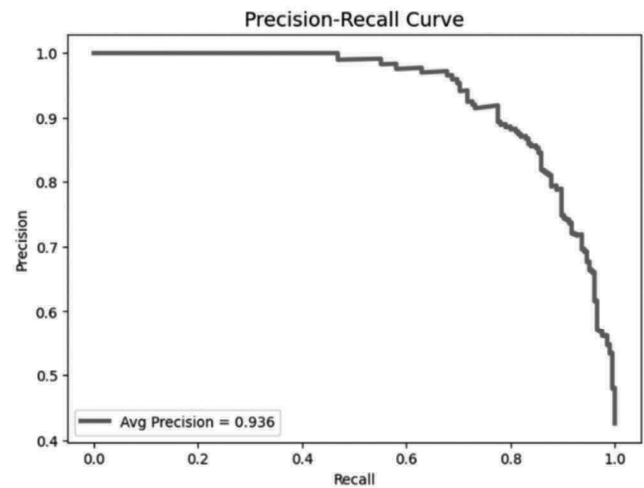


Рис. 8. Precision-Recall крива моделі для англійської мови

Для англійської моделі показник середньої точності становить  $Avg\ Precision = 0.953$ . Графік демонструє ідеальну траєкторію: точність утримується на рівні, близькому до 1.0, майже до моменту досягнення повноти 0.8. Це свідчить про те, що архітектура BERT надзвичайно чітко розрізняє специфічні маркери ПТСР в оригінальних текстах, мінімізуючи ймовірність помилкової діагностики при високому охопленні пацієнтів.

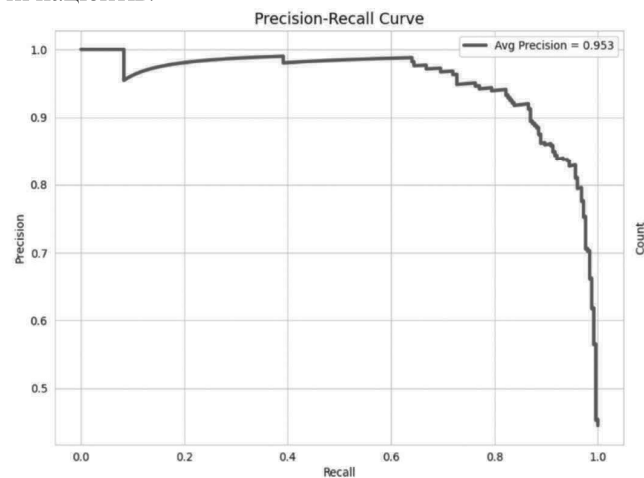


Рис. 9. Precision-Recall крива моделі для української мови

Україномовна модель продемонструвала показник середньої точності, що дорівнює 0.936. Крива зберігає характерну форму якісної моделі, хоча після рівня

повноти 0.6 спостерігається дещо стрімкіше зниження точності порівняно з англійською версією. Це підтверджує, що модель успішно ідентифікує ключові патерни розладу в перекладених текстах, хоча машинний переклад вносить певну неоднозначність, яка трохи знижує точність при максимальному охопленні.

Ретельне зіставлення графіків точності-повноти для обох мовних сегментів виявляє надзвичайно малий розрив у ключових показниках ефективності, зокрема мінімальну різницю у значеннях середньої точності, яка становить лише 1,7%. В оригінальному англійськом середовищі модель продемонструвала показник середньої точності у 0,953, тоді як україномовна версія, попри всі виклики автоматизованої адаптації, утримала високий рівень на позначці 0,936. Таке незначне відхилення підкреслює виняткову архітектурну стійкість обраних трансформерних моделей та їхню здатність зберігати діагностичну сутність даних під час крос-лінгвального перенесення знань. Особливої уваги заслуговує той факт, що обидві моделі демонструють максимальну ефективність саме у зоні високої повноти (Recall), що є безумовним пріоритетом для будь-якої системи медичного або психологічного скринінгу. Оскільки головною метою розробленого інструменту є ідентифікація якомога більшої кількості потенційних випадків ПТСР для надання своєчасної допомоги, стабільно високий рівень повноти гарантує мінімізацію кількості пацієнтів, чий стан міг би залишитися поза увагою автоматизованої системи. Також помітно незначне просідання україномовної кривої у правій частині графіка, саме в той момент, коли показник повноти наближається до свого максимуму. Це є закономірним наслідком процесів мовної адаптації та токенизації. Цей ефект відображає накопичення певного лексичного шуму та семантичних неоднозначностей, що виникають при перекладі складних емоційних наративів через автоматизований API, що іноді призводить до легкого спотворення синтаксичних конструкцій або некоректної передачі специфічних відтінків морфології української мови. Проте, попри ці технічні похибки, загальна форма кривої та високий поріг середньої точності свідчать про те, що сформовані семантичні вкладки залишаються достатньо дискретними для ефективного розмежування цільового та контрольованого класів. Таким чином, отриманий результат є цілком достатнім та валідним для практичного застосування у вітчизняних системах моніторингу ментального здоров'я, пропонуючи надійне рішення для масштабованого скринінгу в українському цифровому просторі, де верифіковані дані для навчання залишаються дефіцитним ресурсом. Це доводить, що розроблена технологія здатна забезпечити професійну якість діагностики, нівелюючи вплив машинного перекладу на фінальний прогноз моделі.

Таблиця 1

## Порівняння класифікаційних звітів моделей

Метрики	BERT (англійська мова)	XLM-RoBERTa (українська мова)
Accuracy	0.90	0.85
Precision	0.88	0.79
Recall	0.88	0.88
F1-Score	0.88	0.84
ROC-AUC	0.962	0.940

Таблиця 1 демонструє кількісні показники ефективності розроблених рішень, що базуються на архітектурах BERT та XLM-RoBERTa. Англійська модель, яка виступає у цьому дослідженні як еталонний показник (baseline), досягла високої точності класифікації (Accuracy) у 0,90, що узгоджується із результатами сучасних досліджень щодо ефективності трансформерів у виявленні психологічних маркерів. Стабільність показників Precision, Recall та F1-Score на рівні 0,88 свідчить про високу здатність моделі до диференціальної діагностики у мовному середовищі.

Найбільш вагомим та методологічно значущим результатом проведеного порівняльного аналізу є вражаюча стійкість показника повноти (Recall) для україномовної моделі, який залишився на рівні 0,88, що є ідентичним показником до оригінальної англійської версії. Такий результат є прямим підтвердженням того, що ключові лінгвістичні та емоційні маркери ПТСР, такі як описи тригерів, нав'язливих спогадів та специфічних фобій, успішно зберігаються та коректно транслюються навіть за умови використання автоматичного машинного перекладу. Збереження високого рівня Recall є критично важливим параметром для систем первинного психологічного скринінгу, оскільки головним пріоритетом у цій сфері є мінімізація хибнонегативних результатів, тобто недопущення випадків пропуску пацієнта, який дійсно потребує фахової допомоги. Хоча при переході до українського сегмента спостерігається певне зниження загальної точності до 0,85 та показника прецизійності (Precision) до 0,79, таке зниження при використанні інструментів автоматизованої мовної адаптації є цілком прийнятним та виправданим в умовах гострого дефіциту верифікованих локалізованих датасетів.

Зниження показника F1-Score з 0,88 до 0,84 для української мови відображає закономірне зростання кількості хибнопозитивних результатів, що безпосередньо зумовлено специфічним лексичним шумом та артефактами, які вносяться алгоритмами Google Translate під час обробки складних речень. Машинна конвертація може дещо спотворювати синтаксичні конструкції або вибирати менш влучні синоніми, що змушує мультимовну модель xlm-roberta-base часом

ідентифікувати ознаки травматичного досвіду там, де їх немає. Проте аналіз балансу помилок показує, що основний негативний вплив перекладу зосереджений саме у зоні специфічності, а не чутливості, що дозволяє системі стабільно виявляти реальні ознаки розладу навіть крізь призму мовної адаптації.

Виключно високі показники площі під кривою (ROC-AUC), що становлять 0,962 для англійської та 0,940 для української мови, дозволяють класифікувати обидві моделі як «відмінні» згідно із загальноприйнятою шкалою прогностичної якості медичних та психологічних тестів. Мінімальна деградація метрики AUC на 2,3% при зміні мовного середовища є фундаментальним доказом того, що трансформерні архітектури здатні розпізнавати глибинні, інваріантні семантичні структури людського горя та стресу, які залишаються незмінними незалежно від мови вираження. Отримані експериментальні дані переконливо доводять, що використання мультимовних моделей у синергії з машинним перекладом є не просто життєздатною, а й найбільш раціональною стратегією для швидкого розгортання інтелектуальних інструментів психологічної підтримки в Україні за умови подальшої відсутності реальних підтверджених даних українською мовою. Це дозволяє забезпечити високу діагностичну якість при мінімальних часових та ресурсних витратах на підготовку даних, що є критично важливим для надання оперативної допомоги військовослужбовцям та цивільним особам у режимі реального часу.

### Висновки

У межах проведеного науково-практичного дослідження розроблено та всебічно верифіковано комплексну методику автоматизованої класифікації текстових даних для ідентифікації ознак посттравматичного стресового розладу (ПТСР). Основний фокус роботи зосереджено на порівняльному аналізі ефективності сучасних архітектур трансформерів, зокрема моделей *distilbert-base-uncased* та *xlm-roberta-base*, в умовах роботи з оригінальними англійськими текстами та адаптованим україномовним сегментом. Результати проведених експериментів переконливо доводять, що технології глибинного навчання здатні успішно розпізнавати складні лінгвістичні та емоційні маркери психічних розладів, зберігаючи високу прогностичну цінність навіть за умови використання автоматичного машинного перекладу для підготовки навчальних корпусів. Це відкриває значні перспективи для швидкого розгортання систем психологічного моніторингу в українському цифровому середовищі, де спостерігається гострий дефіцит верифікованих локальних наборів даних. Репрезентативність отриманих результатів базується на використанні збалансованого корпусу даних обсягом у 4822 записи,

де цільовий клас розміром 2042 тексти охоплює широкий спектр вербального вираження травматичного досвіду. Важливою методологічною перевагою стало формування контрольної групи з 2780 записів за комбінованим принципом, що включав не лише нейтральний контент, а й описи інших психологічних станів. Такий підхід дозволив моделям вийти за межі простої ідентифікації негативного сентименту та зосередитися на специфічних клінічних маркерах ПТСР, що підтверджується семантичним аналізом хмар слів. Візуалізація даних продемонструвала високий ступінь ізоморфізму семантичних структур: англійське ядро («feel», «know», «ptsd») знайшло точне відображення в українських відповідниках («відчуваю», «знаю», «стресовий»), що гарантує збереження діагностичної цінності текстів після їхньої мовної адаптації. Кількісні показники ефективності продемонстрували високу стабільність обраних архітектур. Англійська модель досягла еталонної точності на рівні 0.90 при збалансованому показнику F1-score 0.88. Найбільш значущим результатом є продуктивність україномовної версії на базі *xlm-roberta-base*, яка показала загальну точність 0.85, зберігши при цьому ідентичний до оригіналу рівень повноти (Recall) дорівнює 0.88. Це свідчить про те, що незважаючи на певну деградацію специфічності та зростання кількості хибно-негативних результатів через вплив машинного перекладу, модель не втрачає здатності виявляти реальні випадки розладу. В умовах первинного медичного скринінгу такий пріоритет повноти над точністю є критично важливим, оскільки ціна пропущеного випадку (хибного негативу) є значно вищою за помилкову гіпердіагностику. Аналіз ROC-кривих підтверджує відмінну прогностичну якість обох моделей: показник AUC склав 0.962 для англійської та 0.940 для української мов. Мінімальне зниження метрики AUC на 2,3% при переході між мовами є ключовим доказом стійкості трансформерних архітектур до міжмовної конвертації. Це підтверджує гіпотезу про те, що нейронні мережі успішно розпізнають інваріантні семантичні структури травматичного досвіду, незалежно від морфологічної складності мови вираження. Виявлене зміщення україномовної моделі у бік помилок типу «false positive» є закономірним наслідком автоматизованої адаптації професійної лексики, проте загальний результат залишається цілком достатнім для практичного впровадження. Практичне значення роботи полягає у можливості створення інтелектуальних систем підтримки ментального здоров'я, здатних функціонувати в режимі реального часу. Розроблений конвеєр, що включає багаторівневу стратегію перекладу через Google Translate API для коротких записів та LLM для складних структур, забезпечує високу швидкість та стабільність обробки даних. Такі інструменти можуть стати основою для цифрових платформ

допомоги військовослужбовцям та цивільним особам, що постраждали внаслідок збройної агресії. Подальший розвиток дослідження вбачається у донавчанні моделей на нативних україномовних корпусах та інтеграції методів пояснювального штучного інтелекту (ХАІ) для підвищення прозорості та інтерпретованості прийнятих діагностичних рішень.

#### Список використаної літератури

- [1] Sawalha J., Yousefnezhad M., Shah Z. et al. Detecting Presence of PTSD Using Sentiment Analysis From Text Data // *Frontiers in Psychiatry*. — 2022. — Vol. 12. — Art. 811392. — DOI: <https://doi.org/10.3389/fpsy.2021.811392> (дата звернення: 05.01.2026).
- [2] Blekic W., D'Hondt F., Shalev A. Y., Schultebraucks K. A systematic review of machine learning findings in PTSD and their relationships with theoretical models // *Nature Mental Health*. — 2025. — Vol. 3. — P. 139-158. — DOI: <https://doi.org/10.1038/s44220-024-00351-4> (дата звернення: 07.01.2026).
- [3] Yuan Y., Liu Z., Miao W., Tian X. Automatic screening for posttraumatic stress disorder in early adolescents following the Ya'an earthquake using text mining techniques // *Frontiers in Psychiatry*. — 2024. — Vol. 15. — Art. 1439720. — DOI: <https://doi.org/10.3389/fpsy.2024.1439720> (дата звернення: 10.01.2026).
- [4] Srinivasan M., Bhatia A. Using NLP to Detect Mental Health Disorders Through Patient Narratives // *ResearchGate*. — 2024. — Режим доступу: <https://www.researchgate.net/publication/380123456> (дата звернення: 11.01.2026).
- [5] Шевчук А. В., Яновець А. І., Стернічук В. Б. Нейромережеві моделі у перекладі німецькомовних і англомовних текстів // *Академічні студії. Серія «Гуманітарні науки»*. — 2025. — № 3. — С. 12-18. — DOI: <https://doi.org/10.52726/as.humanities/2025.3.12> (дата звернення: 13.01.2026).
- [6] Ольховська А., Андрущенко Ю. Експериментальне дослідження впливу машинного перекладу на якість перекладу // *Вісник Харківського національного університету імені В. Н. Каразіна. Серія «Іноземна філологія»*. — 2024. — № 99. — С. 76-84. — Режим доступу: <https://periodicals.karazin.ua/foreignphilology/article/view/23867> (дата звернення: 18.01.2026).
- [7] Wu Y., Chen J., Mao K., Zhang Y. Automatic Post-Traumatic Stress Disorder Diagnosis via Clinical Transcripts: A Novel Text Augmentation with Large Language Models // *IEEE WIVM*. — 2023. — P. 1957-1962. — DOI: <https://doi.org/10.1109/WIVM59613.2023.10385704> (дата звернення: 16.01.2026).
- [8] Chen F., Ben-Zeev D., Sparks G. et al. Detecting PTSD in Clinical Interviews: A Comparative Analysis of NLP Methods and Large Language Models // *arXiv preprint*. — 2025. — arXiv:2504.01216. — Режим доступу: <https://arxiv.org/abs/2504.01216> (дата звернення: 18.01.2026).
- [9] Sharma C., Vaid A., Sharma K. Natural Language Processing in SAP: Enhancing User Interactions and Data Analysis through NLP // *IJCSET*. — 2024. — Vol. 2. — P. 58-76 — ISSN: 3048-5355 (дата звернення: 19.01.2026).
- [10] Журко Д., Білоус І. Використання моделей вкладання слів в обробці природної мови // *Технічні науки та технології*. — 2025. — № 1 (39). — С. 151-160. — DOI: [https://doi.org/10.25140/2411-5363-2025-1\(39\)-151-160](https://doi.org/10.25140/2411-5363-2025-1(39)-151-160) (дата звернення: 20.01.2026).
- [11] Gardazi N. M., Daud A., Malik M. K. et al. BERT applications in natural language processing: a review // *Artificial Intelligence Review*. — 2025. — Vol. 58, No. 166. — DOI: <https://doi.org/10.1007/s10462-025-11162-5> (дата звернення: 21.01.2026).
- [12] Черняк І. О. Дослідження можливостей моделі DistilBERT для подальшого використання в системі автоматизації електронного документообігу // *Вчені записки ТНУ імені В. І. Вернадського. Серія: Технічні науки*. — 2024. — Т. 35 (74), № 5. — УДК 004.01. — DOI: <https://doi.org/10.32782/2663-5941/2024.5.1/47> (дата звернення: 28.01.2026).
- [13] He Q., Veldkamp B. P., Glas C. A. W., de Vries J. Automated Professional Diagnosis of Posttraumatic Stress Disorder Using Text Mining Techniques // *Journal of Psychopathology and Behavioral Assessment*. — 2012. — Vol. 34. — P. 40-47. — DOI: <https://doi.org/10.1007/s10862-011-9259-3> (дата звернення: 24.01.2026).
- [14] He Q., de Vries J., Veldkamp B. P. et al. Text Mining for Posttraumatic Stress Disorder Risk Detection // *In: Healthcare Data Analytics*. — 2015. — P. 349-362. — DOI: <https://doi.org/10.1201/b18597-17> (дата звернення: 25.01.2026).
- [15] Miranda J., Kim Y., Shing H.-C. et al. Real-time extraction of RDoC criteria from clinical notes of PTSD patients using NLP // *Scientific Reports*. — 2024. — Vol. 14. — Art. 12543. — DOI: <https://doi.org/10.1038/s41598-024-63124-y> (дата звернення: 26.01.2026).
- [16] Lorenzoni G., Silva D., Santos R. et al. Depression and PTSD detection: an assessment of machine learning and NLP techniques // *Journal of Medical Systems*. — 2025. — Vol. 49. — Art. 102. — DOI: <https://doi.org/10.1007/s10916-024-02102-w> (дата звернення: 28.01.2026).

*Received (Надійшла) 20.12.2025*

*Accepted for publication (Прийнята до друку) 10.01.2026*

*Publication date (Дата публікації) 27.03.2026*

УДК 004.8

DOI: [https://doi.org/10.30837/bi.2026.1\(104\).07](https://doi.org/10.30837/bi.2026.1(104).07)І. В. Кириченко<sup>1</sup>, Г. Ю. Терешенко<sup>2</sup>, О. В. Лазаренко<sup>3</sup><sup>1</sup>Харківський національний університет радіоелектроніки, м. Харків, Україна, [iryna.kyrychenko@nure.ua](mailto:iryna.kyrychenko@nure.ua), ORCID iD: 0000-0002-7686-6439<sup>2</sup>Харківський національний університет радіоелектроніки, м. Харків, Україна, [hlib.tereshchenko@nure.ua](mailto:hlib.tereshchenko@nure.ua), ORCID ID: 0000-0001-8731-2135<sup>3</sup>Харківський національний університет радіоелектроніки, м. Харків, Україна, [oleksandr.lazarenko@nure.ua](mailto:oleksandr.lazarenko@nure.ua), ORCID iD: 0009-0005-8957-2770

## ДОСЛІДЖЕННЯ ІНСТРУМЕНТАРІЮ СТАТИСТИЧНОГО МОДЕЛЮВАННЯ У КОНТЕКСТІ ЦИФРОВОЇ МАРКЕТИНГОВОЇ АНАЛІТИКИ

Предметом дослідження є процеси оцінювання та прогнозування ефективності цифрового маркетингу з використанням сучасних інструментів статистичного моделювання.

Метою роботи є проведення всебічного аналізу статистичних моделей оцінки маркетингової ефективності, виявлення їх функціональних переваг та недоліків, а також виробка рекомендацій щодо вибору оптимального методу залежно від бізнес-сценарію.

У статті вирішуються такі завдання: систематизація методів статистичного моделювання, визначення можливостей їх застосування до домену цифрової маркетингової аналітики, вироблення критеріїв вибору актуального інструментарію для різних бізнес-сценаріїв, визначення найбільш ефективного статистичного підходу шляхом проведення багатокритеріального аналізу прийняття рішень (MCDA) методом зваженої адитивної функції корисності.

Були задіяні такі методи: аналіз наукової літератури для систематизації підходів, метод експертних оцінок для ранжування методів, та інструментарій теорії корисності для вирішення багатокритеріальної задачі.

Висновки: обґрунтовано, що висока ефективність маркетингової аналітики досягається не використанням універсального алгоритму, а адаптивним підбором статистичного інструментарію відповідно до типу бізнес-моделі та специфіки наявних даних. Отримані результати експерименту дозволяють стверджувати доцільність використання інструментів аналізу виживаності для підприємств із контрактною моделлю збуту цифрових продуктів.

Наукова цінність: робота систематизує наявні підходи до статистичного моделювання, заповнюючи прогалину у вітчизняній літературі, де бракує методологічного порівняння альтернатив методам машинного навчання у маркетинговій аналітиці.

АНАЛІЗ ВИЖИВАНOSTI, АТРИБУЦIЙНI МОДЕЛI, ВЕЛIКI ДАНI, МАРКЕТИНГОВА АНАЛІТИКА, МЕДІА-МІКС, СПОЖИВАЦЬКИЙ АНАЛІЗ, СТАТИСТИЧНЕ МОДЕЛЮВАННЯ? СТОХАСТИКА, ТЕОРІЯ ЙМОВІРНOSTЕЙ, ЦИФРОВИЙ МАРКЕТИНГ

**I.V. Kyrychenko, G.Yu. Tereshchenko, O.V. Lazarenko. A Study of Statistical Modelling Tools in the Context of Digital Marketing Analytics.** The study examines the processes of assessing performance and predicting the effectiveness of digital marketing using statistical modelling tools.

The purpose of the work is to conduct a comprehensive analysis of statistical models for assessing marketing effectiveness, identify their functional strengths and weaknesses, and formulate recommendations for selecting the optimal method based on the business scenario.

The article addresses the following tasks: systematising statistical modelling methods, assessing their applicability to digital marketing analytics, developing criteria for selecting relevant tools for various business scenarios, and determining the most effective statistical approach by solving a multi-criteria decision-making (MCDM) problem using a weighted additive utility model (weighted sum model).

The following methods were used: analysis of the scientific literature to systematise approaches, expert assessment to rank methods, and a toolkit for utility theory to support multi-criteria decision-making.

Conclusions: It is substantiated that high marketing analytics efficiency is achieved not by using a universal algorithm, but by adaptive selection of statistical tools in accordance with the business model and the specifics of the available data. The obtained experimental results allow us to assert the feasibility of using survival analysis tools for enterprises with a contract sales model.

Scientific value: the work systematises existing approaches to statistical modelling, filling a gap in the domestic literature by providing a methodological comparison of alternatives to machine learning methods in marketing analytics.

ATTRIBUTION MODELS, BIG DATA, CONSUMER ANALYSIS, DIGITAL MARKETING, MARKETING ANALYTICS, MEDIA MIX, PROBABILITY THEORY, STATISTICAL MODELING, STOCHASTICS, SURVIVAL ANALYSIS

### Вступ

Провідною тенденцією сучасної економіки є цифровізація, яку слід розглядати не просто як технологічний перехід, а як фактор фундаментального переосмислення бізнес-моделей [1], зокрема через динаміку збільшення обсягу вхідних сигналів, притаманну будь-яким системам інтелектуального аналізу

[2]. Найбільш динамічних змін зазнає сфера маркетингу. Вплив технологій ще в середині 1990-х став каталізатором [3] появи терміну "цифровий маркетинг", що охоплює збутову діяльність в онлайн-площині.

Завдяки інноваційним інструментам (платформи, таргетинг) та новим каналам комунікації [4], цифровий маркетинг демонструє найвищі темпи зростання.

За прогнозами [5], ринок цієї галузі зросте до 1,3 трлн дол. до 2033 р. із показником CAGR 9,4%. Цей факт підтверджується активною експансією на ринок аналітичних систем провідних технологічних корпорацій (Google, Microsoft, Adobe та ін.).

Однак розвиток галузі супроводжується новими викликами. Цифровізація зрівняла доступ до каналів збуту для компаній різного масштабу, що істотно посилило конкуренцію та ринкову насиченість. З іншого боку, інформаційне перевантаження трансформувало сприйняття реклами: формується "кліпове мислення", а споживачі частіше приймають емоційні, а не раціональні рішення. Це, разом із мінливістю цифрових інструментів, вимагає переходу від постфактум-оцінки до безперервного моніторингу воронки продажів.

Динамічність ринку стимулює зміщення фокусу на розбудову довірчих відносин. Відбувся перехід від "транзакційного маркетингу" до "маркетингу відносин" [6], де аналізується весь шлях клієнта, а не окрема покупка. Це зумовлює потребу в інтелектуальному аналізі великих масивів даних та автоматизації процесів через платформи маркетингової аналітики. Тому вдосконалення їх теоретичного підґрунтя — моделей статистичного моделювання — є важливою задачею розвитку домену.

### 1. Аналіз останніх досліджень та публікацій

Можливості використання теорії статистики в цифровому маркетингу пройшли значну еволюцію. В епоху класичної моделі (1960–1980 рр.) дослідження спиралися на концепцію "маркетинг-міксу" (4P) [7] та методи описової статистики. Ключовим інструментом був аналіз кореляцій між витратами на рекламу та продажами, а метрики ефективності (ROI, CPO) розглядалися як незмінні у часі константи. Згодом (у 1990-х рр.) було доведено [8] перспективність інтелектуального аналізу даних: кластерного, факторного аналізу та методів перевірки гіпотез. Однак ці інструменти забезпечували переважно ретроспективний аналіз, тоді як сучасна аналітика потребує проспективного прогнозування — екстраполяції минулого досвіду на майбутнє. Ці задачі вирішуються в межах школи маркетингової економетрики (А. Агарвал, О. Вільчинська, Р. Фріш та ін.). Їхніми зусиллями виокремлено чотири групи алгоритмів:

- макроаналіз інкрементального ефекту маркетингу (моделі часових рядів);
- макроаналіз сукупної стратегії товаропросування (моделювання медіа-міксу);
- мікроаналіз точок рекламного контакту клієнтів (атрибуційні моделі);
- мікроаналіз сегментів споживачів та можливості їх утримання (моделювання утримання).

Підхід на основі часових рядів є найменш проробленим у літературі, хоча має перевагу м'яких вимог до вихідних даних. Сучасні моделі структурних

часових рядів, інтегровані з елементами байєсівської статистики [9], здатні виявляти стохастичні залежності та тренди навіть за умов ринкової нестабільності. Проте даному підходу все ще бракує єдиного методологічного фреймворку, а тому він не набув широкого застосування.

Натомість атрибуційні моделі є глибоко опрацьованими, особливо в контексті споживацької поведінки. Дослідження [10] підтверджує ефективність залучення теорії ігор (вектор Шеплі) для оцінки внеску каналів, а також теорії марківських ланцюгів, що реалізовано на платформах "Adobe Analytics" та "Google Analytics" [11]. Головним недоліком підходу є вимога до значного обсягу та цілісності даних для побудови повного графа точок контакту клієнта з бізнесом [12].

Найбільш популярним у сучасній теорії та практиці є фреймворк "моделювання медіа-міксу" (МММ). Переважний час свого розвитку він покладався на нелінійні регресійні моделі для моделювання впливу контрольованих факторів впливу, а також спеціальні параметричні перетворення для моделювання неконтрольованих коваріат (напр., ф-її Хілла для моделювання "згасання" реклами), але поступово втрачав актуальність через відсутність еволюції.

Ренесанс методу відбувся завдяки дослідженням корпорації "Google" [13], що інтегрували в МММ принципи байєсівської статистики. Це стало причиною ренесансу даного підходу, оскільки змогло повернути інтерес наукової спільноти до прикладних можливостей МММ. Вказані вище та інші релевантні роботи стали основою створення програмних інструментів "LightweightMMM" та "Meridian" від тієї ж компанії "Google", які значно спрощують процес побудови байєсівських МММ-моделей. Враховуючи високу значимість даного фреймворку в останні роки, він дедалі частіше стає предметом дослідження у сучасних наукових працях, як іноземних, так і вітчизняних (див., зокрема, роботу Д. Чана та М. Перрі [14], де детально розглянуто передумови повернення індустрії до МММ через кризу традиційної атрибуції та втрату точності індивідуального трекінгу користувачів).

Розглядаючи інші можливі варіанти вирішення задач маркетингової аналітики, слід зазначити існування ще одного перспективного для проблематики напряму — аналізу виживаності. В економетриці він найчастіше використовується для моделювання субпроцесів утримання та відтоку клієнтської бази, та вивчає зміну її обсягу з часом у розрізах споживацьких сегментів. Зокрема, його впровадження у поширеному бізнес-сценарії неперервних покупок та неконтрактної моделі збуту (напр., готельний бізнес) стало каталізатором появи статистичного фреймворку "Buy-till-you-Die" (BTYD). Зі своїм розвитком, останній набув низку фундаментальних моделей, які спрощують побудову процесів МА. Проте наразі

застосування його інструментарію у сфері маркетингової аналітики вивчено недостатньо глибоко. Серед нечисленних дослідників слід виділити П. Фейдера та Б. Харді, які протягом останніх декількох десятиліть видали низку праць з цієї проблематики. Так, їх робота [15] підкреслює можливість перенесення представлення відтоку клієнтів у контексті задачі сегментації споживачів на проспективну модель прогнозування КРІ життєвої цінності споживачів (*Customer Lifetime Value*) та клієнтського капіталу (*Customer Equity*) у розрізах гомогенних когорт останніх.

## 2. Визначення не вирішених аспектів проблематики.

### Мета, завдання роботи

З середини 2010-х рр. домінантним підходом у маркетинговій аналітиці стала галузь машинного навчання. Ринковий аналіз [16] фіксує високий попит на алгоритми градієнтного бустингу, рекурентні нейромережі та ансамблі дерев рішень. Галузеві дослідження [17] підтверджують, що штучний інтелект забезпечує вищу точність прогнозування на великих масивах даних. Проте цей підхід часто призводить до втрати інтерпретованості результатів (ефект "чорної скриньки"), унеможливаючи пояснення причинно-наслідкових зв'язків, критичних для планування маркетингової діяльності.

Водночас розвиток теоретичного фундаменту — статистичного моделювання — суттєво загальмувався. Це зумовлено відносною простотою інтеграції ML-моделей у наявну інфраструктуру та специфікою ринку праці: перенасиченістю фахівцями зі штучного інтелекту на тлі дефіциту експертів із знаннями теорії статистики.

У вітчизняній літературі наразі відсутні роботи, присвячені систематизації та методологічному порівнянню підходів статистичного моделювання. Відтак, питання їх продуктивності на специфічних ринках, зокрема цифрового маркетингу, залишається відкритим. Цим і зумовлена головна мета та задачі даної роботи.

Метою роботи є проведення вичерпного аналізу статистичних моделей оцінки маркетингової ефективності у контексті галузі цифрового маркетингу та задач прикладної маркетингової аналітики, виявлення їх функціональних переваг та недоліків, а також формулювання рекомендацій щодо вибору оптимального методу залежно від типу бізнес-моделі, очікувань від маркетингу та обсягу й формату збутових даних.

Об'єктом вивчення є маркетингові процеси оцінювання продуктивності реалізованих та прогнозування ефективності майбутніх збутових проектів підприємств, що просувають цифрову продукцію в онлайн-середовищі засобами цифрового маркетингу.

Предметом дослідження є сучасні економетричні підходи до аналізу, а також інформаційні методи

оцінки результатів маркетингової діяльності у цифровій площині, які ґрунтуються на підходах статистичного моделювання різних субпроцесів маркетингової діяльності бізнесу.

Задля досягнення наведеної мети, був визначений наступний перелік актуальних завдань:

- дослідити процес МА та виокремити етап проспективного аналізу інформації;
- систематизувати методи моделювання та інтелектуального аналізу маркетингових даних;
- оцінити можливості і обмеження досліджуваних методів в умовах різних бізнес-сценаріїв;
- визначити критерії вибору оптимального статистичного інструментарію для різних типів задач маркетингової аналітики;
- розробити план проведення теоретичного експерименту порівняння альтернативних підходів;
- вирішити багатокритеріальну задачу вибору релевантного методу для проведення експерименту та досягнення поставленої мети.

## 3. Матеріали та методи

### 3.1. Декомпозиція процесу маркетингової аналітики.

Маркетингова аналітика як предмет дослідження включає два різні аспекти — практичний та інженерний. З точки зору бізнесу маркетингова аналітика забезпечує чіткий зв'язок між інвестиціями в рекламну діяльність та очікуваною фінансовою віддачею від неї. З інженерної точки зору, маркетингова аналітика, з метою вивчення даних про клієнтів, виявлення інформативних закономірностей, прогнозування майбутньої поведінки споживачів та оптимізації рентабельності інвестицій у маркетинг, застосовує широкий спектр інструментів статистичного аналізу і сценарного моделювання. При цьому ефективність таких інструментів критично залежить від якості та своєчасності даних, що забезпечується впровадженням архітектурних шаблонів типу «Shift-Left», які дозволяють здійснювати очищення та валідацію потоків великих даних безпосередньо поблизу джерел їх виникнення [18].

Процес маркетингової аналітики [19] охоплює шість послідовних етапів (рис. 1):

1. Етап планування маркетингового аналізу.
2. Консолідація збутової інформації та її передобробка.
3. Сегментація або поділ споживацької бази.
4. Виконання двоступінної процедури аналізу.
5. Формування звітності та спрощення результатів для бізнесу.
6. Прийняття бізнес-рішень щодо майбутнього маркетингового проекту.

Для цілей даного дослідження, найбільш значущими етапами виступають перший (планування аналізу) та четвертий (його поетапне проведення).



Рис. 1. Схематичне зображення процесу цифрової МА

Планування маркетингового аналізу передбачає виконання двох обов'язкових процедур: визначення очікувань від збутового проекту та формалізацію критеріїв оцінки ефективності аналізу. Ці критерії набувають форми KPI, зорієнтованих на певні бізнес-цілі — наприклад, залучення нових клієнтів та збереження лояльності наявних споживачів. Залежно від цілей показники ефективності навіть у межах одного ринку збуту у різних підприємств можуть значним чином відрізнятися.

На четвертому етапі, маркетингова аналітика, незалежно від її галузевої спрямованості, повинна вирішувати дві послідовні задачі:

- оцінювати, наскільки добре була здійснена маркетингова діяльність відносно плану;
- визначати стратегію адаптивних змін для більшої рентабельності у майбутньому.

Послідовне вирішення цих задач дозволяє отримати цілісну картину операційної ефективності маркетингового проекту та його подальших перспектив, а також здобути цінні для бізнесу результати (у формі KPI) як у тактичному, так і у стратегічному часовому горизонті.

Обидві задачі вирішуються шляхом інтелектуального аналізу збутових даних, проте з використанням різних інструментів. Перша оцінка поточної ефективності маркетингу реалізується проведенням ретроспективного аналізу інформації. Друга ж, визначення подальших змін, застосовує більш складний тип аналізу — проспективний.

Ретроспективний аналіз намагається узагальнити результати минулої збутової діяльності на поточну шляхом емпіричної екстраполяції наявних спостережень. Він значним чином покладається на базові алгоритми та метрики статистичного аналізу та загалом слугує як фундамент для вироблення релевантних передбачень іншими методами. Прикладом такого результату може бути показник вартості залучення клієнта (CAC), що виражається у формі відношення загальних маркетингових інвестицій до обсягу залучених споживачів певної послуги — вже відомих консолідованих метрик.

Прспективний аналіз базується на ретроспективному аналізі та спрямований на формування

обґрунтованих даними прогнозів цільових показників на певний майбутній період часу. На відміну від нього, цей тип аналізу здатний знаходити часові залежності між періодами спостережень, відстежувати глобальні та локальні тренди, виокремлювати сезонність, а також контролювати кількісний вплив неконтрольованих зовнішніх факторів на показники ефективності маркетингової діяльності. Він представляє більш комплексний аналітичний підхід, а тому залучає більш складний підхід статистичного моделювання (а не аналізу) для вирішення поставлених задач. Прикладом результату даного аналізу може бути показник вартості клієнта (CLTV), що залежить від часу, протягом якого суб'єкт буде вважатися клієнтом, та прибутку, що залучить останній за решту часу взаємодії з бізнесом — ще невідомих предиктивних метрик.

Отже, застосування статистичного моделювання у процесі маркетингової аналітики є доречним лише на етапі безпосереднього аналізу збутових даних, а у ньому — при виконанні процедури проспективного аналізу інформації та формуванні прогнозів по обраних KPI.

### 3.2. Огляд атрибуційних моделей на основі даних.

Почати огляд визначених в аналізі літератури методів варто з найбільш простого для сприйняття підходу, атрибуції. У контексті маркетингової аналітики атрибуція визначається як процес розподілу внеску окремих елементів маркетингової стратегії у здійсненні перспективним клієнтом певної цільової дії (найчастіше кліку або конверсії). Методологічно цей підхід базується на побудові та аналізі орієнтованого графу (рис. 2) взаємодії клієнтів із точками рекламного контакту бізнесу, де вузлами виступають цифрові канали просування, а ребрами — ймовірнісні переходи між ними у часі. Вивчення причинно-наслідкових зв'язків у такій структурі дозволяє вирішувати задачі як ретроспективного, так і проспективного аналізу.

Еволюція методів атрибуції пройшла два етапи: від евристичних правил до ймовірнісного моделювання. Перше покоління представлено моделями єдиного каналу (англ. Single-Touch Attribution, STA), такими як атрибуція за останньою (англ. Last-Touch) або першою (англ. First-Touch) взаємодією, а також простими позиційними моделями (U-shaped) та кількісними метриками типу RCQ (Reach, Cost, Quality), спирається на жорстко детерміновані евристики розділення внесків. Ці методи ігнорують складну топологію шляху користувача, зводячи весь процес прийняття рішення до однієї або декількох ціннісних подій. Натомість друге покоління — багатоканальна атрибуція (англ. Multi-Touch Attribution, MTA) забезпечує більшу об'єктивність спостережень, оскільки розподіляє вагу цільової дії між усіма точками контакту на основі їхнього фактичного впливу на результат. Перевагою останнього є здатність виявляти скриті взаємодії каналів, які залишаються непоміченими при використанні первинної евристики.

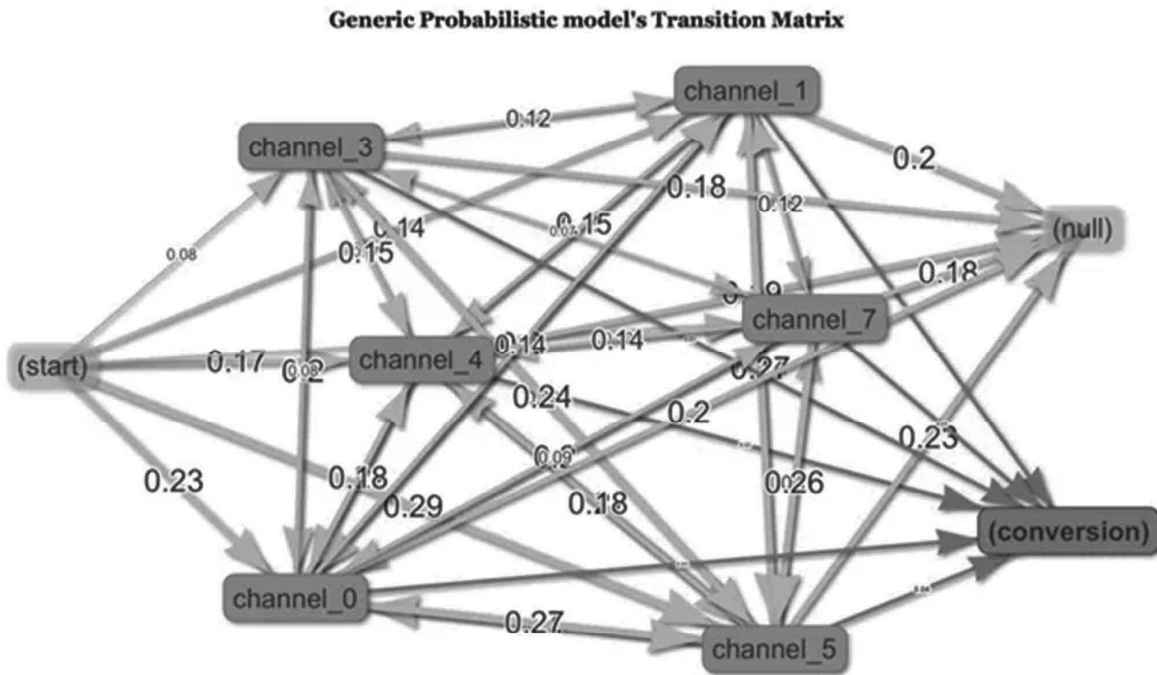


Рис. 2. Схематичне зображення атрибутивного графу споживацької активності

Далі розглянутий математичний апарат трьох найпоширеніших реалізацій МТА.

Першою є фундаментальна модель логістичної регресії, яка є загальновідомим інструментом вирішення задач бінарної (або багатокласової, якщо регресія є множинною) класифікації. Формально атрибуція тут зводиться до оцінки ймовірності настання конверсії  $P(c)$  залежно від наявності точок контакту  $X_i$  у ланцюжку взаємодії. Модель описується рівнянням (1):

$$P(c) = \frac{e^{\beta_0 + \sum \beta_k X_k}}{1 + e^{\beta_0 + \sum \beta_k X_k}}, \quad (1)$$

де коефіцієнти  $\beta_k$  інтерпретуються як міра якісного впливу  $k$ -го каналу на зважений результат.

Перевагою методу є відносна простота інтерпретації вагових коефіцієнтів, проте суттєвим недоліком виступає припущення про лінійну адитивність факторів. Останнє не дозволяє коректно моделювати синергію каналів, якщо не вводити спеціальні перехресні ознаки, що призводить до комбінаторного вибуху розмірності. Цей метод також не враховує стохастичну динаміку клієнтської взаємодії.

Іншим методом реалізації є побудова вектора Шеплі-цінності. Цей підхід походить із теорії кооперативних ігор, де рекламні канали розглядаються як гравці, що формують коаліцію для отримання певного виграву (цільової дії). Метод обчислює граничний внесок каналу "d" за всіма можливими перестановками його участі у графах медіа-міксу (2):

$$\phi_d = \sum_{P \subseteq D \setminus \{d\}} \frac{|P|!(n-|P|-1)!}{n!} (v(P \cup \{d\}) - v(P)), \quad (2)$$

де  $v(P)$  – цінність підмножини каналів.

Це забезпечує математично справедливий розподіл цінності та врахування перехресних ефектів. Головним недоліком методу є його складність: кількість підмножин для "n" каналів становить  $2^n$ , що дає часову складність  $O(2^n)$  відносно перебору перестановок. Для  $n > 20$  точний розрахунок стає неможливим і потребує ітераційної апроксимації (напр., методом Монте-Карло, що дозволяє знизити складність до поліноміальної при певній похибці).

Альтернативою комбінаторному підходу є ймовірнісне моделювання шляху користувача як марківського процесу. Граф станів включає канали  $S_1, \dots, S_n$ , стартовий стан "Start" та поглинаючі стани "Conversion" і "Null" (відтік). Матриця перехідних ймовірностей будується на основі емпіричних частот переходів між каналами у шляхах користувачів. Для оцінки важливості каналу використовується метрика ефекту видалення (Removal Effect, RE). Нехай  $P(Conv|M)$  – ймовірність досягнення стану "Conversion" у вихідному графі. Тоді якщо видалити вузол "d" з графа, всі переходи до нього будуть прямувати у "Null".

Обчислимо нову ймовірність (3)

$$P(Conv|M \setminus \{d\}) \quad (3)$$

і отримаємо (4):

$$RE_i = 1 - \frac{P(Conv|M \setminus \{d\})}{P(Conv|M)}. \quad (4)$$

Цінність конверсії розподіляється пропорційно нормалізованим ефектам видалення. Марківські моделі, особливо вищих порядків (де стан залежить від "k" попередніх кроків), дозволяють врахувати послідовність подій через переходи між станами.

Складність обчислення матриці переходів представляється як  $O(LP)$ , де  $L$  – середня довжина шляху,  $P$  – кількість шляхів, що робить метод масштабованим для довільних обсягів збутових даних. Однак "пам'ять" ланцюга обмежена його порядком. Для гарантії збіжності необхідне виконання першої марківської властивості, що, своєю чергою, зумовлює необхідність у великій (але дискретній) кількості ітерацій для побудови стійкого простору переходів між станами, що зумовлює більші вимоги до вихідних даних.

Предиктивний потенціал атрибуційного підходу обмежений тактичним горизонтом планування. Оскільки він спирається на минулі шаблони поведінки та не адаптує процес моделювання до можливих майбутніх змін, достовірність вироблених їм прогнозів знижується при зміні ринкової кон'юнктури або структури медіа-міксу. Проте при стабільності ринкових умов, методи атрибуції являють вигідну альтернативу більш комплексним підходам. Вихідним результатом моделей атрибуції є набір нормалізованих коефіцієнтів значущості  $\omega_i \in [0,1]$  для кожної точки контакту, сума яких дорівнює загальній кількості здійснених цільових дій. Для трансформації цих абстрактних величин у значимі метрики виконується агрегація зважених конверсій в розрізі окремих кампаній або каналів. Головним ретроспективним КРІ виступає атрибутована вартість за дію (Attributed CPA), яка розраховується як відношення фактичних витрат на канал до суми його фракційних внесків (напр., канал отримав 0,4 від однієї конверсії та 0,2 від іншої: кумулятивно 0,6 атрибутованої ваги). Для оцінки перспективної рентабельності розраховується атрибутований ROAS, де внесок від кожної цільової дії розподіляється пропорційно вагам каналів у графі станів.

Також варто зауважити, що застосування атрибуції не стикається зі значними обмеженнями в сучасних бізнес-сценаріях. Критичним фактором є фрагментація даних внаслідок змін у регулюванні доступу до інформації користувачів, зокрема впровадження ініціативи App Tracking Transparency (ATT) від Apple та відмова від сторонніх cookie-файлів у екосистемі трекінгу користувачів від Google. Це призводить до втрати цілісності ідентифікації користувача на різних пристроях, що унеможливує наскрізну атрибуцію відразу по цифрових каналах несумісних екосистем без порушення регуляцій обробки приватних даних. Але ж дослідження продуктивності каналів у межах одного джерела є цілком обґрунтованим та здійсненим.

### 3.3. Огляд інкрементальних моделей часових рядів.

На відміну від атрибуційного моделювання, яке фокусується на дискретних подіях взаємодії на рівні окремих користувачів, аналіз часових рядів (АЧР) розглядає маркетингову діяльність як

неперервний стохастичний процес на макрорівні. Фундаментальною відмінністю підходу є перехід від питання розподілу конверсійних внесків до задачі вимірювання чистого інкрементального приросту (відомого як "lift") – різниці між фактичними показниками та базовим попиту, який існував би навіть за відсутності маркетингового втручання. Такий контрфактний аналіз дозволяє виокремити досягнення саме рекламної діяльності, та вивчати її кількісний вплив на комерційний успіх бізнесу в ізоляції від органічної складової. Передумовою застосування АЧР є наявність у збутових даних внутрішньої причинної наслідковості: напр., обсяг продажів залежить не лише від миттєвих маркетингових стимулів, а й від інерції ринку, сформованої у попередні періоди спостережень. Така структура, шляхом дослідження часової динаміки, дозволяє виокремити стійкі закономірності, очищені від короткострокових "шумів".

Базовим інструментом моделювання даного підходу виступає клас авторегресійних моделей. Вони базуються на гіпотезі, що майбутнє значення змінної є лінійною функцією її минулих значень. Формально авторегресійна модель порядку " $p$ " описується рівнянням (5):

$$Y_t = \alpha + \sum_{i=1}^p \phi_i Y_{t-i} + \epsilon_t \quad (5)$$

де  $Y_t$  – значення відгуку (найчастіше обсяг продажів або конверсій) у момент часу " $t$ ",  $\phi_i$  – коефіцієнти автокореляції, що характеризують силу впливу лагових значень.

Для коректного визначення рекурентних закономірностей у даних в моделях АЧР існує механізм декомпозиції часового ряду на структурні компоненти. До таких компонентів належать: тренд (глобальний напрям розвитку ринку або локальні зміни попиту), сезонність (періодичні календарні коливання), цикли (хвилеподібні атипічні зміни) та випадковий шум (стохастична компонента без детермінованої часової динаміки). Виокремлення цих складових є критичним для забезпечення стаціонарності ряду, на якому можливо проводити подальший регресійний аналіз.

У маркетинговій аналітиці авторегресія розширюється моделями з розподіленним лагом (AR+DL) для врахування змін зовнішніх факторів, ковзного середнього (ARMA) для врахування помилок минулих прогнозів, а також інтегрованими моделями (ARIMA) та їх сезонними модифікаціями (SARIMA/(X)). Проте суттєвим обмеженням даного класу моделей є припущення про незмінність автокореляції моделі протягом усього періоду спостережень, що часто не відповідає динамічній природі ринку та збутової діяльності.

Більш адаптивним методом опису динамічних процесів (у т.ч. маркетингової діяльності) є використання моделей простору станів, що представляють

часовий ряд через систему двох рівнянь спостереження та стану (6):

$$\begin{aligned} - y_t &= Z_t \alpha_t + \epsilon_t, \quad \epsilon_t \sim N(0, H_t), \\ - \alpha_t &= T_t \alpha_{t-1} + R_t \eta_t, \quad \eta_t \sim N(0, Q_t), \end{aligned} \quad (6)$$

де  $y_t$  – вектор спостережуваних значень,  $\alpha_t$  – вектор прихованих станів (напр., рівень тренду, сезонності), що еволюціонують у часі згідно з матрицею переходу  $T_t$ .

Найбільш адаптивним представником класу моделей простору станів у маркетинговій аналітиці виступає метод Байєсівських структурних часових рядів (Bayesian Structural Time Series, BSTS). Узагальнено модель BSTS представляється у вигляді рівняння регресії, компоненти якого є досліджуваними стохастичними процесами (7):

$$y_t = \mu_t + \tau_t + \sum_{m=1}^M \beta_{m,t} X_{m,t} + \epsilon_t, \quad (7)$$

де  $\mu_t$  – стохастичний тренд,  $\tau_t$  – сезонна або циклічна компонента, а сума  $\beta_{m,t} X_{m,t}$  – регресійний блок, що враховує вплив екзогенних факторів  $X_{m,t}$  (витрат на канали просування) із регресійними коефіцієнтами впливу  $\beta_{m,t}$ , які можуть змінюватися у часі.

Ключовою особливістю методу є формалізація кожного компонента через систему прихованих станів, що еволюціонують за власними імовірнісними законами. Так, локальний лінійний тренд описується системою рівнянь для рівня ряду  $\mu$  та нахилу тренду  $\lambda$  (8):

$$\begin{aligned} - \mu_t &= \mu_{t-1} + \lambda_{t-1} + \eta_t^{(\mu)}, \\ - \lambda_t &= \lambda_{t-1} + \eta_t^{(\lambda)}, \end{aligned} \quad (8)$$

де  $\eta_t^{(\mu)}$  та  $\eta_t^{(\lambda)}$  – незалежні помилки, що відповідають за стохастичну зміну рівня та кута нахилу тренду відповідно.

Аналогічним чином моделюються динамічні регресори, що дозволяє враховувати змінну ефективність маркетингових каналів у часі. Еволюція коефіцієнтів  $\beta_t$  задається як процес випадкового (Брунівського) блукання:  $\beta_t = \beta_{t-1} + \eta_t^{(\beta)}$ . Це дозволяє моделі адаптуватися до структурних зрушень ринку без порушення припущень стаціонарності. Байєсівська природа методу полягає у тому, що всі невідомі параметри та стани розглядаються як випадкові величини з апіорними розподілами. Замість точкових оцінок (характерних для частотних методів типу ММП або ЗМК), BSTS генерує повний апостеріорний розподіл ймовірностей для кожного параметра, використовуючи алгоритми Марківських ланцюгів Монте-Карло (MCMC), що забезпечують коректну оцінку невизначеності прогнозів.

Повертаючись до АЧР, предиктивний потенціал даного підходу є високим для середньострокового горизонту планування, де інерція ринку та сезонні патерни залишаються стабільними. Хоча класичні

авторегресійні моделі забезпечують надійне прогнозування стаціонарних процесів, жорсткі вимоги до незмінності параметрів обмежують їх перспективи за умови великих часових діапазонів та виникнення непередбачуваних екзогенних шоків. Для задач контрфактичного аналізу, де необхідно апіорі моделювати динамічні структурні зміни, доцільно залучати більш гнучкі моделі, які розглядають параметри системи як змінні величини.

Фундаментальним результатом методів АЧР є побудова контрфактичної траєкторії – часового ряду очікуваних продажів  $\hat{Y}_t$  за сценарію відсутності маркетингової активності. Ключовим КРІ тут виступає інкрементальний ліфт, що представляється різницею між фактичними спостереженнями  $Y_t$  та змодельованим базовим рівнем на досліджуваному проміжку часу (9):

$$\Delta = \sum (Y_t - \hat{Y}_t). \quad (9)$$

На основі цієї величини розраховується інкрементальний ROAS (iROAS) або інкрементальний CPA (iCPA) – відношення чистого приросту доходу (або обсягу конверсій) до маркетингового бюджету періоду. На відміну від атрибуційних метрик, ці показники відображають лише ту частину результату, яка була отримана безпосередньо маркетинговим втручанням та очищена від сезонності та трендів.

#### 3.4. Огляд моделювання медіаміксу (MMM).

Моделювання медіа-міксу (Marketing Mix Modeling, MMM) – це економетричний підхід до кількісної оцінки внеску маркетингових та немаркетингових факторів у динаміку цільового КРІ. Методологія дозволяє вирішувати завдання ретроспективного аналізу (оцінка рентабельності) та проспективного планування (бюджетна оптимізація) на основі агрегованих історичних даних.

Тут варто підкреслити, що хоча MMM та атрибуційні моделі вирішують суміжну задачу розподілу цінності між каналами просування, обидва підходи оперують на різних рівнях абстракції. Якщо атрибуція фокусується на дискретних подіях у індивідуальних ланцюжках користувачів, то MMM аналізує агреговані величини (часові ряди витрат та продажів), що усуває необхідність ідентифікації та відстеження конкретних клієнтів. Абстраговано MMM можна розглядати як концептуальне узагальнення двох розглянутих раніше підходів: вирішення задачі атрибуції методологічним апаратом аналізу часових рядів.

У класичному вигляді MMM формалізується як адитивна лінійна модель множинної регресії:

$$Y_t = \alpha + \sum_{m=1}^M \beta_m f(X_{m,t}) + \sum_{c=1}^C \gamma_c Z_{c,t} + \epsilon_t, \quad (10)$$

де  $Y_t$  – значення КРІ у момент часу "t", " $\alpha$ " – базовий рівень продажів (незалежно від реклами),

$X_{m,t}$  а – маркетингова активність  $m$ -го каналу,  $Z_{c,t}$  – змінні контролю (сезонність, макроекономіка тощо), а  $\beta_m$  та  $\rho_c$  – коефіцієнти регресії.

Проте базова лінійна структура суперечить емпіричній природі маркетингового впливу. У фреймворку

МММ для адекватного відображення реальності у рівняння регресії вводяться специфічні параметричні перетворення змінних  $X_{m,t}$ , які моделюють два ключові стохастичні ефекти: перенос та насичення реклами (див. рис. 3).

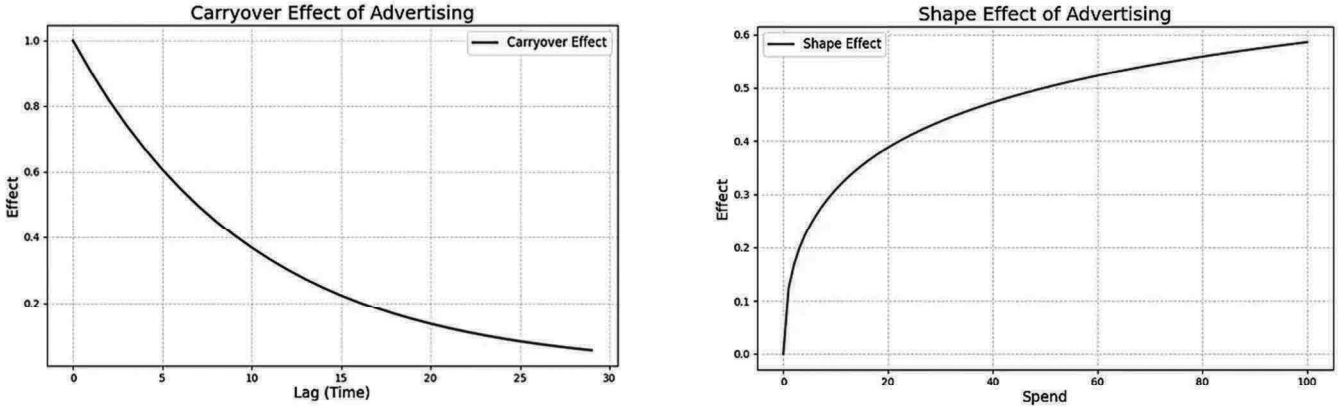


Рис. 3. Графічне зображення кривих ефектів переносу (зліва) та насичення (справа)

Передумовою формалізації цих ефектів є розуміння двох типів нелінійності регресійних моделей: за змінними та за параметрами. Нелінійність за змінними може бути лінеаризованою і вирішеною класичним МНК. Натомість нелінійність за параметрами передбачає, що похідна функції відгуку за параметром сама залежить від цього параметра, що унеможливує знаходження розв'язку через систему рівнянь та вимагає складніших методів.

Перший ефект перенесення описує інерційність впливу реклами, коли витрати у поточний період генерують відгук у майбутньому (феномен «пам'яті» ринку). Найчастіше він моделюється через трансформацію Adstock у формі геометричного загасання (11):

$$Adstock(X_{m,t}) = X_{m,t} + \lambda_m \cdot Adstock(X_{m,t-1}), \quad (11)$$

де  $\lambda_m \in [0,1]$  – параметр швидкості згасання цінності реклами. Ця рекурентна властивість фактично перетворює статичну регресію на модель із нескінченним розподіленням лагом (*Infinite Distributed Lag*, IDL) типу перетворення Койка. Наявність  $\lambda_m$  у виразі робить модель нелінійною за цим параметром, оскільки вплив  $X_{t-k}$  на  $Y_t$  дорівнює  $\beta \cdot \lambda^k$ , а параметр " $\lambda$ " неможливо вивести як лінійний множник.

Другий ефект насичення відображає фундаментальний економічний закон спадної віддачі: зі збільшенням обсягів інвестицій граничний приріст ефективності поступово зменшується, прямуючи до нуля. Ця залежність описується S-подібними (сігмоїдними) кривими, які демонструють повільний старт, фазу швидкого зростання та поступовий вихід на асимптоту. Універсальним інструментом для моделювання такої динаміки у МММ є функція Хілла (12):

$$Hill_m(X_{m,t}, K_m, \gamma_m, S_m) = \frac{S_m \cdot X_{m,t}^{\gamma_m}}{K_m^{\gamma_m} + X_{m,t}^{\gamma_m}}, \quad (12)$$

де  $S_m$  – максимальна пропускна здатність каналу (асимптота насичення),  $K_m$  – точка напівнасичення, а  $\gamma_m$  – параметр форми, що визначає крутизну кривої.

Введення степеневого показника  $\gamma_m$  робить модель повністю нелінійною за змінними  $X_{m,t}$ , що вимагає застосування складних методів апроксимації для знаходження розв'язків рівняння.

Врахування цих трансформацій перетворює задачу на проблему нелінійної оптимізації у багатовимірному просторі. Сучасне вирішення полягає у розширенні теоретичного базису МММ шляхом інтеграції байєсівської статистики. Байєсівська економетрична регресія для МММ набуває вигляду (доданки спрощено):

$$Y_t = \sum_{m=1}^M \left( \beta_m \cdot Hill(X_{m,t}, \lambda_m, \dots) \right) + \sum_{c=1}^C \gamma_c Z_{c,t}.$$

У цій формалізації рівняння параметри перетворень ( $\lambda$ ,  $K$ ,  $\gamma$ ,  $S$ ) та коефіцієнти нахилу ( $\beta_m$ ) є випадковими величинами з апіорними розподілами. Це дозволяє інтегрувати експертні знання (напр., про неможливість від'ємного впливу реклами) та отримувати стійкі оцінки параметрів через апостеріорний розподіл за загальновідомою формулою Байєса.

Вихідними результатами байєсівської моделі МММ є апостеріорні розподіли параметрів кривих відгуку та коефіцієнтів ефективності для кожного медіа-каналу. Ці параметри трансформуються у криві насичення, що дозволяє розрахувати два критично важливі KPI: середній ROAS (mROAS) для ретроспективної оцінки (внесок каналу / витрати на нього) та граничний ROAS (Marginal ROAS) для проспективного планування. Граничний ROAS визначається як перша похідна функції відгуку в точці поточних витрат і показує ефективність наступної витраченої гривні. Саме на основі вирівнювання граничних ROAS між каналами будується крива оптимального розподілу

бюджету, що є інструментом прийняття стратегічних рішень. Результати представляються у вигляді агрегованих тижневих або місячних звітів у розрізі медіаканалів. Важливо зазначити, що КРІ з МММ (зокрема внесок каналів) є "очищеними" від впливу інших факторів та базового попиту ( $\alpha$ ), тому вони, як правило, нижчі за показники атрибуції, але ближчі до істинної інкрементальності, яку вимірює АЧР.

Предиктивний потенціал МММ виходить за межі простої екстраполяції трендів. Завдяки розділенню базового рівня попиту ( $\alpha$ ) та інкрементального впливу маркетингу ( $\beta \cdot f(X)$ ), модель дозволяє проводити сценарне моделювання різних розподілів каналів. Це дає змогу оцінювати наслідки зміни структури медіаміксу та знаходити точку глобального оптимуму функції відгуку (обсягу продажів), максимізуючи ROI при обмеженому бюджеті. Часовий горизонт достовірності прогнозів для МММ є стратегічним (від кварталу до року), оскільки модель спирається на стійкі шаблони, ігноруючи короткостроковий шум.

У висновку варто зауважити, що МММ є найбільш комплексним та перевіреним часом інструментом макроаналізу, останні розширення якого поєднують математичну строгість нелінійної економетрики з гнучкістю байєсівської статистики та висновку. Також, оскільки МММ є частковим випадком АЧР та оперує агрегованими (за часом) даними, він, на відміну від атрибуції, не залежить від персональних даних користувачів, що робить його де-факто стандартом оцінки маркетингової ефективності в умовах обмежень трекінгу клієнтів.

### 3.5. Огляд моделювання споживацьких процесів.

Альтернативним до каналоцентричної атрибуції та макроекономічного МММ виступає клієнт-орієнтований підхід, що зміщує фокус аналізу з рекламних подій на суб'єктів, які їх генерують. Методологічним фундаментом цього напряму є аналіз виживаності, математичний апарат якого був запозичений з доменів медичної статистики та інженерії надійності. У першому ці методи (відомі як "frailty" моделі) використовуються для прогнозування часу до настання смерті пацієнта, а в останньому — для оцінки часу безвідмовної роботи обладнання під впливом певних навантажень. У контексті маркетингової аналітики поняття "смерті" або "відмови" трансформується у подію відтоку — припинення комерційної взаємодії клієнта з бізнесом, час настання якої є об'єктом прогнозування.

Для усунення високого стохастичного шуму, притаманного індивідуальній поведінці кожного клієнта, а також уникнення надмірного узагальнення агрегованих даних, аналіз даного підходу завжди проводиться на рівні окремих когорт: груп гомогенних клієнтів, об'єднаних певними спільними поведінковими характеристиками. Вхідним вектором даних виступає

історія транзакцій, найчастіше у агрегованому форматі RFM (Recency, Frequency, Monetary Value), що, як і випадку з МММ, дозволяє уникати використання персональних даних.

Вибір коректного статистичного інструментарію моделювання залежить від двох фундаментальних характеристик бізнес-моделі: типу відносин із клієнтом та характеру розподілу здійснених покупок у часі.

#### 1. Тип відносин:

— контрактні: відтік є явною, спостережуваною подією (напр., скасування підписки);

— неконтрактні: відтік є прихованим; бізнес не знає, чи клієнт пішов назавжди, чи просто робить паузу перед наступною транзакцією.

#### 2. Характер часу:

— неперервний: покупка може відбутися у будь-який момент (напр., покупка у магазині).

— дискретний: подія можлива лише у фіксовані моменти (напр., відвідування конференції).

Таке бінарне розділення утворює матрицю рішень (табл. 1), де кожному квадранту відповідає специфічний клас ймовірнісних моделей.

Таблиця 1

Таксономія інструментів аналізу виживаності у контексті МА

Тип відносин	Контрактні відносини (відтік спостерігається)	Неконтрактні відносини (відтік латентний)
<b>Неперервний час</b> (транзакції у будь-який момент)		
<b>Приклади бізнес-сценаріїв</b>	Мобільний зв'язок, банкові перекази	Роздрібна торгівля, електронна комерція
<b>Статистичний інструмент</b>	Класичний аналіз виживаності (не-/напів параметричні методи)	Параметричне моделювання: Покупки — процес купівлі Пуассона; Вживання клієнта — експон. розподіл
<b>Актуальні методи</b>	Процедура Каплана-Мейера, пропорційних ризиків Кокса	Фреймворк BTYD: моделі Pareto/NBD, BG/NBD
<b>Дискретний час</b> (транзакції у певні фіксовані періоди)		
<b>Приклади бізнес-сценаріїв</b>	Щомісячні підписки (SaaS), журнальні передплати	Благодійні внески, запис на регулярні події
<b>Статистичний інструмент</b>	Дискретний аналіз виживаності (параметричні моделі)	Параметричне моделювання: Покупки — процес купівлі Бернуллі; Вживання клієнта — геометр. розподіл
<b>Актуальні методи</b>	Модель клієнтського відтоку: sBG	Фреймворк BTYD: модель BG/BB

Ключовим результатом застосування чинного підходу є KPI життєвої цінності клієнта та клієнтського капіталу, динаміка яких слугує головним індикатором довгострокової ефективності маркетингової стратегії. Також методи генерують криві утримання та ймовірність активності клієнта у майбутньому, які використовуються для сегментації споживацької бази. Отримані показники відрізняються від попередньо розглянутих підходів: вони є прогностичними і вимірюють цінність відносин, а не ефективність витрат на конкретний рекламний засіб. Тому CLTV не можна прямо порівнювати з CPA або mROAS (вартість активу – ефективність операційних витрат).

Розглянувши ключові підходи до статистичного моделювання задач маркетингової аналітики, можемо перейти до формування плану проведення експерименту.

### 3.6. Постановка теоретичного експерименту.

Передумовою здійснення процедур інтелектуального аналізу інформації є виконання першого етапу процесу маркетингової аналітики – планування процесу аналізу. У чинній роботі він буде виконаний шляхом синтезу експериментального бізнес-сценарію. У ході виконання експерименту буде вирішена задача багатокритеріального вибору найбільш релевантного статистичного методу для означеного бізнес-сценарію з поміж цільових альтернатив.

Спочатку опишемо сферу функціонування підприємства-суб'єкта експерименту. Для експерименту обрано наступну предметну галузь: підприємство є комерційним B2C-постачальником цифрових продуктів, наповнення яких передбачає вивчення мов програмування в інтерактивному, "гейміфікованому" форматі. Ці продукти реалізуються за контрактною моделлю збуту та надаються кінцевим споживачам у формі SaaS-послуг. Період повторного стягнення коштів за контрактом складає один місяць. Просування товарів здійснюється засобами цифрового маркетингу моноканально: із залученням соціальної мережі "TikTok".

Наступним кроком у формуванні плану експерименту є визначення критеріїв ефективності порівнюваних методів в контексті задач маркетингової аналітики. Шляхом аналізу літературних джерел та думок провідних експертів маркетингової економетрики, було сформовано наступний вичерпний набір характеристик:

1. Вимоги методу до обсягу, повноти та структурованості вхідних даних.
2. Тривалість "прогнозного горизонту", який здатен забезпечити відповідний метод.
3. Доступність результатів аналітичного методу для бізнесу та глибина їх інформаційного критерію.
4. Операційна складність інтеграції в існуючий аналітичний процес.

5. Стійкість до зовнішніх факторів та врахування їхнього впливу.

6. Масштабованість реалізації методу відносно розміру вхідних даних.

Для представлення відповідності методів визначеним критеріям було розроблено низку якісних, інтервальних та порядкових шкал вимірювання. Приведення всіх шкал до оптимальності здійснювалося "за максимумом" – тобто в межах всіх критеріїв більше значення вважається кращим у контексті поставленої багатокритеріальної задачі.

– Вимоги до формату вхідних даних: номінальні; рівні: високі (1), помірні (2), мінімальні (3).

– Актуальність прогнозів: інтервальна, тижні; рівні: тактична (1–4), перспективна (5–13), стратегічна (16+).

– Доступність результатів: номінальна; рівні: низька (1), середня (2), висока (3).

– Складність процесної інтеграції: номінальна; рівні: дуже висока (3), висока (2), стандартна (1).

– Стійкість до зовнішніх факторів: номінальна; рівні: вбудована (1), можлива (2), відсутня (3).

– Обчислювальна масштабованість: порядкова; рівні: обмежена (0–60), пропорційна (60–80), лінійна (80–100).

## 4. Результати та їх обговорення

Для об'єктивного ранжування досліджуваних методів за обраними якісними шкалами в експерименті був задіяний метод експертних оцінок. Для цього шляхом анонімного анкетування були опитані 32 представники зацікавлених у послугах маркетингової аналітики підприємств (фахівців з галузевої аналітики і менеджерів відповідних відділів), компетентні в питаннях маркетингового аналізу та ознайомлені з проблематикою опитування. Критерій вибору підприємств ґрунтувався на їхньому впровадженні у власні аналітичні процеси цифрових платформ маркетингової аналітики – як зарубіжних (напр., Microsoft Dynamics або Google Analytics), так і вітчизняних (напр., Netpeak, Ringostat). Зміст анкети складався з 36 питань закритої форми, кожне з яких слідує загальному шаблону: "Як би Ви оцінили метод <№ 1-6> за критерієм <№ 1-6>", та варіанти якого відповідають категоріям оцінювання, викладеним вище. Якщо респондент не мав досвіду з використанням певного методу, то таке питання можливо було пропустити (такі відповіді не брали участі у розрахунку фінальних оцінок). Загальна кількість отриманих анкетних записів складає 192 одиниці, що відповідає повному охопленню матриці "експерт-метод" (32×6); проте на рівні окремих критеріїв спостерігалася нерівномірна щільність даних (див. табл. 2).

Таблиця 2  
Аналіз структури отриманих відповідей

Альтернатива (метод)	Кількість ціннісних оцінок	Середнє відхилення	Коефіцієнт варіації
Атрибуційні моделі	186	8,4	14,0
Регресійні ЧР	174	12,2	29,8
Регресійний МММ	168	10,5	16,1
Байєсівській МММ	142	15,8	21,8
Байєсівські ЧР (BSTS)	138	19,2	34,9
Моделі клієнтських субпроцесів	178	9,6	18,3

Отримана відносна частка пропусків у відповідях прямо корелювала зі складністю методу. Для класичних атрибуційних моделей цей показник склав лише 3,1%, тоді як для байєсівських структурних часових рядів (BSTS) цей показник сягнув 28,4%, що підтверджує один із тезисів даної роботи - дефіцит експертів із глибоким знанням теорії статистики є серйозною проблемою.

На табл. 3 нижче можна побачити фрагмент агрегованої матриці відповідей:

Для зведення інформації, отриманої від респондентів, було застосовано метод усередненого значення, за яким кожна ціннісна відповідь була спочатку приведена до кількісного діапазону [0;100] (евристика наведена нижче), після чого було розраховано середнє арифметичне за кількістю таких відповідей. Отримане дробове значення було округлене вгору до супремуму найближчої категорії за кожним критерієм.

Так, для умовного значення "57,8" за критерієм № 6, за визначеною евристикою, результат буде зведений до мінімального значення найближчої порядкової категорії "пропорційна" та отримає оцінку у 60 балів (див. табл. 4).

Таблиця 3  
Фрагмент відповідей респондентів на питання анкети експерименту

Досвід з плат-формами	Обсяг даних (P1-3)	Горизонт (P1-3)	Інформаційність (P1-3)	Інтегрованість (P1-3)	Стійкість (P1-3)	Масштаб (P1-3)
<b>Метод 1</b>						
Adobe Analytics, HubSpot МН	P3	P1	P3	P1	P3	P1
HubSpot МН	P3	P1	P3	P1	P3	P1
<b>Метод 2</b>						
Microsoft Dynamics 365, Netpeak	P2	P2	P2	P1	P2	P1
Netpeak	—	—	—	—	—	—
Netpeak, Ringostat	P2	P1	P2	P1	P2	P1
...						
<b>Метод 6</b>						
GA 4, GA 360, Ringostat	—	—	—	—	—	—
GA 4, Microsoft Dynamics	P1	P2	P2	P2	P2	P2
...	...	...	...	...	...	...

Таблиця 4

Відображення альтернатив за якісними категоріями шкал вимірювання

Критерій/ Альтернатива	Атрибуційні моделі	Регресійні ЧР	Регресійний МММ	Байєсівській МММ	Байєсівські ЧР (BSTS)	Моделі клієнтських субпроцесів
Вимоги до вхідних даних	Мінімальні	Помірні	Помірні	Високі	Високі	Мінімальні
Тривалість прогнозного горизонту	Тактична	Тактична	Перспективна	Стратегічна	Стратегічна	Перспективна
Інформаційність результатів	Низька	Низька	Середня	Висока	Висока	Середня
Складність інтеграції в існуючий процес	Стандартна	Стандартна	Стандартна	Висока	Дуже висока	Стандартна
Стійкість до зовнішніх факторів	Відсутня	Можлива	Можлива	Вбудована	Вбудована	Можлива
Обчислювальна масштабованість	Лінійна	Пропорційна	Пропорційна	Обмежена	Обмежена	Лінійна

Зведення якісних шкал до кількісного вигляду.

Для формування справедливих кількісних оцінок методів за критеріями відповідні їм якісні шкали необхідно перевести у кількісний формат. Для якісних шкал критеріїв № 1, 3, 5, 6, кожна з яких має 3 допустимих категорії (спрощено – низьку, середню та високу), буде застосоване наступна конвертація в порядку площину [0;100]: низький: 0, середній: 50, високий: 100. Для критерію № 2 категорії відображені

числом у діапазоні за принципом "довший прогноз – більша корисність для стратегічного аналізу":

Рівень 1. Короткий, тактичний (1–4 тижня): 25.

Рівень 2. Середній, перспективний (5–13 тижнів): 60.

Рівень 3. Довгий, стратегічний (16+ тижня): 100.

Після застосування евристики перетворення якісних шкал у кількісні було отримане фінальне представлення оцінок альтернатив (див. табл.5).

Таблиця 5

Відображення альтернатив за кількісними категоріями шкал вимірювання

Критерій/ Альтернатива	Атрибуційні моделі	Регресійні ЧР	Регресійний МММ	Байєсівській МММ	Байєсівські ЧР (BSTS)	Моделі клієнтських субпроцесів
Вимоги до вхідних даних	100	50	50	0	0	100
Тривалість прогнозного горизонту	25	25	60	100	100	50
Інформаційність результатів	0	0	50	100	100	50
Складність інтеграції в існуючий процес	100	100	100	50	0	100
Стійкість до зовнішніх факторів	0	50	50	100	100	50
Обчислювальна масштабованість реалізації	100	50	50	0	0	100

У якості методу вирішення заданої багатокритеріальної задачі, був задіяний інструментарій Теорії Корисності, зокрема модель лінійної адитивної згортки з ваговими коефіцієнтами (див. табл. 6), які відображають відносну важливість кожного критерія у процесі прийняття стратегічно-обґрунтованого рішення. Як і у випадку з цільовими критеріями, для визначення відносних ваг факторів було проведено експертне опитування серед тої ж самої групи доменних експертів.

Таблиця 6

Відображення відносних ваг критеріїв (за спаданням показника, зверху-вниз)

Критерій № 1	0,25
Критерій № 2	0,30
Критерій № 3	0,13
Критерій № 4	0,10
Критерій № 5	0,12
Критерій № 6	0,10
Усього	1,00

Для визначення остаточного показника кожного методу, відповідні критеріальні значення були зведені до зваженої суми за стовпцями. Заради більш сприйнятого порівняння, отримані зважені оцінки були

нормалізовані за відношенням показника до еталонного кількісного значення (100). Після здійснення відповідних розрахунків було отримано остаточне ранжування (табл. 7).

Таблиця 7

Ранжування зважених оцінок кожного з методів (за спаданням показника, зверху-вниз)

Альтернатива № 6	0,725
Альтернатива № 4	0,600
Альтернатива № 3	0,580
Альтернатива № 5	0,550
Альтернатива № 1	0,525
Альтернатива № 2	0,410

### Висновки

Метою даної роботи було дослідження ефективності використання різних інструментів статистичного моделювання для вирішення ключових задач маркетингової аналітики. Задля цього був здійснений методичний аналіз присвячених проблематиці іноземних і вітчизняних наукових публікацій, який дозволив систематизувати статистичні алгоритми на чотири групи, а також виділити з них ті, які мають застосування в маркетинговій аналітиці.

У якості цільових було розглянуто наступні статистичні моделі:

- стохастичний аналіз часових рядів: клас авторегресійних моделей та модель BSTS;
- моделювання медіа-міксу: економетрична лінійна та байєсівська регресійна моделі;
- атрибуція на основі даних: вектор Шеплі-цінності та модель на основі ланцюгів Маркова;
- аналіз виживаності: моделі процесів утримання та відтоку сегментів клієнтської бази.

Для визначення найбільш релевантного підходу статистичного моделювання був визначений набір критеріїв порівняння альтернатив. Для формування об'єктивної оцінки підходів за всіма критеріями був застосований метод експертної оцінки, у контексті якого проведено анкетування серед маркетингових аналітиків та інженерів даних, що представляли підприємства з налагодженими аналітичними процесами.

У ретроспективі проведення експерименту можна однозначно стверджувати, що використання інструментів аналізу виживаності, зокрема у поєднанні з алгоритмами передобробки вихідних даних та їх сегментуванням по гетерогенним когортам, дозволяють ефективно вирішувати дві задачі маркетингової аналітики: проведення ретроспективного та обґрунтованого результатами останнього проспективного типів аналізу. Отримані результати теоретичного експерименту відкривають перспективні можливості підприємствам зі схожими бізнес-сценаріями – а саме: збутом цифрових товарів за контрактною моделлю та просуванням останніх засобами середовища цифрового маркетингу.

Напрямок подальших досліджень може бути сфокусований на гібридизації різних стратегій статистичного моделювання у єдиний аналітичний конвеєр для вирішення більш комплексних завдань цифрової маркетингової аналітики, а також поєднанні їх базисного математичного апарату з навчальними можливостями штучного інтелекту, зокрема рекурентних нейронних мереж, для більшої узагальнювальної здатності отриманих моделей та покращення продуктивності прогнозів релевантних маркетингових КРІ.

#### Список літератури:

- [1] Малярчук Н. Роль диджиталізації у розвитку маркетингу та її вплив на організаційну структуру підприємств // Sustainable Socio-Economic Development Journal, V. 1(3-4), P. 90–97. DOI: [https://doi.org/10.31499/2786-7838.ssedj.2023.1\(3-4\).90-97](https://doi.org/10.31499/2786-7838.ssedj.2023.1(3-4).90-97)
- [2] Зберігання великих даних в інтелектуальних системах: монографія / І. В. Кириченко, Г. Ю. Терещенко, Н. О. Шанідзе, Н. В. Шаронова. Харків : ХНУРЕ, 2026. 222 с. DOI: <https://doi.org/10.30837/978-966-659-380-4>
- [3] Окландер М. А., Романенко О. О. Специфічні відмінності цифрового маркетингу від інтернет-маркетингу // Економічний вісник Національного технічного університету України "Київський політехнічний інститут". 2015. № 12. С. 362-371. DOI: <https://doi.org/10.20535/2307-5651.12.2015.45715>
- [4] Omar A., Nermine A. The Impact of Digital Marketing on Consumer Buying Decision Process in the Egyptian Market // International Journal of Business and Management. Vol. 15. P. 120. DOI: <https://doi.org/10.5539/ijbm.v15n7p120>
- [5] Bagekari A. Digital Marketing Service Market Analysis. Annual Report. URL: <https://www.cognitivemarketresearch.com/digital-marketing-service-market-report>
- [6] Borisavljevic K., Egan J.: Relationship marketing: Exploring relational strategies in marketing // Ekonomski Horizonti. 2013. DOI: <https://doi.org/10.5937/EKONHOR1303257B>.
- [7] Приймак В. Методи економічного аналізу. Київ : Центр навчальної літератури, 2019. 296 с.
- [8] Linoff G., Berry M. Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management. 2nd ed. 2004. 672 p.
- [9] Karlsson J. Bayesian Structural Time Series in Marketing Mix Modelling. URL: <https://www.diva-portal.org/smash/get/diva2:1752040/FULLTEXT01.pdf>
- [10] Zhao K., Mahboobi S. H., Bagheri S. Shapley Value Methods for Attribution Modelling in Online Advertising. DOI: <https://doi.org/10.48550/arXiv.1804.05327>
- [11] Mehta K., Ekta S. Marketing Channel Attribution Modelling: Markov Chain Analysis // International Journal of Indian Culture and Business Management. 2020. DOI: <https://doi.org/10.1504/IJICBM.2020.109344>
- [12] Danaher P., & van Heerde H. Delusion in Attribution: Caveats in Using Attribution for Multimedia Budget Allocation // Journal of Marketing Research. 2018. Vol. 55. DOI: <https://doi.org/10.1177/0022243718802845>
- [13] Jin Y., Yueqing W. et al. Bayesian Methods for Media Mix Modeling with Carryover and Shape Effects. URL: <https://static.googleusercontent.com/media/research.google.com/ru//pubs/archive/46001.pdf>
- [14] Chan D., Perry M. Challenges and Opportunities in Media Mix Modeling. URL: <https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/2d0395b-c7d4d13ddedf54d744ba7748e8ba8dd1.pdf>
- [15] Fader, P., Hardie, B.. Probability Models for Customer-Base Analysis // Journal of Interactive Marketing. 2009. Vol. 23. P. 61-69. DOI: <https://doi.org/10.1016/j.intmar.2008.11.003>
- [16] Mishra A., Reddy U. S. A comparative study of customer churn prediction in telecom industry using ensemble-based classifiers // Proc. ICICI. 2017. P. 721–725. DOI: <https://doi.org/10.1109/ICICI.2017.8365230>
- [17] Huang, M. H., Rust, R. T. A strategic framework for artificial intelligence in marketing // Journal of the Academy of Marketing Science. 2021. Vol. 49. P. 30-50. DOI: <https://doi.org/10.1007/s11747-020-00749-9>
- [18] Vysotska V., Kyrchenko I. Competency module of shift-left architecture in big data // Proc. AICS-CoLIInS 2025. V. 4015. P. 32-46. DOI: <https://doi.org/10.31110/COLINS/2025-3/003>
- [19] Городецький Ю. Предикивна аналітика та її роль у прийнятті стратегічних рішень у маркетингу // Journal of Strategic Economic Research. 2023. С. 65-72. DOI: <https://doi.org/10.30857/2786-5398.2023.5.7>

Received (Надійшла) 22.12.2025

Accepted for publication (Прийнята до друку) 18.01.2026

Publication date (Дата публікації) 27.03.2026

УДК 004.8

DOI: [https://doi.org/10.30837/bi.2026.1\(104\).08](https://doi.org/10.30837/bi.2026.1(104).08)Н. Б. Гулієв<sup>1</sup>, О. С. Назаров<sup>2</sup><sup>1</sup>ХНУРЕ, м. Харків, Україна, [nural.huliiev@nure.ua](mailto:nural.huliiev@nure.ua), ORCID ID: 0000-0003-2123-0377<sup>2</sup>ХНУРЕ, м. Харків, Україна, [oleksii.nazarov1@nure.ua](mailto:oleksii.nazarov1@nure.ua), ORCID ID: 0000-0001-8682-5000

## ДОСЛІДЖЕННЯ СПОСОБІВ УСУНЕННЯ ДИСБАЛАНСУ КЛАСІВ СЕРЕД МЕДИЧНИХ ТА ПСИХОЛОГІЧНИХ ДАНИХ ДЛЯ ПОБУДОВИ АЛГОРИТМУ ВИПАДКОВИЙ ЛІС

Випадковий ліс є широковідомим методом прогнозування, який застосовується в екології, бізнесі, фінансах, медицині, ІТ. Не дивлячись на те, що він є досить потужним механізмом будування відповідних моделей, алгоритм може надавати некоректні результати, тому потребує удосконалення. Таких способів наразі багато, але необхідними в нашому випадку, а саме у спостереженні попереджень розвитку психологічних розладів серед людей, хворих на гіпо- та гіпертиреоз, є не всі. В минулому дослідженні ми обирали найоптимальніший спосіб налаштувань гіперпараметрів для будування дерев рішень, де найпідходящим та найкращим було виявлено метод байєсівської оптимізації. В цьому експерименті наступним кроком є вибір підходу усунення дисбалансу класів серед медичних даних пацієнтів серед таких, як undersampling, oversampling, SMOTE, RUSBoost, збалансований випадковий ліс (BRF) та ADASYN. На основі даних за критеріями часу, accuracy, precision, recall та f1 побудовано лінійну адитивну згортку задля прийняття рішення. За її показниками видно, що в нашому випадку слід обирати RUSBoost в якості способу боротьби з класами меншості, щоб алгоритм видавав більш точні результати.

ВИПАДКОВИЙ ЛІС, ДИСБАЛАНС КЛАСІВ, КЛАС МЕНШОСТІ, ADABOOST, ADASYN, OVERSAMPLING, RUSBOOST, SMOTE, UNDERSAMPLING

**N. B. Huliiev, O. S. Nazarov. Research on ways to eliminate class imbalance among medical and psychological data in building a random forest algorithm.** Random forest is a widely known forecasting method used in ecology, business, finance, medicine, and IT. Despite being a powerful mechanism for building relevant models, the algorithm can produce incorrect results and therefore needs improvement. There are many such methods available, but not all of them are necessary in our case, namely in monitoring the development of psychological disorders among people with hypo- and hyperthyroidism. In a previous study, we selected the most optimal method of hyperparameter settings for building decision trees, where the Bayesian optimization method was found to be the most suitable and best. In this experiment, the next step is to choose an approach to eliminate class imbalance among patient medical data, such as undersampling, oversampling, SMOTE, RUSBoost, balanced random forest (BRF), and ADASYN. Based on the data, a linear additive convolution was constructed for decision making according to the criteria of time, accuracy, precision, recall, and f1. Its indicators show that in our case, RUSBoost should be chosen as a way to combat minority classes so that the algorithm produces more accurate results.

ADABOOST, ADASYN, CLASS IMBALANCE, MINORITY CLASS, OVERSAMPLING, RANDOM FOREST, RUSBOOST, SMOTE, UNDERSAMPLING

### Вступ

Медична галузь стикається з великою кількістю ускладнень, що виникають на тлі вже наявних захворювань. Це стосується й гіпо- та гіпертиреозу, оскільки ці стани можуть провокувати розвиток супутніх патологій, які ускладнюють лікування основної хвороби. Одним із таких ускладнювальних чинників є психологічні порушення.

У реальних умовах набори даних рідко бувають повністю збалансованими, і проблема дисбалансу дедалі частіше ускладнює процес класифікації, що робить її актуальною темою досліджень. Незбалансований датасет означає, що один клас представлений значно більшою кількістю прикладів, тоді як інші — суттєво меншою. Більшість алгоритмів машинного навчання розрахована на приблизно рівномірний розподіл класів. Тому за наявності дисбалансу моделі схильні «віддавати перевагу» більш чисельному класу, демонструючи високу точність саме для нього, тоді як менш представлений клас часто визначається гірше або

навіть ігнорується. У результаті це негативно впливає на загальну якість класифікації.

Медичні дані зазвичай формуються на основі інформації про пацієнтів, що нерідко й призводить до дисбалансу. Зокрема, кількість здорових людей у вибірці часто перевищує кількість пацієнтів із певним захворюванням. Крім того, різна поширеність окремих хвороб також зумовлює нерівномірний розподіл класів. Попри те що менші класи зазвичай мають нижчу точність прогнозування, у медицині саме вони становлять найбільший інтерес. Тому помилки при класифікації менш чисельного класу можуть мати значно серйозніші наслідки, ніж помилки щодо більшості [1].

Об'єкт даного дослідження є саме процес прогнозування психологічних розладів у пацієнтів із гіпо- та гіпертиреозом на основі медичних і психологічних показників.

Предмет дослідження — методи усунення дисбалансу класів даних алгоритму Random Forest.

## 1. Дослідження оптимізації алгоритму випадковий ліс для аналізу даних пацієнтів

Random Forest — це метод ансамблевого навчання, який застосовує дерева рішень задля того, щоб згенерувати більш точні та надійні прогностичні результати. Кожне з дерев будується за рахунок випадкових підмножин ознак вибірки та бутстреп-множини навчальних даних. Остаточний результат визначається після процесу голосування серед них. Перевагою алгоритму є можливість зменшення ступеня перенавчання та ефективність у дослідженнях з великим обсягом вхідних атрибутів вибірки.

Алгоритм випадкового лісу демонструє достатню стійкість до шуму та аномальних значень, адже кожне дерево будується на різних підвибірках об'єктів і ознак. Завдяки цьому окремі нетипові або екстремальні спостереження мають менший вплив на підсумковий результат моделі.

Дерева рішень є одним із найпоширеніших підходів до побудови алгоритмів класифікації в сучасних інформаційних системах завдяки їх численным перевагам. Зокрема, вони дозволяють працювати як з числовими, так і з категоріальними ознаками, демонструють високу стійкість до аномальних або екстремальних значень спостережень, а також забезпечують наочну інтерпретацію отриманих правил прийняття рішень. Це робить їх корисними для пояснення закономірностей у даних та застосування в експертних системах. Однією з ключових задач, що вирішуються за допомогою експертних систем, є класифікація — процес визначення класу об'єкта шляхом зіставлення його з одним із відомих системі класів.

Попри усе вищезгадане, даний метод алгоритму має низку недоліків. Одним із таких є дисбаланс класів, що і є темою даного дослідження.

У практичних задачах дисбаланс класів є досить поширеним явищем, особливо в тих сферах, де один клас суттєво переважає інші за кількістю прикладів. У таких випадках моделі машинного навчання зазвичай «схиляються» до класу більшості, що погіршує їхню здатність коректно розпізнавати об'єкти менш представленого класу. Це особливо небезпечно в галузях із високим рівнем ризику — зокрема в медицині, системах виявлення фінансового шахрайства чи прогнозуванні надзвичайних подій. Помилки у визначенні рідкісних захворювань, шахрайських операцій або екстремальних погодних явищ можуть мати серйозні наслідки. Наприклад, невчасне виявлення рідкісної хвороби може затримати лікування пацієнта, а пропущені випадки шахрайства — спричинити значні фінансові збитки. Тому подолання дисбалансу класів є не лише технічною задачею, а й важливою умовою забезпечення справедливості, надійності та точності систем ухвалення рішень.

Причиною виникнення дисбалансу класів є велика перевага кількості одного з класів поміж інших,

тому такий клас вважається негативним. Взагалі, передбачається, що в задачах класифікації розподіл даних збалансований. Тому в випадку оберненої ситуації клас із меншою кількістю вибірки ігнорується під час аналізу або досліджується некоректно [2].

## 2. Матеріали і методи досліджень

Розв'язком багатокритеріальних задач є пошук кращого варіанту серед можливих на основі поставлених вимог. В таких випадках застосовують два види множини методів. Перший зменшує кількість критеріїв, а друга — кількість альтернатив. Для цього експерименту краще використати перший, який має такі різновиди: метод відстані, згортки, головні критерії та граничні точки.

Методи згорток поділяються на лінійні, адитивні, мультиплікативні та максимінні. Метою застосування згорток є узагальнення усіх критеріїв аналізу.

Адитивна розраховується за наступною формулою:

$$K(x) = \sum_{j=1}^n a_j K_j(x), \quad (1)$$

де  $K(x)$  — загальний критерій для альтернативи  $x \in X$ ,  $(K_1(x), \dots, K_j(x), \dots, K_n(x))$  — набір вихідних критеріїв;  $n$  — число вихідних критеріїв;  $a_j(x)$  — нормуючий множник, який вказує на вагу альтернативи.

Найкращий із усіх можливих альтернатив задачі обчислюється за допомогою наступної формули:

$$x^n = \arg \max_{x \in X} K(x), \quad (2)$$

Тобто результатом є найбільше значення, отримане методом згортки.

Мультиплікативна згортка розраховується за допомогою такої формули:

$$K(x) = \prod_{j=1}^n K_j^{a_j}(x). \quad (3)$$

Максимінна згортка знаходиться за формулою:

$$K(x) = \max_i \min_j a_{ij} K_j(x). \quad (4)$$

Найкращі результати за мультиплікативними та максимінними згортками обчислюються за формулою (2).

Метод граничних критеріїв застосовується в задачах проектування і планування, в яких порогові значення критеріїв набувають значень  $k_j(x) \geq k_{j0}$ ;  $j = 1, \dots, n$ . Формула обчислення цього способу наступна:

$$K(x) = \min_j \left( \frac{K_j(x)}{K_{j0}(x)} \right) \quad (5)$$

Найкращий результат обирається формулою 2.

Метод відстані використовує відстань, яка є додатковою метрикою. Наприклад, для вибору ідеального рішення цілком достатньо інформації. Обчислимо відстань до значення максимуму  $d(x)$  для кожної альтернативи. Тоді найкраща альтернатива буде відомою із застосуванням формули:

$$x^* = \arg \min_{x \in X} d(x). \quad (6)$$

Навіть обравши метод із першої множини, доцільним може бути застосування принципу Парето (різновид другої вибірки), який полегшує подальше дослідження виключенням наявних альтернатив, якщо усі їхні значення критеріїв менші за значення інших варіантів.

Може бути інший випадок, коли значення неконтрольовані, що ускладнює побудову моделі, тому при цьому використовують метод гарантованого результату.

Для даного дослідження краще обрати метод згортки, бо порогові показники складно вирахувати. Найпростішою та точно підходящою буде лінійна адитивна згортка, якщо буде необхідним – також принцип Парето.

Першим кроком є вибір критеріїв, за якими будуть порівнюватися варіанти.

Другий крок умовний: дані можуть бути кількісними, якими лінійна адитивна згортка коректно оперує, але якщо вони – якісні, то необхідно конвертувати їх у кількісні, що змінить першочергову таблицю даних дослідження.

Якщо значення знаходяться в різних мірах вимірювання або проміжках, необхідно нормалізувати дані.

Під час четвертого етапу деякі з альтернатив видаються методом Парето: цей принцип передбачає виключення варіантів, якщо вони прозоро програють за усіма своїми показниками критеріїв з-поміж інших альтернатив.

П'ятий крок – ранжування значень, що означає розрахунок вагових коефіцієнтів кожного з критеріїв.

Шостим та останнім кроком є обчислення значення згортки для кожної із альтернатив: знаходимо суму добутків пар значень критеріїв та їх вагових коефіцієнтів.

Задачею дослідження є прийняття рішення з приводу вибору найкращого способу видалення дисбалансу даних при реалізації методу випадковий ліс.

### 3. Аналіз літературних джерел

В одному дослідженні задачею було обрати найоптимальніший метод ансамблевого навчання для видалення або принаймні зменшення дисбалансу класів даних за допомогою таких методів, як збалансований випадковий ліс (BRF), SMOTE-випадковий ліс або SMOTE-RF, RUSBoost та SMOTEBoost. Експеримент брав до уваги також AdaBoost та традиційний алгоритм випадкового лісу. Було застосовано 13 наборів множин з різними ступенями дисбалансу класів, кожна з яких мала бінарні класифікаційні значення. Тестова вибірка складала 20%, а навчальна – 80%. Для підтримки пропорційності класів використовували стратифіковану множину. Кожному із вище зазначених способів передувала оптимізація шляхом різного налаштування гіперпараметрів за 10 разів. Найкращий варіант обирався, опираючись на

значення часу обчислення, точності та відтворення. Метод збалансованого випадкового лісу показав найвищі показники відтворення та точності, а звичайний алгоритм виграв у значенні ефективності моделі [3].

У будь-якому бізнесі необхідно проводити аналіз існуючих клієнтів задля розрахунку прогнозування спадання їхньої зацікавленості у продукції. Існує дослідження, в якому розглядалося дане питання стосовно банківської галузі, де також був дисбаланс класів даних, що вирішувалося за допомогою таких способів, як RUS – метод випадкової недостатньої вибірки, ROS – метод випадкової надмірної вибірки та SMOTE – метод синтетичної надмірної вибірки меншин. Дані включали 10 000 записів з 14 характеристиками та були оброблені алгоритмами LASSO Logistic Regression та Random Forest. Ефективність моделей досліджувалася на основі таких показників, як  $f1$ , точності, відтворюваності та прецизійності. Random Forest у поєднанні із ROS-методом показав найвищий результат 86% точності, із SMOTE – 82% та із RUS – 79%. А логістична регресія відповідно 71%, 71% та 73%. В даному випадку кращим виявився спосіб випадкової надмірної вибірки [4].

Алгоритм випадкового лісу є одним із широко використовуваних методів класифікації машинного навчання, оскільки має перевагу у вигляді зменшення ризику перенавчання та покращення загальної ефективності прогнозування. Однак для даних із незбалансованими класами цей алгоритм не дозволяє досягти найкращої ефективності, особливо у прогнозуванні даних у класі меншості. Як результат, у цій статті пропонуються два підходи до повторної вибірки для збалансування даних: техніка синтетичної надмірної вибірки меншості (SMOTE) та техніка синтетичної надмірної вибірки меншості з редагованими найближчими сусідами (SMOTE-ENN). Для техніки класифікації даних алгоритм випадкового лісу застосовується до вихідних даних, а потім до результатів повторної вибірки з використанням як SMOTE, так і SMOTE-ENN. Приклад було застосовано до даних про затримку росту, що склалися з 421 випадку в класі більшості та 79 у класі меншості. Було отримано точність 89% для вихідних даних, 90% для даних, передискретизованих за допомогою SMOTE-ENN, та 91% для даних, передискретизованих за допомогою SMOTE. Найкраща точність була отримана за допомогою техніки передискретизації SMOTE, однак вона не була особливо значущою [5].

Oversampling вирішує проблему дисбалансу класів шляхом генерації даних у класі меншості для покращення ефективності класифікації. Він був застосований, щоб не допустити втрати даних у незбалансованій вибірці. Методи надмірної вибірки меншого з класів SMOTE та ADASYN генерують дані та застосовують принцип суміжності, а також зменшують вірогідність виникнення перенавчання. У даному дослідженні

Random Forest поєднали з цими способами боротьби з дисбалансом класів задля аналізу даних (9 атрибутів та 1 бінарний залежний показник) економічної кризи в Індонезії. Target-значення експерименту представляє собою значення того, чи наявні кризові умови, чи ні. Результати свідчили про те, що обробка дисбалансу класів дійсно приводила до кращих результатів, а саме за допомогою методу ADASYN з такими значеннями Accuracy, Recall, Precision, F1 та ROC AUC, як 98.26%, 66.67%, 72.22%, 65.57% та 82.93% відповідно [6].

Як відомо, однією із головних причин смерті є захворювання серця, тому дослідження в цьому питанні тривають постійно. Особливо результативними виявилися моделі LASSO Logistic Regression, Support Vector Machine та Random Forest у спостереженнях аналізу ризиків та профілактичних процедур. Під час останніх досліджень виявлено, що побудова моделей була незначно оптимізована за допомогою SMOTE, Random Oversampling та Random Undersampling. Проводився новий експеримент, метод якого було відстежити ефективність остаточних методів прогнозування серцево-судинних захворювань із застосуванням вищезгаданих методів усунення дисбалансу класів. В результаті встановлено, що кращим варіантом була розробка моделі Random Forest із SMOTE, що було видно в показниках точності, специфічності та чутливості [7].

Автоматична ідентифікація структури мозку за допомогою магнітно-резонансної томографії є дуже важливою як для досліджень у галузі нейробіології, так і як можливий інструмент клінічної діагностики. У цьому дослідженні представлено нову стратегію повністю автоматизованої сегментації гіпокампу за допомогою MPT. Вона базується на алгоритмі з контролем, який називається RUSBoost, і поєднує випадкову підбірку даних з алгоритмом підсилення. RUSBoost – це алгоритм, спеціально розроблений для незбалансованої класифікації, який підходить для великих наборів даних, оскільки використовує випадкову підбірку більшості класів. Ефективність RUSBoost порівнювали з ефективністю ADABOOST, Random Forest та загальнодоступного пакета сегментації мозку FreeSurfer. Це дослідження було проведено на наборі даних із 50 структурних зображень мозку з вагою T1. Незалежний набір даних із 50 структурних сканів мозку з вагою T1 був використаний для незалежної валідації повністю навчених стратегій. Знову сегментації RUSBoost вигідно відрізнялися від ручних сегментацій, маючи найвищі показники серед чотирьох інструментів. Більше того, коефіцієнт кореляції Пірсона між об'ємами гіпокампу, обчисленими за допомогою ручної сегментації та сегментації RUSBoost, становив 0,83 (0,82) для лівої (правої) сторони, що є статистично значущим і вищим за показники, обчислені за допомогою Adaboost, Random Forest та FreeSurfer. Запропонований метод може бути

придатним для точної, надійної та статистично значущої сегментації гіпокампу [8].

Розглянуто чимало існуючих методів усунення дисбалансу класів, але варто визначити, який саме підходить поставленій задачі.

#### 4. Експериментальні дослідження

Проведемо дослідження та оберемо найпідходящий спосіб усунення дисбалансу класів для алгоритму випадковий ліс, написаного задля аналізу медичних та психологічних показників.

У дослідженні альтернативами виступатимуть наступні:

- random undersampling;
- random oversampling;
- SMOTE;
- RUSBoost (Random Undersampling + AdaBoost);
- збалансований випадковий ліс (BRF);
- ADASYN.

Критеріями розгляду, за якими будуть будуватись три моделі, будуть такі атрибути, як:

- age – вік,
- sex – стать,
- on\_thyroxine – чи приймає тироксин,
- query\_on\_thyroxine – запит на тироксин,
- on\_antithyroid\_meds – чи приймає анти тиреоїдні ліки,
- sick – чи хворий,
- pregnant – вагітність,
- thyroid\_surgery – чи робилась операція на щитоподібній,
- I131\_treatment – лікування радіоактивним йодом,
- query\_hypothyroid – підозра на гіпотиреоз,
- query\_hyperthyroid – підозра на гіпертиреоз,
- lithium – чи приймає літій,
- goitre – зоб,
- tumor – пухлина,
- hypopituitary – гіпопітуїтаризм,
- psych – чи є психічні розлади,
- TSH\_measured – чи вимірювався TSH,
- TSH – значення TSH,
- T3\_measured – чи вимірювався T3,
- T3 – значення T3,
- TT4\_measured – чи вимірювався TT4,
- TT4 – значення TT4,
- T4U\_measured – чи вимірювався T4U,
- T4U – значення T4U,
- FTI\_measured – чи вимірювався FTI,
- FTI – значення FTI,
- TBG\_measured – чи вимірювався TBG,
- TBG – значення TBG,
- referral\_source – джерело направлення,
- target – цільовий клас,
- patient\_id – ID пацієнта.

Написаний код на Python показав, що шість алгоритмів мають такі показники (див. табл. 1).

Таблиця 1

Числові характеристики алгоритмів

Методи та критерії	Accuracy	Precision	Recall	F1	Час
Random under-sampling	0.535754	0.840746	0.535754	0.610383	0,0457
Random over-sampling	0.944794	0.943255	0.944794	0.942957	3,2543
SMOTE	0.941089	0.938677	0.941089	0.939295	13,3661
RUSBoost	0.719155	0.545685	0.719155	0.620273	0,0301
Збалансований випадковий ліс (BRF)	0.696184	0.882207	0.696184	0.750664	0,1383
ADASYN	0.938496	0.935228	0.938496	0.936277	13,1897

Як бачимо, показник часу потребує нормалізації, щоб точно оцінити його, адже чим більше час, тим гірше. Перебудуємо таблицю (див. табл. 2).

Таблиця 2

Змінені числові характеристики алгоритмів

Методи та критерії	Accuracy	Precision	Recall	F1	Час
Random under-sampling	0.535754	0.840746	0.535754	0.610383	21,8818381
Random over-sampling	0.944794	0.943255	0.944794	0.942957	0,30728575
SMOTE	0.941089	0.938677	0.941089	0.939295	0,07481614
RUSBoost	0.719155	0.545685	0.719155	0.620273	33,2225914
Збалансований випадковий ліс (BRF)	0.696184	0.882207	0.696184	0.750664	7,23065799
ADASYN	0.938496	0.935228	0.938496	0.936277	0,07581674

Принцип Парето в даному випадку неможливо застосувати, тому усі альтернативи залишаються.

Тепер розрахуємо значення згортки для кожного із варіантів (див. табл. 3).

Таблиця 3

Значення згортки

Методи	Згортка
Random under-sampling	2,59140864
Random over-sampling	2,43651573
SMOTE	2,41595685
RUSBoost	3,19695142
Збалансований випадковий ліс (BRF)	2,25584928
ADASYN	2,40859136

Найкращий результат належить алгоритму RUSBoost [9-10].

Зображення роботи методу наведено нижче (рис. 1).

RUSBoost — це алгоритм, створений спеціально для задач із незбалансованими класами, який добре підходить для роботи з великими обсягами даних,

оскільки базується на випадковій недовибірці об'єктів класу більшості.

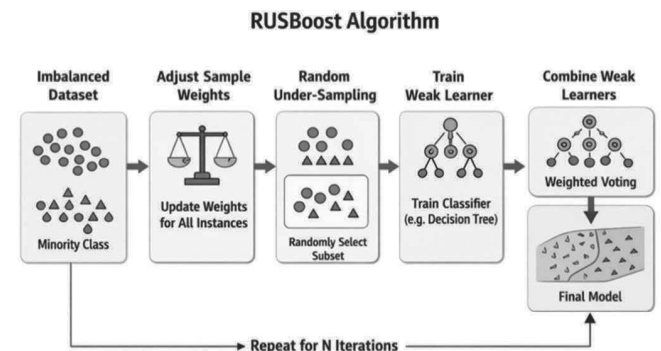


Рис. 1. RUSBoost

Метод поєднує випадкове зменшення кількості прикладів більш представленого класу з механізмом boosting. Він є альтернативою SMOTEBoost, що використовує надвибірку, генеруючи нові зразки меншості шляхом інтерполяції між наявними прикладами та поєднуючи цей процес із підсиленням. Через створення синтетичних даних SMOTEBoost збільшує тривалість навчання моделі. Хоча цей підхід продемонстрував ефективність у низці застосувань, переважно для наборів даних невеликого розміру, зі зростанням обсягу навчальних даних обчислювальні витрати SMOTE істотно підвищуються, що може зробити його непрактичним [11].

У випадку дуже великих датасетів, зокрема 3D MRI, вибір стратегії балансування стає критично важливим. Використання недовибірки, як у RUSBoost, дозволяє значно скоротити час навчання порівняно з методами надвибірки. Водночас основним недоліком цього підходу є можлива втрата частини інформації через видалення окремих прикладів із навчальної вибірки.

Крім того, існують результати досліджень, які свідчать, що RUSBoost може перевершувати SMOTEBoost, будучи простішим і швидшим методом та нерідко забезпечуючи кращі показники класифікації. Наскільки відомо авторам, це перше застосування RUSBoost-класифікаторів для задачі сегментації гіпокампа [12].

Висновки

Алгоритм випадковий ліс вирізняється своєю ефективністю, тому відомий та застосовується в багатьох сферах, але все одно потребує оптимізації. В даному дослідженні аналізувалася проблема дисбалансу класів даних.

Досі цікаво проводити дослідження для розв'язання проблеми усунення дисбалансу класів задля удосконалення роботи алгоритмів. Одним із таких популярних дієвих методів залишається RUSBoost-підхід, який і сам потребує покращення, тому триває робота над видаленням або зменшенням його недоліків.

Розглянуто існуючі методи в різних сферах, а саме: undersampling, oversampling, SMOTE, RUSBoost, збалансований випадковий ліс (BRF) та ADASYN. Метою дослідження було вирішити багатокритеріальну задачу вибору найпідходящого способу усунення класу меншості. Тому методом дослідження обрано лінійну адитивну згортку, яка надала найкращий результат на основі таких показників, як час роботи алгоритму, accuracy, precision, recall та f1. В результаті виявлено, що для усунення дисбалансу класів психологічних та медичних даних слід застосовувати RUSBoost-підхід.

RUSBoost як підхід до балансування даних був запропонований Seiffert і колегами для зменшення обчислювального навантаження, притаманного SMOTEBoost. Метод поєднує випадкову недовибірку (Random Under Sampling, RUS) із алгоритмом підсилення AdaBoost.

Для вирівнювання розподілу класів застосовується стратегія RUS, що полягає у випадковому вилученні частини зразків із класу більшості. Кожна ітерація boosting складається з двох кроків. Спочатку виконується недовибірка для всіх класів, за винятком найменш представленого, і вона триває доти, поки кількість прикладів у кожному класі не стане однаковою та не дорівнюватиме розміру класу меншості. Після цього до сформованої збалансованої підвибірki застосовується алгоритм AdaBoost. Отже, на кожній ітерації використовується оновлена навчальна вибірка, а не фіксований набір даних [13-14].

До недоліків RUSBoost належить потенційна втрата інформації через видалення частини прикладів під час недовибірki. Попри це, алгоритм ефективно застосовується як у задачах бінарної, так і багатокласової класифікації.

За результатами лінійної адитивної згортки RUSBoost має найвищий показник, але варто провести повторне дослідження із ваговими коефіцієнтами важливості критеріїв задля більш точного вибору методу боротьби із дисбалансом класів у дереві рішень [15-16].

У цьому дослідженні було вивчено ефективність різних методів класифікації на наборах даних з різним рівнем дисбалансу класів, що дозволило досягти мети дослідження – визначити оптимальні методи обробки дисбалансованих даних.

Проведено системне порівняння методів балансування (SMOTE, ADASYN, RUS, ROS, RUSBoost, balanced RF) для задачі прогнозування психологічних розладів серед пацієнтів із тиреоїдними порушеннями.

Запропоновано підхід до попередження розвитку психологічних розладів шляхом ранньої стратифікації ризику на основі ансамблевих методів машинного навчання з урахуванням дисбалансу класів.

Запропоновано модель машинного навчання для раннього прогнозування психологічних розладів у пацієнтів із гіпо- та гіпертиреозом на основі комплексного аналізу ендокринних та психометричних показників.

## Список літератури

- [1] Buda, M., Maki, A., & Mazurowski, M.A. (2021). A multiple combined method for rebalancing medical data with class imbalance. *Computers in Biology and Medicine*, 135, 104589. DOI: <https://doi.org/10.1016/j.combiomed.2021.104589>.
- [2] Prasetyo, E., et al. (2018). Evaluating Ensemble Learning Techniques for Class Imbalance Problem. *Scientific Journal of Informatics*, 5(2), pp. 184–193. URL: <https://journal.unnes.ac.id/journals/sji/article/view/15937/2440>
- [3] Implementation of Imbalanced Learning Methods Using RUSBoost. *Mortalita*, 4(2). URL: <https://ejournal.darunnajah.ac.id/index.php/mortalita/article/view/709/390>
- [4] Comparative Analysis of Imbalanced Data Handling Using Ensemble Methods. *RESTIA*, 7(3). URL: <https://journal.aiksauniversity.ac.id/index.php/restia/article/view/1906/853>
- [5] Performance Analysis of RUSBoost for Imbalanced Dataset Classification. *International Journal of Multidisciplinary and Current Research*, 11. URL: <https://ijmcr.in/index.php/ijmcr/article/view/1121/855>
- [6] Application of Ensemble Learning Methods on Imbalanced Medical Dataset. *Jurnal Matematika, Statistika dan Komputasi*, 20(1). URL: <https://journal.unhas.ac.id/index.php/jmsk/article/view/35552/12001>
- [7] Fernández, A., García, S., Galar, M., Prati, R.C., Krawczyk, B., & Herrera, F. (2018). Learning from Imbalanced Data Sets. *Journal of Information and Visualization*, 8(3). URL: <https://joiv.org/index.php/joiv/article/view/2283>
- [8] Iglesias, J.E., Liu, C.Y., Thompson, P.M., & Tu, Z. (2015). Automated hippocampal segmentation in 3D MRI using random undersampling with boosting algorithm. *Pattern Analysis and Applications*, 18, pp. 851–864. DOI: <https://doi.org/10.1007/s10044-015-0492-0>
- [9] Kaur, H., Pannu, H.S., & Malhi, A.K. (2020). Boosting methods for multi-class imbalanced data classification: an experimental review. *Journal of Big Data*, 7, Article 65. DOI: <https://doi.org/10.1186/s40537-020-00349-y>.
- [10] U. Hasanah, A. M. Soleh, and K. Sadik, Effect of Random Under Sampling, Oversampling, and SMOTE on the Performance of Cardiovascular Disease Prediction Models, *Jurnal Matematika, Statistika dan Komputasi* 21 (2024) 88-102.
- [11] R. Hidayat, M. A. Syawaludin, and N. Nurmalitasari, Prediksi Churn Pelanggan Multinational Bank Menggunakan Algoritma Machine Learning, *Simpatik: Jurnal Sistem Informasi dan Informatika* 4 (2024) 89-97.
- [12] F. Ismail and I. I. Lawanda, Implementasi EDMS dalam Penataan Dokumen di Rail Document System PT. Kereta Api Indonesia (Persero) Daerah Operasi 1 Jakarta, *Baca: Jurnal Dokumentasi Dan Informasi* 41 (2020) 143-168.
- [13] S. M. Kim, Y. Kim, K. Jeong, H. Jeong, and J. Kim, Logistic LASSO Regression for the Diagnosis of Breast Cancer Using Clinical Demographic Data and the BI-RADS Lexicon for Ultrasonography, *Ultrasonography* 37 (2018) 36-42.
- [14] M. Marcellina and A. Mukhlason, Analisis Prediktif Churn untuk Meningkatkan Tingkat Retensi Pelanggan pada Perusahaan SaaS Menggunakan Machine Learning, *ILKOMNIKA* 6 (2024) 21-32.
- [15] R. Zhu, Y. Guo, and J.-H. Xue, "Adjusting the imbalance ratio by the dimensionality of imbalanced data," *Pattern Recognit. Lett.*, vol. 133, pp. 217–223, 2020, doi: <https://doi.org/10.1016/j.patrec.2020.03.004>.
- [16] M. Çakır, A. Degirmenci, and O. Karal, "Exploring the Behavioural Factors of Cervical Cancer Using ANOVA and Machine Learning Techniques BT -Science, Engineering Management and Information Technology," A. Mirzazadeh, B. Erdebilli, E. Babaee Tirkolae, G.-W. Weber, and A. K. Kar, Eds., Cham: Springer Nature Switzerland, 2023, pp. 249–260

Received (Надійшла) 14.01.2026

Accepted for publication (Прийнята до друку) 10.02.2026

Publication date (Дата публікації) 27.03.2026

УДК 004.75

DOI: [https://doi.org/10.30837/bi.2026.1\(104\).09](https://doi.org/10.30837/bi.2026.1(104).09)

Яна Даніель<sup>1</sup>, Данііл Максименко<sup>2</sup>, Дмитро Панченко<sup>3</sup>,  
Ольга Калиниченко<sup>4</sup>, Олексій Турута<sup>5</sup>

<sup>1</sup>ХНУРЕ, м. Харків, Україна, [yana.daniiel@nure.ua](mailto:yana.daniiel@nure.ua), ORCID iD: 0000-0002-3895-0744

<sup>2</sup>ХНУРЕ, м. Харків, Україна, [daniil.maksymenko@nure.ua](mailto:daniil.maksymenko@nure.ua), ORCID iD: 0000-0003-3223-5130

<sup>3</sup>ХНУРЕ, м. Харків, Україна, [dmytro.panchenko@nure.ua](mailto:dmytro.panchenko@nure.ua) м. Харків, ORCID iD: 0000-0001-5454-5661

<sup>4</sup>ХНУРЕ, м. Харків, Україна, [olga.kalynychenko@nure.ua](mailto:olga.kalynychenko@nure.ua), ORCID iD: 0000-0003-1466-3967

<sup>5</sup>ХНУ ім. Каразіна, м. Харків, Україна, [Oleksii.Turuta@karazin.ua](mailto:Oleksii.Turuta@karazin.ua), ORCID iD: 0000-0002-0970-8617

## НАБІР ДАНИХ УКРАЇНСЬКИХ НОВИН ЯК БЕНЧМАРК ДЛЯ КЛАСИФІКАЦІЇ ТЕКСТІВ

У статті розглянуто задачу створення україномовних наборів даних для класифікації текстів. Запропоновано підхід для створення простого набору даних. Також створено корпус українських новин, придатний для порівняльного оцінювання моделей. На цьому наборі даних проведено бенчмаркінг сучасних моделей на основі трансформерів (mBERT, Slavic BERT, ukr-RoBERTa, ukr-ELECTRA, XLM-R) та базової моделі NB-SVM у різних режимах навчання. Результати показують, що ukr-RoBERTa, ukr-ELECTRA та XLM-R демонструють найвищу якість. XLM-R, як правило, краще працює з об'ємними текстами, тоді як ukr-RoBERTa — з більш короткими послідовностями.

ОБРОБКА УКРАЇНСЬКОЇ МОВИ, КЛАСИФІКАЦІЯ ТЕКСТУ, ТРАНСФОРМЕРИ, ТЕКСТОВИЙ НАБІР ДАНИХ

**Y. Daniil, D. Maksymenko, D. Panchenko, O. Kalynychenko, O. Turuta. Text Classification Benchmark of Ukrainian News Dataset.** The problem of the lack of high-quality benchmark datasets for Ukrainian text classification is considered. A framework for creating a simple dataset with minimal annotation effort is proposed. In addition, a corpus of Ukrainian news suitable for comparative model evaluation is created. Using this dataset, a benchmarking study of modern transformer-based models (mBERT, Slavic BERT, ukr-RoBERTa, ukr-ELECTRA, XLM-R) as well as a baseline NB-SVM model is conducted under various training settings. The results show that ukr-RoBERTa, ukr-ELECTRA, and XLM-R achieve the best performance. XLM-R generally performs better on long texts, whereas ukr-RoBERTa is more effective for shorter sequences.

UKRAINIAN LANGUAGE PROCESSING, TEXT CLASSIFICATION, TRANSFORMERS, TEXT DATASET

### Вступ

Останнім часом обробка природної мови пережила фазу швидкого розвитку, аналогічну до революції комп'ютерного зору у 2010-х роках. Цей прогрес здебільшого пов'язаний з розвитком архітектур Transformer [2] та BERT [3]. Таких успіхів вдалося досягти переважно завдяки механізму трансферного навчання.

Однак, попереднє навчання таких моделей потребує великої кількості даних та обчислювальної потужності. В результаті більшість найкращих попередньо навчених архітектур існують лише для найпопулярніших мов, таких як англійська, китайська тощо. Для української мови є лише кілька подібних моделей, зокрема ukr-RoBERTa [4] та ukr-ELECTRA [5]. Однак обидві ці моделі не мають належного оцінювання: ukr-ELECTRA орієнтована тільки на POS-теги та завдання NER, тоді як ukr-RoBERTa взагалі не має жодних показників, розрахованих на загальнодоступних наборах даних.

Ця відсутність різноманітних оцінок, яка зазвичай властива науковим роботам з NLP для англійської мови, є результатом недостатньої кількості загальнодоступних, правильно організованих та очищених наборів даних для української мови. Науковцями було

зроблено кілька спроб створити еталонний набір даних для найпоширенішого та найпростішого завдання обробки природної мови — класифікації тексту.

В [6] і [7] автори створюють набір даних для аналізу тональності відгуків про готелі, який може бути еталоном для текстових класифікаторів української мови. Вони також відзначають відсутність українських даних та складність їх збору. Наприкінці автори доповнюють свій набір російськими текстами, перекладеними українською за допомогою алгоритмів машинного перекладу.

Інший відомий ресурс з українськими наборами даних [8] пропонує велику колекцію нерозмічених даних, а також набори даних та попередньо навчені моделі для розпізнавання іменованих сутностей (NER). Проте в ньому відсутні українські набори даних для завдань «послідовність до одного».

В цій роботі запропоновано альтернативний підхід до розв'язання задач обробки української мови. Він полягає у використанні багатомовних моделей. Існують два трансформери, навчені різними мовами, включаючи українську: Multilingual BERT [3] та XLM-R [9]. Такі моделі зазвичай навчаються на комбінованому корпусі, що включає тексти на десятках мов (точніше, mBERT і XLM-R навчені на колекції

104 найбільших різномовних наборах даних Вікіпедії). Після цього вони оцінюються за допомогою міжмовних бенчмарків, таких як XNLI [10]. Однак, XNLI не включає української мови, тому ці моделі не тестувалися конкретно на українських даних. Для тестування розробляється загальний фреймворк, який дозволяє відносно легко збирати велику кількість українських даних та не потребує ручної анотації.

Після цього застосовується методологія для створення набору даних для класифікації новин (хоча її можна застосувати до кількох інших галузей). Цей набір даних використовується для оцінки та порівняння кількох трансформерів з відкритим вихідним кодом, які є доступними та застосовними для української мови. В останньому розділі аналізуються результати та створюються рекомендації щодо вибору потенційної моделі для аналогічних прикладних завдань.

Ми заохочуємо інших дослідників у цій галузі використовувати цей набір даних для подальшої оцінки своїх моделей.

## 1. Набір даних

Створено набір даних українських новин, зібраних з кількох джерел, перелічених у розділі 2.1. Фреймворк підготовки даних описаний в розділі 1.2. Даний набір даних використовується для бенчмаркінгу різних моделей класифікації тексту. Ці моделі або навчені, або тонко налаштовуються на задачі, описані у розділі 1.3.

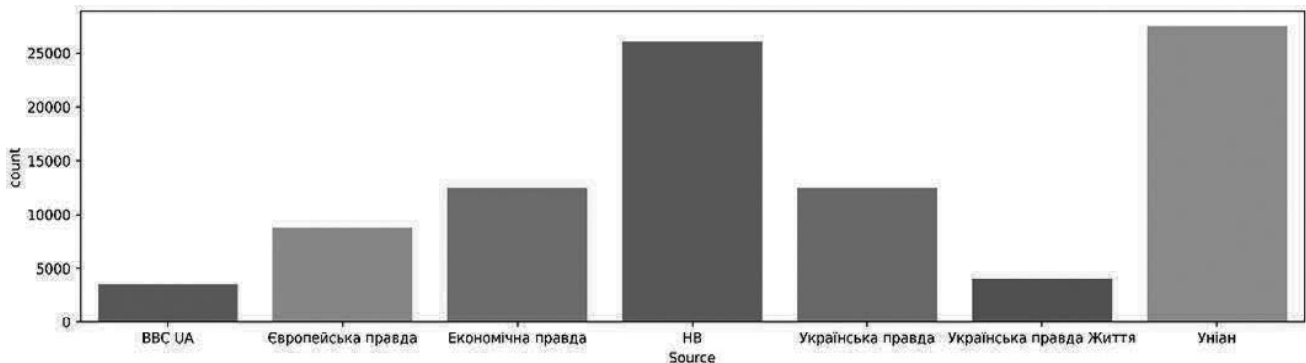


Рис. 1. Розподіл необроблених даних

Після цього використовуємо джерело даних як мету класифікації. Наслідки цього вибору та гіпотези про те, чого модель може навчитися з такої цільової змінної, розглядаються у розділі 2.3.

### 1.2. Підготовка даних

В необробленому наборі даних існує кілька аспектів, які потребують попередньої обробки та очищення. Навіть проста модель «мішок слів» [19], навчена на цьому отриманому наборі, досягає 0,9 бала за F1-мірою, в той час як більш складні підходи, засновані на глибокому навчанні, показують майже ідеальну точність після кількох епох навчання через численні неявні витоки даних [20].

Для досягнення цих результатів було застосовано наступний конвеєр для підготовки даних:

### 1.1. Джерела даних

Щоб створити досить великий бенчмарк-набір, потрібно зібрати і розмітити величезну кількість текстів. Також, якщо тестувати моделі за різних умов (наприклад, стиль тексту, довжина, поєднання різних мов тощо), потрібно створювати окреме маркування для кожного з обраних налаштувань. На жаль, збір даних є досить дорогим і складним процесом, особливо для української мови, оскільки вона має обмежену кількість джерел даних, а більшість легкодоступних медіаджерел (наприклад, повідомлення в соціальних мережах, новини тощо), містять значні домішки російських та англійських фрагментів.

Для вирішення цієї проблеми було розроблено конвеєр, який дозволяє збирати нерозмічений набір даних для класифікації тексту без особливих зусиль щодо підготовки даних. При необхідності, цей набір даних може розширювати тим самим способом.

Для цього було завантажено, проаналізовано та спеціальним чином підготовлено дані з кількох українських новинних сайтів: BBC News Україна [11], НВ [12], Українська правда [13], Економічна правда [14], Європейська правда [15], Українська правда Життя [16] та Уніан [17]. Розподіл зібраних даних показано на рис. 1. Повний набір необроблених даних складається з 94 994 текстів.

– Нормалізація пробілів. Початкові та кінцеві символи пробілів були видалені. Послідовності більш ніж одного пробільного символу були стиснуті.

– Видалення заголовка джерела. Усі згадки будь-якої назви джерела у будь-якій граматичній формі (наприклад, «BBC», «Бібісі» або «Служба новин BBC») були замінені спеціальним токеном [ДЖЕРЕЛО] як у назвах статей, так і в текстах. Була використана модифікована версія коректора друкарських помилок Norvig's [21] для усунення неправильного написання заголовків джерел даних.

– Видалення дублікатів. Для кожного кластера дубльованих або схожих текстів залишався лише один екземпляр. Найбільш яскравими прикладами таких кластерів є шаблонні статті про курси валют,

нові випадки коронавірусу в Україні, або новини з фронту, які відрізняються один від одного лише цифрами та дрібними деталями (наприклад, список регіонів червоної зони). Очевидно, що модель могла б запам'ятовувати такі тексти замість вивчення їхньої семантики. Таким чином, усі подібні випадки були розцінені як витік даних.

– Очищення мови. Мова текстів визначалася автоматично за допомогою langdetect [22]. Усі документи неукраїнського походження були видалені з набору даних.

– Витоки шаблонних даних. Було проведено напівавтоматичний пошук типових патернів, що зустрічаються лише у текстах із певного джерела даних, однозначно ідентифікуючи його формою. Усі такі випадки були видалені з набору даних.

Щоб виконати пошук витоку даних шаблону, спочатку ми створили матрицю TF-IDF [23] всіх термінів та найпопулярніших біграм у наборі даних. Потім, ми використовували вибір функції хі-квадрат, щоб знайти 20 найкращих токенів для кожного класу. Всі вибрані токени були перевірені вручну, і деякі типові речення або фрази, що їх містять, були ідентифіковані як витік даних та очищені. Цей процес повторювався кілька разів, поки в топ-20 не потрапив жоден підозрілий токен.

Найбільш яскравими прикладами таких витоків даних шаблону були клікбейтні фрази (наприклад, «Відвідайте наш канал YouTube для більш детальної інформації») та посилання (наприклад, «Зображення надані...»). Приклади таких витоків з контекстом показано рис. 2. Кожен тип витоку даних шаблону був

або замаскований токеном [SOURCE], або видалено іншим чином.



Рис. 2. Приклади витоку даних через шаблонні фрази у статтях BBC та УНІАН відповідно

Хоча ми визнаємо, що такі зміни спотворюють природний розподіл даних, ми вважаємо їх необхідними, щоб зробити задачу репрезентативною для реальних умов, де моделі повинні вивчати складні семантичні зв'язки, а не шукати набір заздалегідь відомих підказок.

Після вищезазначеної попередньої обробки отриманий набір даних складається з 82 554 текстів (близько 12 000 текстів були повністю виключені з різних причин).

Далі цей набір даних поділено на навчальну та тестову вибірку. Повний навчальний набір містить 57789 заголовків та текстів. Тестовий набір складається з 24765 зразків. Підмножини мають аналогічний цільовий розподіл змінних (рис. 3).

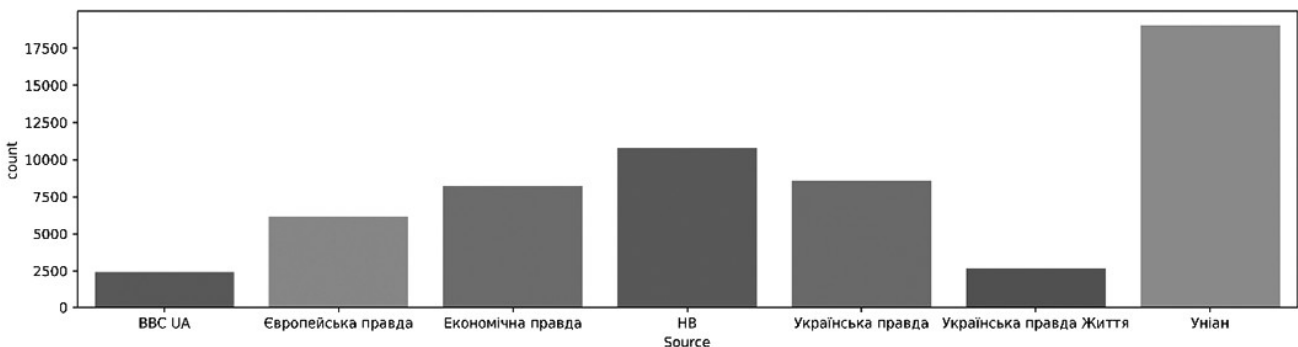


Рис. 3. Розподіл класів навчального набору даних

### 1.3. Постановка задачі

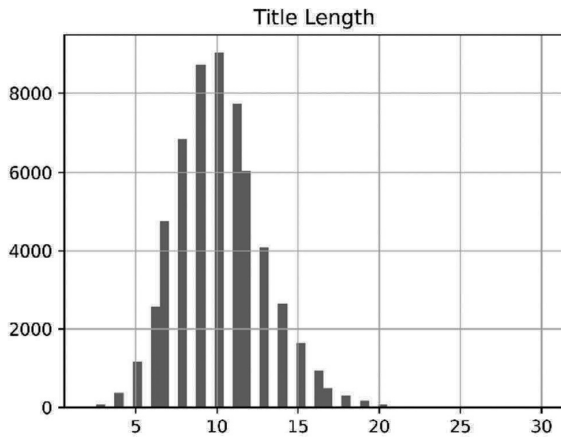
У якості цілі класифікації використано джерело даних. Таким чином, очікується, що моделі навчатимуться суміші стилістичної класифікації або ідентифікації джерела (оскільки кожне джерело даних має унікальні стилістичні та текстові атрибути), а також тематичному моделюванню (оскільки деякі новинні веб-сайти в нашому наборі даних зосереджені на певних наборах тем, хоча будь-яка велика тема представлена мінімум кількома джерелами).

Ця проблема мультикласової класифікації оцінюється за допомогою макроусередненої оцінки  $F_1$ :

$$F_1 = \frac{1}{n} \sum_{i=0}^n \frac{2 \times Precision_i \times Recall_i}{Precision_i + Recall_i}$$

З формули видно, що другорядні класи штрафуються в тій же мірі, що і основні, тим самим заставляючи моделі боротися з дисбалансом класів. Набір даних спеціально зроблений незбалансованим, щоб створити більш складну та схожу на реальний світ задачу. Виходячи з цього, пропонується тестувати моделі в двох режимах.

Перший варіант — навчати моделі або на повному навчальному наборі, або на маленькій підмножині, що складається з 8256 зразків. Це дозволяє моделювати продуктивність моделі при обмеженнях навчальних даних: така ситуація може виникнути в реальних додатках, коли вартість маркування даних висока, наприклад, коли для розмітки даних потрібні вузькоспеціалізовані фахівці [18].



Другий варіант — навчати моделі або на повних статтях, або тільки на заголовках. Таким чином, ми можемо визначити продуктивність моделей на текстах різної довжини.

Розподіл довжини тексту та заголовка показано на рис. 4. Тут і далі тексти та заголовки також називаються довгими та короткими текстами відповідно.

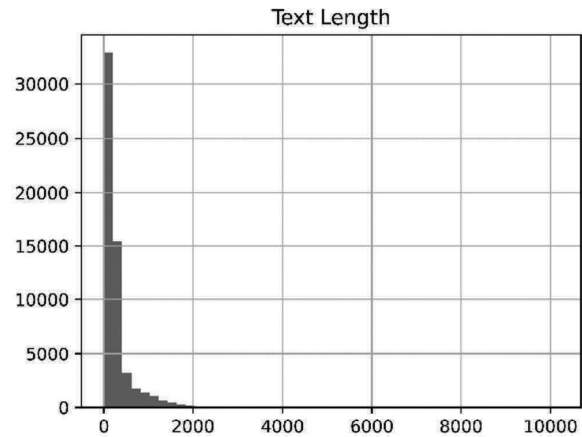


Рис. 4. Архітектура CFFN

Підсумовуючи, можна сказати, що за цих двох умов є чотири навчальні режими. І кожна модель тестується на кожному варіанті.

## 2. Різновид моделей

В цій роботі тестуються п'ять різних моделей-трансформерів: багатомовну BERT [3], слов'янську BERT [24], ukr-RoBERTa [4], ukr-ELECTRA [5] та XLM-R [9]. Розділи 3.1-5 стисло описують ці моделі.

Також навчається простий, але надійний базовий варіант - NB-SVM на функціях TF-IDF, реалізованих згідно [25]. Його результати можна порівняти з результатами трансформерів при деяких налаштуваннях. Ці результати більш детально обговорюються у розділі 5.

### 2.1. Багатомовна BERT

Багатомовна BERT (mBERT) є розширенням класичної BERT, навченим на об'єднаних корпусах для 104 мов, включаючи українську. Її можна використовувати для трансферного навчання завдань, в тому числі для класифікації новин. Зазвичай вона показує набагато нижчу продуктивність, ніж моделі, попередньо навчені для кожної окремої мови. Для наших експериментів було спеціально використано необроблену mBERT-базу.

### 2.2. Слов'янська BERT

Слов'янська BERT є результатом неконтрольованого трансферного навчання від mBERT на об'єднаному корпусі сторінок Вікіпедії, написаних кількома слов'янськими мовами. Хоча цей корпус не включає українську мову, можливо, попереднє навчання на тих самих морфемах може покращити його

продуктивність для розглянутого у статті українського набору даних.

### 2.3 ukr-RoBERTa

ukr-RoBERTa — це версія моделі RoBERTa [26], попередньо навчена спеціально на великомасштабному корпусі, який складається з української Вікіпедії, дедуплікованого набору даних Ukrainian OSCAR [27] та внутрішнього набору даних youscan, зібраного із соціальних мереж. Автори не повідомляють про будь-які результати для цієї моделі в загальнодоступних тестах, хоча й згадують, що вони отримали 2-відсоткове покращення оцінки F1 у своїх внутрішніх наборах даних порівняно з mBERT. Вимірювання продуктивності цієї моделі у загальнодоступному наборі даних має вирішальне значення для її ефективного використання у прикладній науці.

### 2.4 ukr-ELECTRA

ukr-ELECTRA — це модель на основі архітектури ELECTRA [28], попередньо навчена на сторінках української Вікіпедії та дедуплікованому наборі даних українського OSCAR.

Очікується, що вона повинна працювати краще, ніж ukr-RoBERTa, тому що підхід ELECTRA загалом перевершує RoBERTa по більшості завдань. Проте українську версію ELECTRA було попередньо навчено на меншому наборі даних, тому, як показують наші експерименти, вони порівнюються інакше, ніж їхні англійські аналоги.

### 2.5 XLM-R

XLM-R — це модель на основі RoBERTa, яка навчається так само, як mBERT. XLM-R — єдина модель, яка

має версію попередньо вивчених ваг з відкритим вихідним кодом для версії з великою архітектурою. Щоб знайти найкращу модель класифікації текстів для українських текстів, ми використовуємо цю версію замість базової під час наших експериментів, оскільки очікується, що вона дасть максимальну продуктивність для наступних завдань.

### 3. Експерименти

Було проведено серію з чотирьох експериментів для кожної моделі:

- малий навчальний набір, навчання лише на заголовках;
- малий навчальний набір, навчання на повних текстах;
- великий навчальний набір, навчання лише на заголовках
- великий навчальний набір, навчання на повних текстах.

Для кожного з цих експериментів проводиться поверхнє налаштування. Ми обираємо планувальник швидкості навчання на невеликій валідаційній підбірці, відібраній із навчального набору. Після цього ми перенавчаємо модель з найкращими гіперпараметрами на всьому навчальному наборі перед поданням прогнозів.

Щоб порівняти моделі в реальних умовах, замість навчання кожної моделі однаковою кількістю кроків, ми навчаємо кожну модель фіксований час, який дорівнює 24 годинам на одному GPU P100.

Результати бенчмаркінгу наведені в таблиці нижче.

Таблиця 1  
Результати бенчмаркінгу використаних моделей

Модель	Короткий текст/ малий навчальний набір	Короткий текст/ великий навчальний набір	Повний текст/ малий навчальний набір	Повний текст/ великий навчальний набір
Базова NB-SVM	0.533	0.790	0.636	0.900
mBERT	0.626	0.853	0.685	0.910
Слов'янська BERT	0.620	0.840	0.708	0.907
ukr-RoBERTa	<b>0.675</b>	0.903	<b>0.745</b>	0.940
ukr-ELECTRA	0.623	0.909	0.721	0.948
XLM-R	0.624	<b>0.915</b>	0.689	<b>0.950</b>

### 4. Результати

Найвища оцінка серед усіх проведених експериментів у XLM-R, навченої на великій версії повнотекстового навчального набору: вона сягає 0,95 F1.

Далі наведено декілька цікавих спостережень:

– i mBERT, i Slavic BERT досить погано працюють із погляду оцінки F1. Хоча це й очікувалося для mBERT, проте, дивно, що Slavic BERT не показала ніякого підвищення точності, працюючи навіть гірше, ніж mBERT, у трьох налаштуваннях з чотирьох.

– ukr-RoBERTa демонструє значне покращення продуктивності порівняно з mBERT (5-6% для коротких текстів та 3-6% для довгих текстів). Вона також показує менший розрив між налаштуваннями короткого та довгого тексту. Ми пов'язуємо це з тим, що вона була навчена на наборі даних, який включає скопійовані повідомлення з соціальних мереж, які, як правило, коротше, ніж інші типи текстів.

– ukr-ELECTRA в середньому показує трохи гірші метрики, будучи менш точною на коротких текстах і трохи точнішою на довгих.

– XLM-R загалом перевершує всі інші моделі на довгих текстах, але водночас демонструє значно нижчу продуктивність на коротких текстах. Варто зазначити, що XLM-R має 24 енкодерні блоки замість 12, тому потребує майже втричі більшої пропускну здатності пам'яті та має вищу затримку порівняно з іншими трансформерними моделями, що брали участь у бенчмаркінгу.

Попри очікування, базова модель NB-SVM демонструє досить високий показник F1 у режимі навчання на великому наборі даних. Вона поступається середній трансформерній моделі лише на 7% у режимі коротких текстів, і показує майже ті самі результати, як mBERT і SlavicBERT у режимі довгих текстів. Ми припускаємо, що це зумовлено тим, що під час навчання моделей на малому наборі даних ефективність підходу трансферного навчання є значно вищою, ніж у випадку великого набору даних.

Ці результати показують, що ukr-RoBERTa може бути кращою моделлю для коротких текстів, тоді як XLM-R або ukr-ELECTRA - найкращий вибір для більш об'ємних. Варто зазначити, що модель NB-SVM, яка не вимагає ні графічного процесора для навчання, ні дорогого обладнання для прогнозування в реальному часі, досягає порівняної продуктивності, якщо набір навчальних даних є досить великим. Результат всього на 5% нижче кращої моделі, а на впровадження та навчання займає п'ятнадцять хвилин, що є прийнятним у багатьох прикладних випадках.

### Висновки

У рамках цієї статті розроблено просту та ефективну послідовність дій, яка дозволяє створити набір даних для класифікації тексту з мінімальними витратами.

Використовуючи цей підхід, було створено набір даних для класифікації новин, що складається з майже 60 тисяч навчальних прикладів та дозволяє

проводити бенчмаркінг моделей, використовуючи кілька налаштувань для більш глибокого розуміння плюсів та мінусів моделей.

Набір даних розміщено на платформі Kaggle [1] і є доступним для бенчмаркінгу нових алгоритмів машинного навчання для української мови.

В цій роботі було оцінено кілька наявних моделей з відкритим вихідним кодом на вищезазначеному наборі даних у стандартизованих умовах експерименту. Результати показують, що ukr-RoBERTa та ukr-ELECTRA є найпродуктивнішими моделями середнього розміру, тоді як XLM-R демонструє кращі результати на довгих текстах за відсутності обчислювальних обмежень.

Водночас, NB-SVM демонструє порівнянні результати. Це спостереження, а також той факт, що міжмова модель є однією з найефективніших, означає, що попередньо навченим трансформерним моделям для української мови ще доведеться пройти довгий шлях. Збір великих наборів даних для неконтрольованого попереднього навчання та попереднього навчання великих моделей (наприклад, RoBERTa-large) здаються найбільш перспективними напрямками розвитку.

#### Список літератури:

- [1] <https://www.kaggle.com/c/ukrainian-news-classification/>
- [2] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17). Curran Associates Inc., Red Hook, NY, USA, 6000–6010.
- [3] Jacob Devlin, Ming-Wei Chang, Kenton Lee, & Kristina Toutanova. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.
- [4] Vitalii Radchenko. We Trained the Ukrainian Language Model. <https://youscan.io/blog/ukrainian-language-model/>
- [5] Stefan Schweter, Ukrainian ELECTRA model <https://github.com/stefanit/ukrainian-electra> <https://doi.org/10.5281/zenodo.4267880>
- [6] Babenko, Dmytro. Determining sentiment and important properties of Ukrainian language user reviews : Master Thesis : manuscript rights / Dmytro Babenko ; Supervisor Vsevolod Dyomkin ; Ukrainian Catholic University, Department of Computer Sciences. — Lviv : [s.n.], 2020. — 35 p. : ill.
- [7] Babenko, D., & Dyomkin, V. (2019). Determining Sentiment and Important Properties of Ukrainian Language User Reviews. <http://ceur-ws.org/Vol-2566/MS-AMLV-2019-paper39-p106.pdf>
- [8] NER annotation corpus <https://lang.org.ua/en/corpora/>
- [9] Alexis Conneau and Kartikay Khandelwal and Naman Goyal and Vishrav Chaudhary and Guillaume Wenzek and Francisco Guzmán and Edouard Grave and Myle Ott and Luke Zettlemoyer and Veselin Stoyanov (2019). Unsupervised Cross-lingual Representation Learning at Scale. CoRR, abs/1911.02116.
- [10] Conneau, A., Rinott, R., Lample, G., Williams, A., Bowman, S., Schwenk, H., & Stoyanov, V. (2018). XNLI: Evaluating Cross-lingual Sentence Representations. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics.
- [11] <https://www.bbc.com/ukrainian>
- [12] <https://nv.ua/>
- [13] <https://www.pravda.com.ua/>
- [14] <https://www.epravda.com.ua/>
- [15] <https://www.eurointegration.com.ua/>
- [16] <https://life.pravda.com.ua/>
- [17] <https://www.unian.ua/>
- [18] Shen, Ying et al. “Improving Medical Short Text Classification with Semantic Expansion Using Word-Cluster Embedding.” ArXiv abs/1812.01885 (2018): n. pag.
- [19] Zhang, Yin & Jin, Rong & Zhou, Zhi-Hua. (2010). Understanding bag-of-words model: A statistical framework. International Journal of Machine Learning and Cybernetics. 43-52. <https://doi.org/10.1007/s13042-010-0001-0>
- [20] Kaufman, Shachar & Rosset, Saharon & Perlich, Claudia. (2011). Leakage in Data Mining: Formulation, Detection, and Avoidance. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 6. 556-563. <https://doi.org/10.1145/2020408.2020496>
- [21] Peter Norvig. How to Write a Spelling Corrector. url: <http://norvig.com/spellcorrect.html>.
- [22] Shuyo, N. (2010). Language Detection Library for Java.
- [23] (2011) TF-IDF. In: Sammut C., Webb G.I. (eds) Encyclopedia of Machine Learning. Springer, Boston, MA. <https://doi.org/10.1007/978-0-387-30164-8>.
- [24] Arkhipov, A. (2019). Tuning Multilingual Transformers for Language-Specific Named Entity Recognition. In Proceedings of the 7th Workshop on Balto-Slavic Natural Language Processing (pp. 89–93). Association for Computational Linguistics.
- [25] Wang, C. (2012). Baselines and Bigrams: Simple, Good Sentiment and Topic Classification. In Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers) (pp. 90–94). Association for Computational Linguistics.
- [26] Yinhan Liu and Myle Ott and Naman Goyal and Jingfei Du and Mandar Joshi and Danqi Chen and Omer Levy and Mike Lewis and Luke Zettlemoyer and Veselin Stoyanov (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. CoRR, abs/1907.11692.
- [27] Ortiz Suárez, P., Sagot, B., & Romary, L. (2019). Asynchronous Pipeline for Processing Huge Corpora on Medium to Low Resource Infrastructures. In 7th Workshop on the Challenges in the Management of Large Corpora (CMLC-7). Leibniz-Institut für Deutsche Sprache.
- [28] Kevin Clark, Minh-Thang Luong, Quoc V. Le, & Christopher D. Manning (2020). ELECTRA: Pre-training Text Encoders as Discriminators Rather Than Generators. In International Conference on Learning Representations

Received (Надійшла) 15.02.2026

Accepted for publication (Прийнята до друку) 05.03.2026

Publication date (Дата публікації) 27.03.2026



І. О. Лещинська

ХНУРЕ, м. Харків, Україна, [iryna.leshchynska@nure.ua](mailto:iryna.leshchynska@nure.ua), ORCID iD: 0000-0002-8737-4595

## УЗАГАЛЬНЕНА СИМВОЛЬНА МЕНТАЛЬНА МОДЕЛЬ РІШЕННЯ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ ДЛЯ ЗОВНІШНІХ КОРИСТУВАЧІВ

Розглянуто проблему побудови ментальних моделей рішень інтелектуальних систем для зовнішніх користувачів, вирішення якої потребує побудови прозорих пояснень, адаптованих до предметної області та рівня підготовки користувача. Запропоновано узагальнену тривірневу символічну ментальну модель, що включає концептуальний рівень абстракцій рішень, рівень пояснень на основі каузальних правил і базу фактів, які об'єднані функцією вертикальної композиції, що забезпечує можливість керованої деталізації пояснень. Розроблено алгоритм автоматизованої побудови узагальненої моделі на основі доменної онтології, каузальних послідовностей міркування та даних користувачів, який забезпечує персоналізовані пояснення з урахуванням рівня підготовки цих користувачів. Модель забезпечує можливість формування пояснень для новачків, користувачів середнього рівня та експертів, що дає змогу узгоджувати глибину деталізації з когнітивними можливостями аудиторії та вимогами до верифікації рішень у високоризикових предметних областях, включаючи медичну діагностику, фінансові системи підтримки рішень та інші критичні застосування, де необхідні як інтерпретованість, так і відтворюваність ланцюжків міркувань.

УЗАГАЛЬНЕНА СИМВОЛЬНА МЕНТАЛЬНА МОДЕЛЬ, ПОЯСНЮВАНИЙ ШТУЧНИЙ ІНТЕЛЕКТ, ДОМЕННІ ОНТОЛОГІЇ, КАУЗАЛЬНІ ПРАВИЛА, БАЗА ФАКТІВ, ПЕРСОНАЛІЗОВАНІ ПОЯСНЕННЯ

**I. O. Leshchynska. Generalised symbolic mental model of an intelligent system decision for external users.** The paper addresses the problem of constructing mental models of intelligent system decisions for external users, which requires transparent explanations adapted to the application domain and the user's expertise level. A generalised three-layer symbolic mental model is proposed, comprising a conceptual layer of decision abstractions, an explanation layer based on causal rules, and a fact base, which are linked by a vertical composition function that enables controlled explanation granularity. An algorithm for automated construction of the generalised model is developed using a domain ontology, causal reasoning paths, and user data, providing personalised explanations tailored to users' expertise levels. The model supports explanation modes for novices, intermediate users, and experts, aligning the depth of detail with cognitive capabilities and verification requirements in high-risk domains where trust in intelligent systems is essential.

GENERALIZED SYMBOLIC MENTAL MODEL, EXPLAINABLE ARTIFICIAL INTELLIGENCE, DOMAIN ONTOLOGIES, CAUSAL RULES, FACT BASE, PERSONALIZED EXPLANATIONS

### Вступ

Сучасні інтелектуальні системи використовують складні непрозорі моделі формування результатів, що потребують розробки підсистем пояснень для підвищення довіри зовнішніх користувачів до отриманих рішень. Побудова підсистем пояснень виконується в рамках наукового напрямку пояснювального штучного інтелекту. Розробка систем пояснювального штучного інтелекту потребує побудови формального опису сприйняття рішень інтелектуальних систем користувачами. Вирішення даної проблеми пов'язано із створенням символічних структур для побудови ментальних моделей рішень інтелектуальних систем. Такі структури мають забезпечувати адаптацію ментальної моделі до індивідуальних когнітивних особливостей зовнішніх користувачів-непрофесіоналів. Зовнішні користувачі – це особи, які не є розробниками системи та не володіють глибокими технічними знаннями про механізми роботи інтелектуальної системи, наприклад пацієнти, чиї дані використовуються в медичних діагностичних системах, клієнти банків у фінансових системах підтримки рішень, студенти, що використовують в освітні рекомендаційні системи.

Традиційні неперсоналізовані підходи до побудови пояснень не враховують індивідуальні ментальні

моделі користувачів та рівень їхньої експертизи, що утруднює розуміння пояснень зовнішніми користувачами, особливо коли ці користувачі мають різний фаховий бекграунд та когнітивні здібності. Також побудова ментальних моделей виконується із залученням експертів на основі трудомістких структурованих інтерв'ю. Наслідком використання трудомістких інтерв'ю при побудові опису сприйняття рішень системи людиною стає формування специфічних для предметної області ментальних моделей, які не можуть бути повторно використані в інших доменах через відсутність уніфікованої структури представлення знань. Побудова ментальних моделей експертами через структуровані інтерв'ю та картування концептів потребує значних витрат часу та характеризується низькою узгодженістю між різними експертами внаслідок суб'єктивної інтерпретації когнітивних структур користувачів.

Для побудови інтегрального опису ментальних моделей потрібно сформувані символічну архітектуру, яка явно виділяє концептуальний шар опису абстракцій рішень, пояснювальний шар каузальних правил виведення та базу знань доменних фактів, інтегруючи їх через онтологічний опис.

Внаслідок відсутності інтегрального опису ментальних моделей і відповідної відсутності персоналізованих пояснень знижується довіра до рішень інтелектуальних систем і можуть виникати обмеження при практичному застосуванні таких систем у високоризикових предметних областях.

Відповідно, проблема розробки узагальненого символічного опису ментальних моделей рішень інтелектуальної системи для персоналізації сприйняття пояснень є актуальною.

Серед існуючих підходів до побудови пояснень і до побудови ментальних моделей в рамках пояснень доцільно виділити три напрямки досліджень: пояснення на основі важливості ознак; нейро-символьний підхід; пояснення з використанням ментальних моделей користувачів інтелектуальних систем.

Перший напрямок охоплює методи пояснювального штучного інтелекту, які розраховують внесок окремих ознак у рішення моделі машинного навчання, яка використовується в інтелектуальній системі. В рамках даного напрямку в роботі [1] розроблено метод SHAP, що використовує значення Шеплі з кооперативної теорії ігор для обчислення справедливого розподілу внеску кожної ознаки в рішення моделі через принцип адитивності вкладів. В дослідженні [2] запропоновано метод LIME, який формує пояснення через локальні лінійні апроксимації в околиці конкретного рішення складних моделей, зокрема глибоких нейронних мереж. В [3] проведено порівняльний аналіз методів SHAP та LIME у контексті біомедичних застосувань та показано їхню залежність від типу моделі машинного навчання та характеристик даних, що обмежує узгодженість пояснень.

У монографії [4] систематизовано підходи до забезпечення інтерпретованості машинного навчання, включаючи методи для ранжування ознак за їх внеском в модель на всьому наборі даних та локальні методи для генерації контрфактичних сценаріїв, які відображають мінімальні зміни вхідних даних, необхідні для зміни рішення моделі. У огляді [5] класифіковано методи забезпечення глобальної та локальної пояснюваності. В [6] проведено аналіз обмежень методів забезпечення пояснюваності при клінічних застосуваннях. Дослідження показало суттєвий розрив між технічними метриками для оцінки пояснюваності в системах машинного навчання та реальним розумінням клініцистами механізмів роботи моделі.

У підсумку ці методи пояснювального штучного інтелекту орієнтовані на побудову неперсоналізованих пояснень без урахування індивідуальних ментальних моделей користувачів з різним рівнем експертизи.

Другий напрямок використовує онтології на нейро-символічному підході для представлення знань у предметних областях та їх інтеграції з методами машинного навчання. В [7] дослідили роль онтологій та

показали, що онтології забезпечують структуровані фреймворки знань для забезпечення інтерпретованості процесів прийняття рішень в інтелектуальних системах. В дослідженні [8] проаналізовано множинні ролі онтологій у пояснювальному штучному інтелекті, в тому числі для формалізації доменних концептів, міркування для заповнення прогалін у знаннях з тим, щоб забезпечити розробку людиноцентричних пояснювальних систем.

В роботі [9] проведено систематичний огляд нейро-символьного штучного інтелекту в охороні здоров'я на основі аналізу біомедичних досліджень та показано потенціал інтеграції символічного міркування з глибоким навчанням для підвищення інтерпретованості рішень в клінічних застосуваннях, де зрозумілість міркування є критично важливою для прийняття рішення клініцистами. В [10] розроблено нейро-символьну систему, що об'єднує великі мовні моделі з експертною системою на основі правил для витягування структурованих даних із текстових радіологічних звітів, що забезпечує аудит ланцюжків виведення від кінцевих міток до початкових токенів моделі через отриману явну трасу міркування.

В дослідженні [11] розробили онтологію людських пояснень щодо поведінки робітників, що дає можливість роботам генерувати контекстно-релевантні пояснення своїх дій для людей – спостерігачів. В [12] запропонували онтологічний фреймворк, що інтегрує машинне навчання та імовірнісне планування в межах міркування на основі здорового глузду, де онтології використовуються для формалізації контексту активності користувача для підтримки рішень щодо допоміжних утручань.

Ключове обмеження даного підходу полягає у використанні специфічних онтологій для окремих галузей застосування без розробки узагальненої незалежної від предметної області схеми, що могла б бути повторно використана для нових предметних областей без повного переписування онтологічної структури.

Третій напрямок орієнтований на моделювання взаємодії людини з інтелектуальними системами з використанням ментальних моделей. Останні описують, як користувачі формують внутрішні представлення процесу роботи та рішень інтелектуальних систем на основі отриманих пояснень. В [13] дослідили ментальні моделі в рамках пояснювального штучного інтелекту, включивши призначення пояснення, аудиторію користувачів та мову представлення, з акцентом на зрозумілість пояснення. Тобто, дослідили ментальні моделі як основу людиноцентричного підходу до пояснень. В дослідженні [14] проведено огляд літератури з соціальних наук щодо процесу формування пояснень людьми в повсякденній комунікації та виявлено істотний розрив між методами

пояснювального штучного інтелекту, що фокусуються на повноті представленої інформації, та когнітивними теоріями людських пояснень, що підкреслюють селективність та контекстуальність представлення інформації.

В роботі [15] досліджували ефект прозорості штучного інтелекту на формування ментальних моделей користувачів та виявили, що відображення рівнів впевненості моделі в своїх передбаченнях зменшує надмірну довіру користувачів у медичних діагностичних системах. В роботі [16] провели дослідження щодо впливу інформації про відмови системи на рівень довгострокової довіри користувачів. Користувачі, що отримували чіткі пояснення причин відмов, демонстрували вищі показники довіри через кілька місяців порівняно з користувачами без досвіду відмов.

В [17] досліджено поведінкові метрики довіри в системах штучного інтелекту, виявивши ключові індикатори, зокрема частоту прийняття рекомендацій системи, час між отриманням рекомендації та дією користувача, час освоєння нового функціоналу та швидкість відновлення довіри після помилок системи. В [18] провели кількісне дослідження впливу інтерактивних та контекстуальних пояснень на довіру користувачів, ра продемонстрували, що залученість користувачів у процесу формування пояснень та розуміння контексту відіграють додаткову роль у формуванні довіри крім забезпечення прозорості механізмів роботи інтелектуальної системи.

Однак даний підхід має обмеження, пов'язане із відсутністю формальних схем побудови символічних ментальних моделей із підтримкою механізмів деталізації від абстрактних концептів до конкретних фактів предметної області з використанням формалізованих правил логічного виведення.

Таким чином, існуючі підходи до побудови ментальних моделей рішень інтелектуальних систем для зовнішніх користувачів мають ряд обмежень щодо формування символічного представлення таких моделей, які перешкоджають побудові персоналізованих пояснень для користувачів-непрофесіоналів, оскільки останні потребують інтерпретованих пояснень рішень у зрозумілій формі без деталізації механізмів роботи алгоритмів машинного навчання. Зазначений недолік обґрунтовує необхідність розробки узагальненої символічної ментальної моделі рішення інтелектуальної системи для зовнішніх користувачів.

### 1. Постановка задачі

Метою дослідження є розробка узагальненої символічної ментальної моделі рішення інтелектуальної системи для зовнішніх користувачів з явним розділенням концептуального шару, пояснювального шару та фундаментальної бази знань для скорочення витрат часу на побудову ментальних моделей і можливості

повторного використання для інших предметних областей за рахунок незалежності від домену схеми онтологічного представлення.

Для досягнення мети вирішуються такі задачі:

- розробити трирівневу символічну ментальну модель рішення інтелектуальної системи для зовнішніх користувачів із концептуальним шаром для представлення абстракцій рішень через класи онтологій, пояснювальним шаром для представлення каузальних зв'язків між властивостями об'єктів онтологій, базою знань для представлення фактів предметної області;

- розробити алгоритм автоматизованої побудови трирівневої символічної ментальної моделі рішення інтелектуальної системи.

### 2. Тривірнева символічна ментальна модель рішення інтелектуальної системи для зовнішніх користувачів

Розроблена модель використовує чотири типи компонентів: класи онтологій, властивості об'єктів, властивості даних та індивіди онтологій.

Тривірнева архітектура моделі має вигляд:

$$M = (L_C, L_E, L_K, \Phi), \quad (1)$$

де  $L_C$  – концептуальний шар високорівневих абстракцій рішень,  $L_E$  – пояснювальний шар каузальних правил виведення,  $L_K$  – база знань фактів предметної області, а  $\Phi$  – функція вертикальної композиції між рівнями для підтримки деталізації пояснень.

Концептуальний шар містить множину класів онтологій:

$$L_C = \{C_1, C_2, \dots, C_n\}, \quad (2)$$

де кожен клас  $C$  представляє категорію рішень інтелектуальної системи з формальним визначенням необхідних та достатніх умов засобами дескриптивної логіки, що дає можливість відносити індивідів до класів на основі їхніх властивостей.

Пояснювальний шар містить множину правил виведення:

$$L_E = \{r_1, r_2, \dots, r_m\}, \quad (3)$$

де кожне правило  $r_j$  представлено традиційною імплікацією:

$$r_j: \text{антецедент}_j \Rightarrow \text{консеквент}_j, \quad (4)$$

де антецедент містить кон'юнкцію атомів із порівняннями властивостей даних з пороговими значеннями з використанням булевих операторів, а консеквент містить або твердження про належність індивіда до класу концептуального шару або ж додаткові властивості об'єкта для формування каузального шляху міркування.

База знань містить множину екземплярів онтологій

$$L_K = \{i_1, i_2, \dots, i_k\}, \quad (5)$$

де кожен екземпляр  $i$  представляє конкретного користувача системи з типізованими властивостями даних,

що відображають значення атрибутів користувача, які відповідають рішенням інтелектуальної системи.

Функція вертикальної композиції  $\Phi$  відображає кожен концепт рішення з концептуального шару на множину пар із каузальних правил виведення пояснювального шару та фактів фундаментальної бази знань:

$$\Phi: L_C \rightarrow 2^{L_E \times L_K}. \quad (6)$$

Вирази (1) – (6) обґрунтовують належність індивіда до класу концепту, забезпечуючи механізм деталізації від абстрактних концептів через каузальні правила до конкретних значень атрибутів користувача, що забезпечує можливість адаптації з урахуванням рівня експертизи останнього.

У сукупності вертикальна композиція рівнів розробленої моделі забезпечує формування повних ланцюжків обґрунтування від високорівневих концептів рішень через каузальні правила до конкретних фактів про атрибути користувачів через функцію обґрунтування та механізм деталізації з адаптацією до рівня експертизи користувача.

Обчислення функції  $\Phi$  базується на аналізі ланцюжків виведення для онтологій, отриманих під час віднесення індивідів до концептуальних класів через застосування правил виведення до властивостей індивідів. Такий підхід забезпечує прозорість процесу прийняття рішення інтелектуальною системою.

Механізм деталізації дозволяє переходити від концептуального шару через пояснювальний шар до бази знань відповідно до когнітивних потреб користувача та рівня його підготовки і практичних задач, які він вирішує.

На концептуальному рівні користувач може бачити лише високорівневий концепт рішення без деталей обґрунтування, що мінімізує когнітивне навантаження для користувачів-новачків.

На пояснювальному рівні користувач може деталізувати концепт для перегляду каузальних правил виведення, які обґрунтовують рішення, що орієнтовано на потреби користувачів середнього рівня експертизи.

На рівні бази знань розробленої моделі користувач може деталізувати правило для перегляду конкретних значень атрибутів свого профілю, які задовольняють умові правила. Така можливість забезпечує повну прозорість для експертів, які потребують верифікації всіх деталей міркування у процесі формування рішення в інтелектуальній системі.

### 3. Алгоритм побудови узагальненої символічної ментальної моделі рішення інтелектуальної системи

Алгоритм автоматизованої побудови узагальненої символічної ментальної моделі рішення інтелектуальної системи виконується для кожної нової предметної області застосування моделі через формування доменних класів без зміни архітектури.

Вхідними даними алгоритму є:

- доменна онтологія із ієрархією класів та відношеннями для відповідної предметної області;
- набір рішень інтелектуальної системи з описами категорій рішень та критеріїв класифікації;
- набір даних користувачів із значеннями атрибутів для заповнення бази знань;
- каузальні послідовності міркування від атрибутів до концептів рішень, отримані на основі експертних знань щодо предметної області.

На першому етапі алгоритму виконується витягання релевантних класів концептів рішень з доменної онтології. Використовується аналіз семантичної подібності між описами категорій рішень системи та класами доменної онтології. Для оцінки семантичної подібності використовують векторні представлення цих описів.

На другому етапі виконується формалізація каузальних шляхів шляхом їх опису з використанням правил виведення. Для кожного каузального шляху від атрибутів до концепту рішення формується правило виведення з антецедентом, що містить булеві умови на значення властивостей даних індивіда, та консеквентом, що містить твердження про належність індивіда до класу концептуального шару.

На третьому етапі виконується заповнення фундаментальної бази знань даними користувачів із типізованими властивостями даних, які відображають конкретні значення атрибутів, релевантних для рішення інтелектуальної системи.

На четвертому етапі виконується автоматизоване виведення через механізм логічного виведення онтологій для віднесення індивідів до класів концептів рішень на основі правил виведення та означень класів, що дає можливість сформулювати повну ментальну модель із ланцюжками виведення від концептів рішень до конкретних значень атрибутів користувачів.

Результатом виконання алгоритму є повна узагальнена символічна ментальна модель рішення інтелектуальної системи для зовнішніх користувачів із класифікованими індивідами, ланцюжками обґрунтування та обчисленими каузальними шляхами міркування.

Дана модель може бути використана для генерації персоналізованих пояснень відповідно до рівня експертизи кожного конкретного користувача.

При імплементації моделі адаптація до рівня експертизи може бути реалізована шляхом конфігурації інтерфейсу користувача відповідно до метаданих із профілю користувача щодо рівня його експертності та практичних задач, що він вирішує.

Новачкам може бути автоматично встановлено спрощений режим із відображенням лише концептуального шару без можливості деталізації для зменшення когнітивного навантаження та запобігання інформаційному перевантаженню.

Користувачам середнього рівня має бути встановлено стандартний режим із можливістю деталізації до пояснювального шару для вивчення каузальних правил обґрунтування рішення та відповідних ланцюжків міркувань.

Експертам встановлюється експертний режим з повною можливістю деталізації від концептуального шару через правила й до бази знань для верифікації конкретних значень даних та повного аудиту ланцюжків міркування системи.

#### 4. Експериментальна перевірка розробленої моделі

Експериментальна перевірка узагальненої символічної ментальної моделі здійснюється на наборі даних медичної діагностики діабету типу 2 шляхом порівняння зрозумілості пояснень із двома базовими підходами SHAP та LIME за метриками частотності термінів у веб-корпусі, семантичної узгодженості з медичною літературою та читабельності текстів пояснень, що забезпечує об'єктивну комплексну оцінку на відміну від суб'єктивних даних за результатами опитування користувачів. Набір даних із репозиторію UCI Machine Learning Repository, що містить записи пацієнток з атрибутами – рівень глюкози, індекс маси тіла, вік, артеріальний тиск, рівень інсуліну, спадкова функція діабету тощо. Для експерименту випадковим чином відібрано підмножину записів для того, щоб забезпечити збалансований розподіл класів у вхідному наборі даних. Для запропонованої моделі пояснення генеруються на основі ланцюжків виведення у середньому шарі. Такий підхід забезпечує побудову структурованих пояснень з явними каузальними зв'язками між атрибутами та рішеннями.

Для методів SHAP та LIME пояснення генеруються через стандартні бібліотеки shap та lime з параметрами за замовчуванням. Ці бібліотеки повертають числові коефіцієнти важливості кожної із ознак або оцінки її впливу на передбачення без явної каузальної інтерпретації. Зрозумілість пояснень оцінюється через три метрики, які використовують веб-пошук та лінгвістичний аналіз, що забезпечує об'єктивність перевірки і не потребує залучення людей – користувачів до опитувань.

Метрика TFWC визначає частотність термінів у веб-корпусі (Term Frequency in Web Corpus). Для розрахунку метрики для кожного текстового пояснення виявляються ключові терміни з використанням алгоритму RAKE (Rapid Automatic Keyword Extraction), що реалізований у бібліотеці rake-nltk. Даний алгоритм ідентифікує 3–5 найважливіших термінів у складі пояснення, наприклад, «high BMI», «diabetes risk».

В подальшому кожен ключовий термін подається у складі пошукового запиту до англійських медичних сайтів. В запиті використовується параметр site:\*.edu або site:diabetes.org або аналогічні. Для кожного

терміну у запиті фіксується кількість результатів пошуку, яка відображає його поширеність в мережі інтернет.

Метрика TFWC обчислюється як середнє логарифмічне значення кількості результатів за всіма ключовими термінами пояснення. Високі значення TFWC свідчать про поширеність термінів пояснення в медичній веб-літературі, що свідчить про зрозумілість вказаних слів для зовнішніх користувачів внаслідок знайомства з термінологією.

Метрика SCMC визначає семантичну узгодженість із медичним корпусом (Semantic Coherence with Medical Corpus). Для кожного текстового пояснення обчислюється векторне представлення через модель sentence-transformers all-MiniLM-L6-v2, що відображає текст у 384-вимірний вектор. Через веб-пошук знаходяться топ-10 результатів пошуку для ключових термінів пояснення з медичних сайтів. Ці результати також відображаються у векторі. Метрика SCMC обчислюється як середня косинусна подібність. Високі значення SCMC свідчать про семантичну узгодженість пояснення з реальною медичною літературою, відображають використання стандартної медичної термінології.

Метрика RS характеризує читабельність текстів пояснень (Readability Score). Для кожного текстового пояснення обчислюється індекс читабельності Flesch Reading Ease з використанням програмної бібліотеки textstat. Індекс читабельності оцінює складність тексту на основі довжини речень та кількості складів у словах. Значення FRE інтерпретуються за шкалою: 90–100 (дуже легко), 80–90 (легко), 70–80 (помірно легко), 60–70 (стандартно), 50–60 (помірно важко), 30–50 (важко), 0–30 (дуже важко). Вищі значення відповідають кращій читабельності для зовнішніх користувачів, що не мають спеціалізованої підготовки у визначеній предметній області.

Результати експериментальної перевірки наведено у таблиці 1.

Таблиця 1  
Результати експериментальної перевірки зрозумілості узагальненої символічної ментальної моделі рішення інтелектуальної системи

Метрика	Запропонована модель	SHAP (XGBoost)	LIME
TFWC (логарифм кількості результатів)	6,42	4,87	5,13
SCMC (косинусна подібність)	0,784	0,521	0,598
FRE (індекс Flesch)	68,3	42,1	51,6
Час побудови моделі (сек)	5,3	142,7	38,4

Запропонована узагальнена модель має найвищі значення TFWC – 6,42. Такі значення відповідають середній кількості результатів пошуку більше одного мільйона ключових термінів пояснень, наприклад, «high BMI diabetes risk», що свідчить про високу поширеність цих термінів у медичній веб-літературі та їхнє використання великою кількістю користувачів.

Базові підходи демонструють значно нижчі значення TFWC, зокрема, для SHAP TFWC становить 4,87, що відповідає десяткам тисяч результатів для термінів типу «SHAP value diabetes prediction».

Для LIME TFWC становить 5,13, що відповідає більше ніж 100 тисяч результатів для термінів типу «BMI contribution diabetes model». Такі результати пояснюються використанням технічної термінології машинного навчання замість стандартної медичної термінології. Остання менш поширена на медичних сайтах.

Запропонована модель має найвищу семантичну узгодженість між векторними представленнями пояснень та топ-10 результатів пошуку з медичних сайтів, що свідчить про використання термінології, яка використовується в популярній медичній літературі.

Запропонована трирівнева символна ментальна модель характеризується незначними обчислювальними витратами. Побудова моделі займає 5 секунд для експериментальної вибірки, що значно менше порівняно з методом SHAP, який потребує майже 143 секунди для перенавчання та обчислення значень Шеплі для всіх екземплярів. І також менше порівняно з методом LIME, який потребує більше 38 секунд для генерації локальних лінійних апроксимацій. Висока швидкість автоматизованої побудови моделі пояснюється ефективністю алгоритмів класифікації на основі логік для онтологій малого і середнього розміру. Також модель не потребує перенавчання для кожного нового набору даних користувачів.

Результати експериментальної перевірки свідчать про ряд переваг розробленої узагальної символної ментальної моделі порівняно з традиційними підходами до побудови пояснень. По-перше, модель дає можливість сформулювати зрозумілі каузальні шляхи міркування через символне представлення знань засобами онтологій та правил виведення, що підтверджується високими значеннями метрик TFWC та SCMC. Використовується термінологія предметної області доменної термінології замість технічної термінології машинного навчання. По-друге, модель забезпечує можливість автоматизації процесу побудови ментальних моделей через механізми логічного виведення онтологій без залучення експертів для структурованих інтерв'ю, що усуває потребу в трудомістких якісних методах побудови ментальних моделей на основі когнітивного картування. В третій, компоненти моделі можуть бути використані повторно для

інших предметних областей внаслідок застосування доменно-незалежної схеми онтологічного представлення, що скорочує витрати на адаптацію моделі до нових варіантів застосування інтелектуальної системи. В четвертих, механізм композиції рівнів для підтримки деталізації пояснень відповідно до рівня експертизи користувача дозволяє адаптувати складність представлення пояснень до індивідуальних особливостей та потреб користувачів. Новачки отримують спрощене представлення на концептуальному рівні без перевантаження деталями. Користувачі середнього рівня можуть деталізувати концепти для вивчення каузальних правил виведення. Експерти отримують повний доступ до конкретних значень атрибутів для верифікації всіх компонентів міркування. На відміну від розробленої моделі, традиційні методи побудови пояснень формують однаковий рівень складності пояснень для всіх користувачів незалежно від їхніх потреб у деталізації.

Проте дана модель має ряд обмежень, що визначають напрямки майбутніх досліджень для підвищення ефективності та розширення можливостей застосування представленого підходу.

Якість моделі залежить від повноти та коректності експертних знань про каузальні шляхи міркування від атрибутів користувачів до концептів рішень. Розроблена модель має суттєві затримки при виконанні механізмів логічного виведення онтологій, що обмежує можливості застосування підходу для систем із суттєвими вимогами до затримок при формуванні пояснень. Також дана модель не призначена для пояснення роботи глибоких нейронних мереж, призначених для обробки візуальної інформації, представлені векторами пікселів. Запропонована модель орієнтована на системи підтримки рішень із табличними даними користувачів, де атрибути мають явне семантичне значення для формування зрозумілих правил виведення.

## Висновки

Розроблена узагальнена символна ментальна модель рішення інтелектуальної системи для зовнішніх користувачів, що задається на основі трирівневої архітектури з явним розділенням концептуального шару високорівневих абстракцій рішень, пояснювального шару каузальних правил виведення та бази знань доменних фактів. Модель передбачає використання онтологічної формалізації та правил виведення ланцюжка міркувань, а також механізму вертикальної композиції рівнів для деталізації інформації відповідно до індивідуального рівня експертизи користувача, що забезпечує формування зрозумілих каузальних шляхів міркування від високорівневих концептів рішень через символні правила логічного виведення з використанням доменної термінології й до конкретних значень атрибутів користувачів.

Модель забезпечує зниження трудомісткості формування пояснень за рахунок виключення трудомісткого процесу побудови ментальних моделей експертами з використанням структурованих інтерв'ю.

Експериментальна перевірка узагальненої символічної ментальної моделі на наборі даних з медичної діагностики діабету типу 2 з використанням метрик на основі веб-пошуку та лінгвістичного аналізу підтвердила покращення зрозумілості пояснень порівняно з базовими підходами до побудови пояснень – методами SHAP та LIME. Практичне значення отриманих результатів полягає у створенні умов для побудови інструментальних засобів розробки пояснювальних інтелектуальних систем у ризикових предметних областях, таких як медичні діагностичні системи із поясненням клінічних рекомендацій для пацієнтів, фінансових систем планування з поясненням інвестиційних рекомендацій для клієнтів тощо.

#### Список літератури

- [1] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems, 30* (pp. 4765–4774). Curran Associates. <https://arxiv.org/abs/1705.07874>
- [2] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1135–1144). ACM. <https://doi.org/10.1145/2939672.2939778>
- [3] Aas, K., Jullum, M., & Løland, A. (2021). Explaining individual predictions when features are dependent: More accurate approximations to Shapley values. *Artificial Intelligence, 298*, Article 103502. <https://doi.org/10.1016/j.artint.2021.103502>
- [4] Molnar, C. (2022). *Interpretable machine learning: A guide for making black box models explainable* (2nd ed.). <https://christophm.github.io/interpretable-ml-book/>
- [5] Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access, 6*, 52138–52160. <https://doi.org/10.1109/ACCESS.2018.2870052>
- [6] Ghassemi, M., Oakden-Rayner, L., & Beam, A. L. (2021). The false hope of current approaches to explainable artificial intelligence in health care. *The Lancet Digital Health, 3*(11), e745–e750. [https://doi.org/10.1016/S2589-7500\(21\)00208-9](https://doi.org/10.1016/S2589-7500(21)00208-9)
- [7] Ambalavanan, R., Snead, R. S., Marczika, J., Towett, G., Malioukis, A., & Mbogori-Kairichi, M. (2025). Ontologies as the semantic bridge between artificial intelligence and healthcare. *Frontiers in Digital Health, 7*, Article 1668385. <https://doi.org/10.3389/fdgh.2025.1668385>
- [8] Guizzardi, G., Fonseca, C. M., Almeida, J. P. A., Sales, T. P., Benevides, A. B., & Porello, D. (2022). Ontology-driven conceptual modeling as a service: Surveys, blueprints, and roadmaps for composable modeling platform. In *Conceptual Modeling* (pp. 3–18). Springer. [https://doi.org/10.1007/978-3-031-17995-2\\_1](https://doi.org/10.1007/978-3-031-17995-2_1)
- [9] Hossain, D., & Chen, J. Y. (2025). A study on neuro-symbolic artificial intelligence: Healthcare perspectives. *arXiv preprint*. <https://arxiv.org/abs/2503.18213>
- [10] Olivares-Alarcos, A., Beßler, D., Khamis, A., Goncalves, P., Habib, M. K., Bermejo-Alonso, J., Barreto, M., Diab, M., Rosell, J., Quintas, J., Olszewska, J., Nakawala, H., Pignaton, E., Gyrard, A., Borgo, S., Alenya, G., Beetz, M., & Li, H. (2019). A review and comparison of ontology-based approaches to robot autonomy. *The Knowledge Engineering Review, 34*, Article e29. <https://doi.org/10.1017/S0269888919000237>
- [11] Bampi, D., Miranda, W. K. de M., & Almeida, J. L. V. (2025). Ontology-driven monitoring system for ambient assisted living. *The Knowledge Engineering Review, 40*, Article e3. <https://doi.org/10.1017/S0269888924000250>
- [12] Hoffman, R. R., Mueller, S. T., Klein, G., & Litman, J. (2018). Metrics for explainable AI: Challenges and prospects. *arXiv preprint*. <https://arxiv.org/abs/1812.04608>
- [13] Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence, 267*, 1–38. <https://doi.org/10.1016/j.artint.2018.07.007>
- [14] Zhang, Y., Liao, Q. V., & Bellamy, R. K. E. (2020). Effect of confidence and explanation on accuracy and trust calibration in AI-assisted decision making. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 295–305). ACM. <https://doi.org/10.1145/3351095.3372852>
- [15] Szymanski, M., Millecamp, M., & Verbert, K. (2021). Visual, textual or hybrid: The effect of user expertise on different explanations. In *Proceedings of the 26th International Conference on Intelligent User Interfaces* (pp. 109–119). ACM. <https://doi.org/10.1145/3397481.3450662>
- [16] Wang, X., & Yin, M. (2021). Are explanations helpful? A comparative study of the effects of explanations in AI-assisted decision-making. In *Proceedings of the 26th International Conference on Intelligent User Interfaces* (pp. 318–328). ACM. <https://doi.org/10.1145/3397481.3450650>
- [17] Weitz, K., Schiller, D., Schlagowski, R., Huber, T., & André, E. (2021). "Let me explain!": Exploring the potential of virtual agents in explainable AI interaction design. *Journal on Multimodal User Interfaces, 15*(2), 87–98. <https://doi.org/10.1007/s12193-020-00332-0>
- [18] Chazette, L., Brunotte, W., & Speith, T. (2021). Exploring explainability: A definition, a model, and a knowledge catalogue. In *2021 IEEE 29th International Requirements Engineering Conference (RE)* (pp. 197–208). IEEE. <https://doi.org/10.1109/RE51729.2021.00025>

Received (Надійшла) 16.01.2026

Accepted for publication (Прийнята до друку) 23.02.2026

Publication date (Дата публікації) 27.03.2026

УДК 004.8:004.9

DOI: [https://doi.org/10.30837/bi.2026.1\(104\).11](https://doi.org/10.30837/bi.2026.1(104).11)С. Ф. Чалий<sup>1</sup>, Р. В. Кравченко<sup>1</sup><sup>1</sup>ХНУРЕ, м. Харків, Україна, serhii.chalyi@nure.ua,  
ORCID iD: 0000-0002-9982-9091<sup>1</sup>ХНУРЕ, м. Харків, Україна, rostyslav.kravchenko1@nure.ua;  
ORCID: 0009-0009-0324-3597

## МЕТОД ПОБУДОВИ ТЕМПОРАЛЬНО УЗГОДЖЕНИХ ПОЯСНЕНЬ В ІНТЕЛЕКТУАЛЬНИХ СИСТЕМАХ НА ОСНОВІ ТЕМПОРАЛЬНИХ ГРАФОВИХ НЕЙРОННИХ МЕРЕЖ

Розглянуто проблему побудови пояснень з урахуванням порядку у часі вхідних даних інтелектуальних систем, вирішення якої потребує забезпечення темпоральної узгодженості каузальних зв'язків між подіями у згенерованих поясненнях. Розроблено інтервальний підхід до формалізації темпоральних відношень між подіями на основі тринадцяти базових відношень інтервальної алгебри з транзитивним поширенням обмежень для виявлення неявних темпоральних протиріч. Запропоновано метод побудови темпорально узгоджених пояснень, що включає обчислення внеску подій через інтегровані градієнти, гібридну генерацію пояснень природною мовою на основі шаблонів та донавченої великої мовної моделі, автоматизовану верифікацію темпоральної коректності з використанням інтервальної алгебри та оновлення фрагментів пояснень із темпоральними порушеннями. Метод забезпечує можливість формування пояснень для систем з упорядкованими у часі вхідними даними в критичних предметних областях, включаючи рекомендаційні системи електронної комерції, медичну діагностику, фінансові системи підтримки рішень, де необхідні як інтерпретованість рішень, так і верифікація коректності темпоральних залежностей між подіями.

ТЕМПОРАЛЬНО УЗГОДЖЕНІ ПОЯСНЕННЯ, ПОЯСНЮВАНИЙ ШТУЧНИЙ ІНТЕЛЕКТ, ТЕМПОРАЛЬНІ ГРАФОВІ НЕЙРОННІ МЕРЕЖІ, ІНТЕРВАЛЬНА АЛГЕБРА, ІНТЕГРОВАНІ ГРАДІЄНТИ, ВЕЛИКІ МОВНІ МОДЕЛІ

**S.F. Chalyi, R.V. Kravchenko. Method for constructing temporally consistent explanations in intelligent systems based on temporal graph neural networks.** The paper addresses the problem of constructing explanations that consider the temporal order of input data in intelligent systems, requiring the temporal consistency of causal relationships in the generated explanations. An interval-based approach for formalizing temporal relationships between events is developed based on thirteen basic relations of interval algebra with transitive constraint propagation to detect implicit temporal contradictions. A method for constructing temporally consistent explanations is proposed, which includes computing event contributions using integrated gradients, hybrid generation of natural-language explanations based on templates and a fine-tuned large language model, automated verification of temporal correctness using interval algebra, and updating explanation fragments with temporal violations. The method enables the generation of explanations for systems with time-ordered input data in critical domains, including e-commerce recommendation systems, medical diagnosis, and financial decision support systems, where both interpretability of decisions and verification of temporal dependency correctness are essential.

TEMPORALLY CONSISTENT EXPLANATIONS, EXPLAINABLE ARTIFICIAL INTELLIGENCE, TEMPORAL GRAPH NEURAL NETWORKS, INTERVAL ALGEBRA, INTEGRATED GRADIENTS, LARGE LANGUAGE MODELS

### Вступ

Інтелектуальні системи на основі глибоких нейронних мереж досягають високої точності прогнозування поведінки користувачів у широкому спектрі предметних областей, включаючи рекомендаційні системи на платформах електронної комерції, медичну діагностику, фінансовий менеджмент. Проте такі системи з точки зору користувача функціонують як непрозора обчислювальна структура, що не відображає причинно-наслідкові зв'язки між вхідними даними та прогнозами системи. Відсутність пояснень прийнятих рішень зменшує довіру користувачів та знижує можливості впровадження пояснень для критичних застосувань, де помилкові прогнози можуть призвести до фінансових втрат, загроз безпеці чи порушень етичних норм.

Забезпечення пояснень для рішень інтелектуальних систем є нормативною вимогою згідно

з Регламентом Європейського Союзу щодо штучного інтелекту, у відповідності до якого мають бути надані зрозумілі пояснення логіки автоматизованого прийняття рішень для систем високого ризику, включаючи рекомендаційні системи електронної комерції та фінансові застосування [1]. Проте існуючі методи побудови пояснень для систем з упорядкованими у часі вхідними даними мають суттєвий недолік, пов'язаний з тим, що описана у поясненні каузальна послідовність подій суперечить фактичним міткам часу виникнення цих подій у вхідних даних системи. Зокрема, великі мовні моделі генерують значну кількість темпоральних невідповідностей у поясненнях звичайною мовою для послідовностей подій через відсутність явної перевірки міток часу.

Таким чином, проблема побудови пояснень з урахуванням порядку у часі у вхідних даних є актуальною.

Побудови пояснень виконуються в рамках наукового напрямку пояснювального штучного інтелекту (XAI). Даний напрямок досліджень спрямований на забезпечення зрозумілості рішень систем штучного інтелекту. Для побудови пояснень з урахуванням темпоральної упорядкованості вхідних даних в рамках даного дослідження розглядаються чотири групи методів. Графові методи обчислюють внески вузлів графової нейронної мережі (GNN) шляхом інтегрування градієнтів, виявляючи структурні патерни. Темпоральні графові методи розширюють статичні підходи механізмами пам'яті та агрегацією з урахуванням темпоральної еволюції. Методи на основі мовних моделей трансформують структуровані атрибути пояснень у пояснення природною мовою з використанням коротких структурованих промптів або гібридної генерації на основі патернів. Методи формалізації послідовностей подій у часі використовують темпоральні правила та інтервальну алгебру для представлення каузальних залежностей з темпоральними обмеженнями, що дає можливість сформувати патерни для генерації пояснень з використанням мовних моделей.

В роботі [2] показано, що градієнтні підходи для GNN забезпечують локальні пояснення, проте не генерують пояснення природною мовою та не верифікують коректність пояснень з урахуванням послідовності вхідних даних у часі. В дослідженні [3] запропонували інструментарій GNNExplainer для зменшення невідповідності між прогнозом та структурами підграфів, проте даний інструментарій не враховує послідовність вхідних подій у часі. В [4] розробили метод інтегрованих градієнтів, який задовольняє аксіоми повноти та чутливості, але не враховує темпоральну структуру вхідних даних. В роботі [5] запропонували Темпоральні графові нейронні мережі (TGNN) з модулем пам'яті для історичних станів та агрегацією з урахуванням часу, проте без механізмів пояснення з верифікацією. Дослідження [6] присвячено розробці TGNN з синусоїдальним кодуванням та багатоголовою увагою для визначення релевантних подій. Робота [7] пропонує метод адаптивного вибору інтервалів на основі відстані редагування графів та спектрального аналізу для детекції структурного дрейфу. В дослідженні [8] запропонували модель GNN з LSTM для виявлення темпоральних патернів та механізм уваги, що комбінує патерни окремих та взаємопов'язаних вершин.

Великі мовні моделі генерують пояснення на основі коротких промптів, однак створюють темпоральні невідповідності. В дослідженні [9] виявили темпоральні зміщення у LLM, тобто асиметрію точності для відношень «before» vs «after». Гібридні підходи комбінують шаблони з донавченими моделями. В [10] Wiegreffe et al. показали, що гібридна генерація підвищує контрольованість та фактуальну коректність.

Дослідження [11] присвячено розробці методу представлення знань на основі темпоральних правил виду «якщо подія A, то через деякий час T відбудеться подія B, що забезпечує інтерпретабельність пояснення.

Інтервальна алгебра [12] на основі 13 базових відношень між інтервалами (before, after, meets, overlaps, during, starts, finishes, equals) дає можливість формалізувати темпоральні відношення. Алгебра забезпечує поширення обмежень на основі властивості транзитивності.

В дослідженні [13] застосували інтервальну алгебру для темпоральних баз даних. Робота [14] Leucker та Schallhart присвячена вирішенню проблеми верифікації в режимі реального часу з використанням лійної темпоральної логіки.

Проведений аналіз показав, що існуючі методи побудови пояснень не забезпечують автоматизовану верифікацію темпоральної узгодженості каузальних зв'язків між подіями. Методи на основі великих мовних моделей генерують текстові пояснення, однак не перевіряють мітки часу і тому генерують значну кількість темпоральних невідповідностей. Методи обчислення внеску вузлів нейронної мережі через градієнти обчислюють показники внеску вузлів, проте не перетворюють їх на пояснення природною мовою з верифікацією послідовності у часі на основі темпоральної логіки. Інтервальна алгебра забезпечує формальне представлення темпоральних відношень, проте не інтегрована в процес генерації пояснень для інтелектуальних систем на основі нейронних мереж. Методи представлення знань через темпоральні правила забезпечують інтерпретовану форму каузальних зв'язків, проте мають обмеження щодо масштабованості для великих графових структур.

Таким чином, необхідною є розробка методу, який інтегрує визначення впливу упорядкованих у часі вхідних подій, гібридну генерацію природною мовою на основі комбінації темпоральних шаблонів та тексту, згенерованого мовною моделлю, верифікацію темпоральних невідповідностей з використанням інтервальної алгебри у єдиний автоматизований процес генерації темпорально узгоджених пояснень.

## 1. Постановка задачі

Метою дослідження є розробка методу побудови темпорально узгоджених пояснень в інтелектуальних системах з автоматизованою верифікацією темпоральної коректності каузальних зв'язків між подіями, що відображають поведінку користувачів, та генерацією пояснень природною мовою.

Для досягнення мети дослідження вирішуються такі задачі: розробити інтервальний підхід до формування та верифікації темпорально узгоджених пояснень; розробити метод побудови темпорально узгоджених пояснень.

## 2. Інтервальний підхід до формування та верифікації темпорально узгоджених пояснень

Розроблений підхід передбачає формалізацію темпоральних відношень з використанням інтервальної алгебри, обчислення внеску подій у пояснення через інтегровані градієнти, генерацію пояснень природною мовою з донавчанням моделі, верифікацію темпоральної узгодженості для пояснення.

Базовий опис складових розробленого підходу наведено в табл. 1.

Таблиця 1

Ключові складові інтервального підходу

Компонент	Коментар	Призначення
1. Представлення темпоральних відношень (інтервальна алгебра)	Система правил для подій у часі одна відносно одної. Наприклад: «подія А ПЕРЕКРИВАЄ подію Б»	Забезпечити однозначне розуміння послідовності подій
2. Показник внеску події (інтегровані градієнти)	Числова оцінка того, наскільки конкретна подія вплинула на фінальне рішення системи. Показник може бути позитивним (подія збільшила ймовірність рекомендації) або негативним (зменшила)	Виявити, які саме події з історії користувача є найбільш важливими для пояснення. Наприклад, для системи e-commerce причиною рекомендації може бути минула покупка, перегляд тощо
3. Функція втрат для донавчання мовної моделі	Міра того, наскільки добре система навчилася генерувати пояснення у стилі та з термінами конкретної галузі	Забезпечити формування пояснень з використанням термінології предметної області
4. Верифікація темпоральної узгодженості	Автоматична перевірка того, чи не містить пояснення логічних помилок у часі. Наприклад, чи не каже система "А призвело до Б", коли насправді Б відбулося раніше А.	Виявити та виправити протиріччя в часі у поясненні

Формалізація темпоральних відношень між інтервалами часу подій здійснюється через інтервальну алгебру Аллена. Інтервал часу події представляється через дві часові мітки — момент початку події та момент завершення події. Алгебра визначає тринадцять базових типів відношень між двома інтервалами, які повністю описують усі можливі варіанти взаємного розташування інтервалів на часовій осі.

Базові відношення включають: before (до), after (після), meets (межує з), met-by (межується з), overlaps (перекриває), overlapped-by (перекривається), during

(всередині), contains (містить), starts (розпочинає), started-by (розпочинається одночасно з), finishes (завершує), finished-by (завершується одночасно з), equals (дорівнює).

Транзитивне поширення обмежень виконується через правила транзитивності інтервальної алгебри, які дозволяють автоматично виводити нові темпоральні відношення з відомих відношень. Наприклад, якщо подія А відбулася before події В, та подія В відбулася before події С, то автоматично виводиться, що подія А відбулася before події С, навіть якщо це відношення не згадано явно.

Показник внеску події обчислюється через метод інтегрованих градієнтів, який визначає вплив події на прогноз моделі.

Метод інтегрованих градієнтів обчислює внесок події у прогноз системи через вимірювання різниці між фактичним станом події  $h_i$  та нейтральним станом  $h_i^{base}$ , помноженої на середнє значення чутливості прогнозу до змін стану події вздовж шляху від нейтрального до фактичного стану. Це дозволяє визначити, наскільки конкретна подія збільшила або зменшила ймовірність прогнозу порівняно з ситуацією, коли ця подія була б відсутня або нейтральна. Відповідно, сума внесків усіх подій дорівнює різниці прогнозів. Тобто весь «ефект прогнозу» щодо системи повністю розподіляється між подіями без втрат. Такий підхід забезпечує інтерпретованість результатів і користувач може бути впевнений, що аналіз охоплює всі фактори, які вплинули на прогноз.

Функція втрат для донавчання мовної моделі обчислюється як:

$$\mathcal{L}_{gen} = -\frac{1}{L} \sum_{t=1}^L \log P(y_t | y_{<t}, S), \quad (1)$$

де  $L$  — довжина тексту у токенах;  $y_t$  — токен на позиції  $t$ ;  $S$  — структуроване пояснення з подіями та показниками внеску.

Функція втрат для донавчання вимірює, наскільки ймовірним є згенерований текст пояснення за умови структурованої інформації про події та їх внески. Мінімізація цієї функції змушує мовну модель навчитися генерувати пояснення природною мовою, які максимально відповідають структурованим даним та стилістичним нормам предметної області. Чим менше значення функції втрат, тим краще модель передбачає наступне слово у поясненні на основі попередніх слів та структурованої інформації.

При верифікації темпоральної узгодженості визначається показник  $TCS$ , який визначає частку коректних темпоральних відношень між подіями у згенерованому поясненні. Для кожної пари подій порівнюється темпоральне відношення, яке впливає з природномовного тексту пояснення, з фактичним темпоральним відношенням, обчисленим на основі міток часу з бази даних системи.

$$TCS = \frac{1}{K(K-1)} \sum_{i=1}^K \sum_{j \neq i}^K \mathbb{I}(R_{ij} = R'_{ij}), \quad (2)$$

де  $K$  – кількість подій;  $R_{ij}$  – фактичне темпоральне відношення;  $R'_{ij}$  – відношення з пояснення;  $\mathbb{I}(\cdot)$  – індикаторна функція, яка приймає значення 1 за умови істинності виразу  $R_{ij} = R'_{ij}$  та 0 в іншому випадку.

Значення  $TCS = 1$  означає повну темпоральну узгодженість – усі темпоральні відношення у поясненні відповідають фактичним міткам часу;  $TCS = 0$  означає повну невідповідність.

### 3. Метод побудови темпорально узгоджених пояснень

Розроблений метод формує темпорально узгоджені пояснення на основі інтеграції внеску подій, гібридної природномовної генерації пояснень та автоматизованої верифікації з використанням інтервальної алгебри. Метод виконується для кожного рішення інтелектуальної системи, що потребує пояснення.

Етап 1. Формування структурованого пояснення через обчислення внеску подій.

Вхідна послідовність подій з часовими мітками подається на темпоральну графову нейронну мережу для формування прогнозу.

Темпоральна графова нейронна мережа обробляє послідовність подій через механізм агрегації сусідів з урахуванням часових міток взаємодій та обчислює приховані репрезентації для кожної події.

Для кожної події обчислюється показник внеску в рішення, який показує, наскільки конкретна подія збільшила або зменшила ймовірність прогнозу порівняно з ситуацією відсутності цієї події. Обчислення виконується через інтегрування градієнтів функції прогнозу вздовж прямолінійного шляху від базової до фактичної репрезентації події, що забезпечує стабільність результатів відносно шуму у градієнтах.

При імплементації етапу формується структуроване пояснення з набору подій з найвищими абсолютними показниками внеску, впорядкованих за зростанням значень часу у часових мітках, з тим, щоб забезпечити коректну темпоральну послідовність.

Етап 2. Гібридна генерація пояснення природною мовою.

Структуроване пояснення подається на гібридний генератор, який комбінує попередньо визначений шаблон з явною темпоральною послідовністю та подає результат комбінування на донавчену велику мовну модель.

Шаблон вибирається з бібліотеки шаблонів залежно від типу прогнозу та кількості подій, з тим, щоб забезпечити правильну темпоральну послідовність через явні темпоральні залежності між заповнювачами місця для опису подій.

Донавчена мовна модель заповнює місця для опису подій шаблоном, генеруючи текстові сегменти з

описом кожної події, зазначенням її внеску у прогноз та доменно-специфічною термінологією. Якість генерації забезпечується мінімізацією функції втрат згідно з формулою, яка змушує модель генерувати пояснення, що максимально відповідають структурованим даним та стилістичним нормам предметної області.

В результаті виконання етапу генерується пояснення природною мовою, яке описує каузальні зв'язки між подіями з дотриманням темпоральної послідовності згідно з шаблоном.

Етап 3. Автоматизована верифікація темпоральної узгодженості подій у поясненні.

Згенероване пояснення аналізується для вилучення всіх згаданих темпоральних відношень між подіями через розпізнавання темпоральних зв'язок типу «після», «внаслідок», «що призвело до».

Для кожної пари подій обчислюється фактичне темпоральне відношення на основі міток часу з бази даних згідно з тринадцятьма базовими відношеннями інтервальної алгебри Аллена.

Виконується транзитивне поширення обмежень через застосування правил транзитивності інтервальної алгебри, що дозволяє виявити неявні темпоральні протиріччя, які не є очевидними при аналізі окремих пар подій. Обчислюється показник темпоральної узгодженості як частка коректних темпоральних відношень у згенерованому поясненні.

Виявляють темпоральні невідповідності через порівняння відношень з тексту та фактичних відношень, обчислених на основі порівняння міток часу. Кожна виявлена невідповідність фіксується із зазначенням проблемного сегмента пояснення.

Етап 4. Перебудова фрагментів пояснення з темпоральними порушеннями.

Для кожної виявленої темпоральної невідповідності виконується перебудова відповідного фрагмента пояснення. Мовна модель отримує доповнений промпт з додатковим обмеженням у вигляді фактичного темпорального відношення між подіями, яке повинно бути відображено у згенерованому тексті з дотриманням семантичної узгодженості згідно з функцією втрат.

Перебудова фрагментів з темпоральними порушеннями виконується ітеративно, тобто після регенерації фрагмента знову виконується верифікація темпоральної узгодженості для перевірки: чи усунена невідповідність. Процес продовжується до досягнення  $TCS = 1$ , що означає повну темпоральну узгодженість, або до досягнення максимальної кількості ітерацій.

У випадку недосягнення повної темпоральної узгодженості після максимальної кількості ітерацій виконується повна перебудова пояснення з використанням альтернативного шаблону з явними вказівками на часові мітки подій.

#### 4. Експериментальна перевірка розробленого методу

Експериментальна перевірка виконана з використанням публічного набору даних Retailrocket Recommender System Dataset, який містить 2 756 101 подію користувачів за період травень—вересень 2015 року.

Для експериментів використано підвибірку з 127 453 транзакцій, що включає послідовності подій: перегляди товарів (view), додавання у кошик (addtocart), покупки (transaction). Кожна подія містить ідентифікатор користувача, ідентифікатор товару, тип події та мітку часу у форматі UNIX timestamp.

Виділено фрагмент набору даних, який розділено на навчальну вибірку (80%), валідаційну вибірку (10%) та тестову вибірку (10%) з дотриманням темпоральної послідовності.

Для генерації пояснень використовується відкрита мовна модель LLaMA-2-7B з донавчанням методом LoRA з квантизацією 4 біти (Low-Rank Adaptation – метод ефективного донавчання великих моделей через низькорангові матриці адаптації) на наборі з 100 пар структурованих пояснень та референсних природномовних пояснень. Вибір LLaMA-2-7B та донавчання через LoRA обумовлений можливістю локального розгортання на мобільній робочій станції та ефективністю навчання з обмеженими обчислювальними ресурсами.

Експериментальний стенд реалізовано на мобільній робочій станції HP ZBook Studio з графічним прискорювачем NVIDIA. Програмна реалізація виконана з використанням фреймворку PyTorch 2.1 для навчання нейронних мереж, бібліотеки PyTorch Geometric для обробки графових структур, бібліотеки Transformers від Hugging Face для роботи з мовними моделями, бібліотеки PEFT для донавчання через LoRA.

Для оцінювання якості пояснень використано метрики UTS, TCS, TEC. User Trust Score (UTS – показник довіри користувачів) обчислюється як середнє значення оцінок довіри за шкалою від 1 до 10, де 1 – повна недовіра, 10 – повна довіра. Дана оцінка була адаптована для виключення суб'єктивної складової. Замість опитування користувачів було виконано аналіз результатів пошукових запитів. Значення оцінки розраховувалося як нормована кількість відповідей пошукової системи для вибірки із випадково вибраних прогнозів системи. Temporal Coherence Score (TCS) – як відношення кількості коректних темпоральних відношень до загальної кількості відношень. Temporal Error Count (TEC) обчислюється як абсолютна кількість виявлених темпоральних невідповідностей у 100 згенерованих поясненнях.

Для порівняння використовувались такі базові методи.

Метод 1. Генерація пояснень через GPT з типовими промптами без донавчання та верифікації отриманих пояснень.

Метод 2. Використання інструменту GNNExplainer для обчислення показників важливості і подальшої трансформації у текст через статичні шаблони.

Метод 3. Використання адаптованої LLaMA-2-7B через LoRA без гібридної генерації з шаблонами та без верифікації.

Метод 4. Генерація пояснень на основі темпоральних правил, що відображають зв'язки у часі між послідовностями подій.

Результати експериментальної перевірки розробленого методу представлено у табл. 2.

Таблиця 2

Порівняльний аналіз результатів методів генерації пояснень з темпорально упорядкованими даними

Метод	UTS (1–10)	TCS (0–1)	TEC (на 100 пояснень)
1. GPT	6,1	0,58	67
2. GNNExplainer та шаблон	5,3	0,45	78
3. LLaMA-2 -7B	6,8	0,65	52
4. Темпоральні правила	7,2	0,84	24
<b>Розроблений метод</b>	<b>8,3</b>	<b>0,92</b>	<b>6</b>

Розроблений метод досяг значення UTS 8,3 порівняно з 6,1 для великої мовної моделі.

Результати аналізу ефективності складових методу представлено у табл. 3.

Таблиця 3

Порівняльний аналіз ефективності етапів методу

Варіант методу	UTS (1–10)	TCS (0–1)	TEC (на 100 пояснень)
Повний метод	8,3	0,92	6
Метод без верифікації темпоральної узгодженості подій у поясненні	7,1	0,70	48
Метод без етапу перебудови фрагментів з темпоральними порушеннями	6,9	0,73	42
Метод без гібридної генерації (тільки з використанням шаблонів)	5,8	0,88	18
Метод без інтегрованих градієнтів (прості градієнти)	7,9	0,90	8

З таблиці 3 видно, що невиконання верифікації призводить до зниження TCS з 0,92 до 0,70 (тобто майже на 25%), збільшення TEC з 6 до 48 (збільшення кількості помилок у 8 разів), що підтверджує важливість автоматизованої верифікації для забезпечення темпоральної узгодженості пояснень. Заміна інтегрованих градієнтів на прості градієнти призводить

до незначного зниження всіх метрик, що вказує на корисність інтегрованих градієнтів для стабільності обчислень показників внеску, проте не є критичним компонентом для темпоральної узгодженості.

Результати експериментальної оцінки обчислювальної складності розробленого методу на основі часу виконання методу наведено у табл. 4.

Таблиця 4

Витрати часу на виконання етапів методу

Етап методу	Час (CPU)	Час (GPU)
Формування структурованого пояснення	28,3 с	2,1 с
Гібридна генерація	3,2 с	0,8 с
Верифікація темпоральної узгодженості	4,1 с	0,5 с
Регенерація сегментів (в середньому 1,8 ітерацій)	5,7 с	1,4 с
Загальний час	41,3 с	4,8 с

З таблиці 4 видно, що використання графічного процесора з об'ємом пам'яті 8 ГБ знижує час розрахунку пояснення у 8 разів, до 4,8 секунди, що свідчить про придатність розробленого методу до практичного застосування в режимі роботи ніарлайн.

Найбільш обчислювально затратним етапом є формування структурованого пояснення через обчислення інтегрованих градієнтів для всіх подій у послідовності. Використання графічного прискорювача знижує час виконання цього етапу з 28,3 секунди до 2,1 секунди (прискорення більш ніж у 13 разів).

Таким чином, розроблений метод формує пояснення на основі узгодженості між тим, як інтелектуальна система пояснює свої рекомендації, та реальною послідовністю подій у часі. Наприклад, коли користувач спочатку додає товар у кошик (подія А, момент часу  $t_1$ ), потім переглянув схожий товар (подія В, момент часу  $t_2$ ), і нарешті здійснив покупку (подія С, момент часу  $t_3$ ), то інтелектуальна система не має генерувати пояснення «Ви купили товар С, тому що переглянули товар В», що є логічним, але порушує порядок у часі, оскільки подія В відбулася після додавання товару у кошик (подія А), яку система проігнорувала.

Розроблений метод має ряд переваг за рахунок того, що використовує нейросимвольний підхід до побудови пояснень. Нейронна мережа забезпечує природність мови та адаптивність до контексту. Символьний компонент, заданий інтервальною алгеброю, забезпечує перевірку коректності послідовності дій у часі. Тобто система не може згенерувати пояснення виду «подія А спричинила подію В», якщо насправді подія В відбулася раніше А.

Процедура автоматизованого виправлення помилок, виконує ітеративне переформування пояснення до тих пір, поки всі темпоральні відношення не стануть узгодженими, на відміну від традиційних

підходів, де згенероване пояснення є фінальним результатом.

Метод забезпечує адаптацію до предметної області за рахунок донавчання мовної моделі на обмеженому наборі прикладів пояснень, що дає можливість використовувати в поясненнях терміни з предметної області. Наприклад, для e-commerce це будуть терміни «сезонний тренд», «promotional campaign», «категорія товарів» у відповідному контексті. Такий підхід підвищує довіру користувачів, оскільки пояснення формуються аналогічно тому, що напише фахівець e-commerce.

Кожне пояснення може бути доповнене списком перевірених темпоральних відношень та їх статусом (виконується/не виконується). Такий підхід має переваги при побудові пояснень в критичних предметних областях, пов'язаних із ризиками для здоров'я, фінансів тощо. Система повинна пояснити не лише те, що вона рекомендує, а й деталізувати, чому ці пояснення є коректними з використанням темпоральної логіки.

Розроблений метод має ряд обмежень. По-перше, обчислювальні витрати суттєво залежать від кількості упорядкованих у часі подій, оскільки вклад обчислюється для кожної події у вхідному наборі даних. Відповідно, витрати часу на CPU склали 42,5 секунди, що не дає можливості використати метод в онлайн-режимі. По-друге, якість текстових пояснень значною мірою залежить від навчальних даних, які використовуються для донавчання мовної моделі. Доновчання потребує декілька сотень прикладів, і тому може бути проблематичним для предметних областей з обмеженою кількістю експертних пояснень.

Також інтервальна алгебра використовує обмежену кількість темпоральних відношень, що дає можливість формалізувати стандартні сценарії, проте не дає можливість описати деякі специфічні патерни. Наприклад, при неточному визначенні міток у вхідних даних розподіленої системи відбувається часткове перекриття подій на декілька хвилин, що не дає можливості точно врахувати темпоральну упорядкованість.

Зазначені обмеження визначають напрямки подальших досліджень, пов'язані із розпаралелюванням обчислень внеску подій, використанням ймовірнісної темпоральної логіки тощо.

## Висновки

Розроблено метод побудови темпорально узгоджених пояснень в інтелектуальних системах, який інтегрує обробку послідовностей подій користувачів з часовими мітками з використанням графових нейронних мереж та автоматизовану верифікацію каузальної коректності згенерованих пояснень з використанням інтервальної алгебри. Метод включає етапи обчислення внеску подій через інтегровані

градієнти, гібридну генерацію пояснень звичайною мовою на основі шаблонів та донавченої великої мовної моделі, автоматизовану верифікацію обмежень з використанням інтервальної алгебри, таргетовану регенерацію сегментів з темпоральними порушеннями. Розроблений метод створює умови для впровадження пояснювальних інтелектуальних систем з можливістю інтерпретації рішень у критичних застосуваннях, де коректність темпоральних залежностей має високе значення при прийнятті управлінських рішень користувачами.

Подальші дослідження доцільно спрямувати на розширення методу для підтримки темпоральних логік з темпоральними кванторами та обмеженнями на тривалість інтервалів між подіями.

Експериментальна валідація на наборі даних рекомендаційної системи електронної комерції підтвердила підвищення показника довіри користувачів до пояснень, збільшення темпоральної узгодженості пояснень та зниження кількості помилок щодо послідовності в часі.

#### Список літератури

- [1] Regulation (EU) 2024/1689 of the European Parliament and of the Council on Artificial Intelligence (AI Act) // Official Journal of the European Union. 2024. L 1689. P. 1–144. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689>
- [2] Longo L., Goebel R., Lecue F., Kieseberg P., Holzinger A. Explainable Artificial Intelligence: Concepts, Applications, Research Challenges and Visions // Machine Learning and Knowledge Extraction. Springer, 2020. Vol. 1. P. 1–16. DOI: [https://doi.org/10.1007/978-3-030-57321-8\\_1](https://doi.org/10.1007/978-3-030-57321-8_1)
- [3] Ying R., Bourgeois D., You J., Zitnik M., Leskovec J. GNNExplainer: Generating Explanations for Graph Neural Networks // Advances in Neural Information Processing Systems (NeurIPS). 2019. Vol. 32. P. 9240–9251. URL: <https://proceedings.neurips.cc/paper/2019/hash/d80b7040b773199015de6d3b4293c8ff-Abstract.html>
- [4] Sundararajan M., Taly A., Yan Q. Axiomatic Attribution for Deep Networks // Proceedings of the 34th International Conference on Machine Learning (ICML). 2017. Vol. 70. P. 3319–3328. DOI: <https://doi.org/10.48550/arXiv.1703.01365>
- [5] Rossi E., Chamberlain B., Frasca F., Eynard D., Monti F., Bronstein M. Temporal Graph Networks for Deep Learning on Dynamic Graphs // Proceedings of the 37th International Conference on Machine Learning (ICML). 2020. P. 8230–8240. DOI: <https://doi.org/10.48550/arXiv.2006.10637>
- [6] Xu K., Hu W., Leskovec J., Jegelka S. How Powerful are Graph Neural Networks? // Proceedings of the 7th International Conference on Learning Representations (ICLR). 2019. P. 1–17. DOI: <https://doi.org/10.48550/arXiv.1810.00826>
- [7] Чалий С. Ф., Кравченко Р. В. Метод адаптивного вибору інтервалів часу для побудови графів темпоральних графових нейронних мереж // Вісник Національного технічного університету "ХПІ". Серія: Системний аналіз, управління та інформаційні технології. 2025. № 2(14). С. 129–139. DOI: <https://doi.org/10.20998/2079-0023.2025.02.19>
- [8] Чалий С. Ф., Кравченко Р. В. Графова нейронна мережа для темпорально упорядкованих даних в задачі побудови пояснень в інтелектуальній системі // Автоматизовані системи управління та прилади автоматики. 2025. Вип. 185. С. 77–86. DOI: <https://doi.org/10.30837/0135-1710.2025.185.077>
- [9] Tan Q., Ng H. T., Bing L. Towards Benchmarking and Improving the Temporal Reasoning Capability of Large Language Models // Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL). 2023. Vol. 1. P. 14820–14835. DOI: <https://doi.org/10.18653/v1/2023.acl-long.828>
- [10] Wiegreffe S., Marasović A. Teach Me to Explain: A Review of Datasets for Explainable NLP // Proceedings of the 35th Conference on Neural Information Processing Systems (NeurIPS). 2021. P. 1–14. DOI: <https://doi.org/10.48550/arXiv.2102.12060>
- [11] Levykin V., Chala O. Method of Determining Weights of Temporal Rules in Markov Logic Network for Building Knowledge Base in Information Control Systems // EUREKA: Physics and Engineering. 2018. No. 5. P. 3–10. DOI: <https://doi.org/10.21303/2461-4262.2018.00713>
- [12] Allen J. F. Maintaining Knowledge about Temporal Intervals // Communications of the ACM. 1983. Vol. 26. No. 11. P. 832–843. DOI: <https://doi.org/10.1145/182.358434>
- [13] Chomicki J., Toman D. Temporal Databases // Handbook of Temporal Reasoning in Artificial Intelligence. Elsevier, 2005. P. 429–467. DOI: [https://doi.org/10.1016/S1574-6526\(05\)80015-6](https://doi.org/10.1016/S1574-6526(05)80015-6)
- [14] Leucker M., Schallhart C. A Brief Account of Runtime Verification // Journal of Logic and Algebraic Programming. 2009. Vol. 78. No. 5. P. 293–303. DOI: <https://doi.org/10.1016/j.jlap.2008.08.004>

*Received (Надійшла) 08.02.2026*

*Accepted for publication (Прийнята до друку) 02.03.2026*

*Publication date (Дата публікації) 27.03.2026*

І. А. Редозуб<sup>1</sup>, Н. Ф. Хайрова<sup>2,3</sup><sup>1</sup>НТУ «ХПІ», м. Харків, Україна, [ivan.red707@gmail.com](mailto:ivan.red707@gmail.com), ORCID ID: 0009-0008-3312-4700<sup>2</sup>НТУ «ХПІ», м. Харків, Україна;<sup>3</sup>Umea University, Sweden, [nina.khairova@umu.se](mailto:nina.khairova@umu.se), ORCID ID: 0000-0002-9826-0286

## ОЦІНЮВАННЯ ЗДАТНОСТІ МОДЕЛЕЙ GENAI ВИЯВЛЯТИ ЕТИЧНІ ВИМІРИ ПУБЛІЧНОГО ДИСКУРСУ: КЕЙС УКРАЇНСЬКИХ БІЖЕНЦІВ

У статті досліджується здатність моделей генеративного штучного інтелекту (GenAI) виявляти й інтерпретувати етичні виміри публічного дискурсу в цифровому середовищі. Актуальність роботи зумовлена зростанням ролі соціальних мереж як джерела суспільних наративів і потребою в аналізі ціннісно навантажених, нормативно чутливих категорій. Зокрема, ми використовуємо модель GPT-3.5-turbo для класифікації текстових даних за визначеними категоріями етичних аспектів міграції. З позицій підходу «етика в дизайні» оцінюється потенціал застосування GenAI для зіставлення громадянських і урядових наративів та аналізу їх узгодженості в контексті підтримки процесів суспільної згуртованості.

ГЕНЕРАТИВНИЙ ШТУЧНИЙ ІНТЕЛЕКТ, ВЕЛИКІ МОВНІ МОДЕЛІ, ЕТИЧНІ ВИМІРИ, АНАЛІЗ СОЦІАЛЬНИХ МЕРЕЖ, ТАКСОНОМІЯ

**I. A. Redozub, N. F. Khairova. Evaluating the Capability of GenAI Models to Identify Ethical Dimensions of Public Discourse: The Case of Ukrainian Refugees.** This paper investigates the capability of generative artificial intelligence (GenAI) models to identify and interpret ethical dimensions of public discourse in the digital environment. The relevance of the study is driven by the growing role of social media as a source of societal narratives and the need to analyze value-laden and normatively sensitive categories. In particular, we employ the GPT-3.5-turbo model to classify textual data according to predefined categories of ethical aspects of migration. From the perspective of the Ethics-in-Design approach, the study evaluates the potential of GenAI to compare citizen and governmental narratives and to analyze their alignment in the context of supporting processes of social cohesion.

GENERATIVE ARTIFICIAL INTELLIGENCE, LARGE LANGUAGE MODELS, ETHICAL DIMENSIONS, SOCIAL MEDIA ANALYSIS, TAXONOMY

### Вступ

В останні роки бурхливий розвиток генеративно-штучного інтелекту (Generative artificial intelligence (GenAI)) суттєво розширив інструментарій аналізу даних соціальних мереж. Платформи соціальних медіа щоденно генерують величезні обсяги мультимедіального контенту, що залучає текстові повідомлення, коментарі, зображення та відеоматеріали, які в свою чергу відображають реакції користувачів на соціальні, політичні та культурні події. Аналіз цих даних має значний інтерес для дослідників у галузі соціальних наук, обчислювальної лінгвістики та аналізу даних.

Великі мовні моделі (Large Language Models, LLMs) демонструють здатність відображати те, як люди структурують знання та смисли в межах конкретних предметних областей, втілюючи ці знання через концептуальну організацію мови. Згідно з [1], ця здатність проявляється в умінні моделей відповідати на запитання, генерувати зв'язний текст та виводити інформацію про сутності, об'єкти й властивості, з якими вони безпосередньо не стикалися під час навчання. Хоча мовні моделі не володіють знаннями у позамовному досвіді, логічні структури, які вони уловлюють між лінгвістичними поняттями, можуть наближатися до когнітивної організації концептуальних знань людини. Це свідчить про потенціал використання LLMs не лише як інструментів обробки тексту, а і як засобів моделювання структур смислу в публічному дискурсі.

У цьому зв'язку в останні роки дедалі більше уваги приділяється застосуванню великих мовних моделей для класифікації та відображення абстрактних концептуальних категорій, зокрема поведінкових емоцій, етичної відповідності, соціальних і культурних норм, а також правових оцінок у великих текстових корпусах. Цей новий напрям досліджень демонструє поступовий зсув від поверхневого мовного моделювання до більш ефективної інтерпретації складних, абстрактних, взаємоперекривних та морально значущих категорій.

Здатність великих мовних моделей працювати з абстрактними та ціннісно навантаженими концептуальними категоріями створює підґрунтя для їх застосування в аналізі соціально значущих дискурсів. Зокрема, LLMs можуть бути використані для інтерпретації суспільних настроїв і позицій, що формуються в онлайн-просторах, де мовні, емоційні та нормативні аспекти тісно переплітаються. Сучасні моделі дають змогу глибше досліджувати суспільні настрої, думки та ставлення до різних соціальних питань, забезпечуючи масштабований і більш контекстно чутливий аналіз порівняно з класичними методами машинного навчання.

Цей перехід від імовірнісного до генеративного моделювання підвищує аналітичну точність, оскільки дає змогу дослідникам виявляти приховані смислові структури та динамічні взаємодії між темами, які

більш адекватно відображають складність публічного дискурсу. Але підхід до аналізу соціально значущих дискурсів відкриває можливості не лише для опису поточних станів суспільних настроїв, а й для виявлення їхньої динаміки в часі та просторових відмінностей між спільнотами. Як показано в роботі [2], застосування генеративних моделей у цьому контексті дозволяє переходити від статичного аналізу окремих текстових фрагментів до системного вивчення еволюції тем, позицій і форм колективного осмислення соціальних проблем у цифровому середовищі. Таким чином, аналітичний потенціал LLMs може бути безпосередньо пов'язаний із ширшими соціальними цілями, зокрема з підтримкою процесів суспільної взаємодії та згуртованості.

Водночас, попри те що наявні дослідження демонструють ефективність підходів GenAI для виявлення структурованих суспільних настроїв, ідентифікації впливових позицій та картографування широкого публічного дискурсу, вони здебільшого не зосереджуються на тому, яким чином такі інструменти можуть бути використані для систематичного зіставлення громадянських наративів із конкретними державними нормами крізь призму етичних, правових, моральних і соціальних вимірів, а також для оцінювання потенціалу таких зіставлень у контексті зміцнення суспільної згуртованості.

Особливої актуальності ці обмеження набувають у випадках, коли предметом аналізу є публічний дискурс маргіналізованих спільнот, зокрема груп мігрантів, де питання представлення інтересів і суспільної згуртованості є критично важливими. Ці труднощі зумовлені не лише обмеженою доступністю та видимістю власних позицій мігрантів у публічному дискурсі, а й складністю інтеграції таких даних з офіційними наборами даних.

У відповідь на ці методологічні виклики в даному дослідженні виконується перевірка здатності моделей GenAI виявляти та інтерпретувати етичні виміри в текстовому контенті. При цьому емпірична перевірка цього потенціалу здійснюється на матеріалі публічних і офіційних наративів, пов'язаних із проблематикою вимушеної міграції, зокрема українських біженців, які були змушені залишити країну внаслідок повномасштабного російського вторгнення.

Випадок українських біженців становить репрезентативний і водночас методологічно релевантний кейс з огляду на масштаб і терміновість гуманітарної кризи: за даними Управління Верховного комісара ООН у справах біженців (УВКБ ООН, 2024), до кінця 2024 року понад шість мільйонів українців зазнали вимушеного переміщення. Такий контекст формує насичене та різноманітне середовище даних для перевірки здатності моделей GenAI працювати з нормативно навантаженими та ціннісно чутливими категоріями.

З позицій підходу «етика в дизайні» (Ethics-in-Design) дослідження спрямоване на оцінювання того, наскільки інструменти штучного інтелекту можуть бути використані для аналізу соціально значущих проблем [3] і, ширше, для підтримки процесів суспільної згуртованості шляхом виявлення потенційних розбіжностей між офіційними нормативними рамками та наративами маргіналізованих груп.

### 1. Аналіз наукових підходів до проблеми

У сучасних дослідженнях дедалі більше уваги приділяється питанням етики застосування штучного інтелекту в гуманітарній сфері та контексті міграції. Зокрема, у статті Міжнародного Червоного Хреста [4] наголошується, що у кризових ситуаціях використання алгоритмів повинно ґрунтуватися на нормах міжнародного права з прав людини, а також доповнюватися етичними принципами прозорості, підзвітності та справедливості. Подібні рекомендації містяться й у матеріалі ASAT Labs — незалежної дослідницько-аналітичної ініціативи у сфері етичного використання штучного інтелекту в гуманітарних проєктах. В цих матеріалах особлива увага приділяється людиноцентризму дизайну, захисту даних біженців та інклюзивності під час проєктування систем [5]. Приклад практичного впровадження таких підходів можна знайти у ініціативах Міжнародного комітету порятунку (IRC), який застосував чатботи у програмі Signpost для підтримки переміщених осіб, проте водночас визначив необхідність анонімізації даних та залучення людини до перевірки відповідей, аби уникнути ризиків поширення неправдивої або шкідливої інформації [6].

Дослідники також звертають увагу на проблему репрезентації біженців у цифровому середовищі. Так, у блозі Університету Організації Об'єднаних Націй (United Nations University, UNU) зазначається, що створення AI-агентів чи цифрових аватарів біженців може призвести до їх відчуження від власного голосу та посилення стереотипів, якщо такі технології не контролюються самими постраждалими спільнотами [7].

У роботі [8] продемонстровано, як загальні принципи етики штучного інтелекту (автономія, справедливість, запобігання шкоді, прозорість) можна трансформувати у конкретні технічні вимоги для систем прогнозування у сфері міграційного менеджменту. Це відповідає ширшій тенденції до розвитку підходів Value Sensitive Design, спрямованих на вбудовування людських цінностей у процес проєктування технологій [9]. Подальші емпіричні дослідження засвідчують, що для практиків і законодавців ключовими проблемними аспектами залишаються прозорість, підзвітність і захист конфіденційності, тоді як одним із найбільших викликів є брак ефективних механізмів регуляторного забезпечення [10].

Важливим для розуміння загальної картини є також Звіт про стан етики штучного інтелекту Монреальського інституту етики ШІ, який відстежує ключові тенденції в екосистемі, зокрема питання упередженості, регуляторного середовища та суспільного сприйняття. [11].

У ряді останніх досліджень значну увагу також приділено етичним рамкам і практикам, що стосуються мігрантів та біженців, особливо молоді та вразливих груп. Наприклад, стаття [12] аналізує, як методологічні підходи до роботи з молоддю-біженцями можуть враховувати етичні дилеми, такі як асиметрія влади, питання відповідальності, приватності й довірчих відносин між дослідниками та учасниками. Інше дослідження [13] звертається до інклюзивності та етичних норм у контексті складної доступності мігрантів під час гуманітарних криз, підкреслюючи необхідність забезпечення захисту прав, гідності та врахування голосів тих, хто зазвичай маргіналізується.

Також важливо звернути увагу на праці, що розглядають етичні аспекти застосування ШІ в міграційному менеджменті. Зокрема, стаття [8] обговорює, як принципи етики, так як прозорість, справедливість, недопущення шкоди можуть бути вбудовані у предикативні інструменти, які використовуються гуманітарними організаціями, щоб мінімізувати ризики і захистити права мігрантів.

Крім того, дослідження [14] ілюструє, як медіа через візуальні образи відображають біженців, надаючи їм різні суб'єктивні позиції (вразливі жертви, особи, що зазнають страждань, резиліентні), що також має етичне значення у формуванні суспільного сприйняття.

У локальному українському контексті прикметним є проведення Літньої школи «Ethics&AI Summer School»<sup>1</sup> в Українському католицькому університеті у 2023 році, де обговорювалися питання довіри, корупційних ризиків та синтетичного контенту. Крім того, у 2025 році Українська національна асоціація адвокатів приєдналася до розробки глобальних стандартів етики використання штучного інтелекту спільно з міжнародними організаціями<sup>2</sup>.

## 2. Методологія дослідження

Дослідження проходило в кілька етапів: спочатку було визначено географічні межі роботи, далі - знайдено та сформовано перелік офіційних джерел інформації й Telegram-каналів та чатів, після чого здійснено збір та обробку даних. Схематично повна методологія відображена на рис. 1.

Створення датасету передусім починається з визначення напрямків міграції біженців, що є надзвичайно важливим для перевірки точності даних і достовірності інформації. У межах дослідження були визначені основні напрями міграції українців у Європі та Північній Америці, зокрема: Україна, Польща, Канада, Сполучені Штати Америки, Словаччина, Молдова, Німеччина, Чехія, Іспанія, Велика Британія, Італія, Швеція та Ірландія<sup>3</sup>. Отримані результати

- 1 UCU Ethics&AI Summer School 2023 – <https://apps.ucu.edu.ua/events/ucu-ethics-ai-summer-school-2023/>
- 2 Ukrainian advocates join efforts to develop AI ethics standards – <https://en.unba.org.ua/activity/news/10435-ukrainian-advocates-join-efforts-to-develop-ai-ethics-standards.html>
- 3 Eurostat – [https://ec.europa.eu/eurostat/databrowser/view/MIGR\\_ASYTPSM\\_\\_custom\\_7394287/default/table?lang=en](https://ec.europa.eu/eurostat/databrowser/view/MIGR_ASYTPSM__custom_7394287/default/table?lang=en)

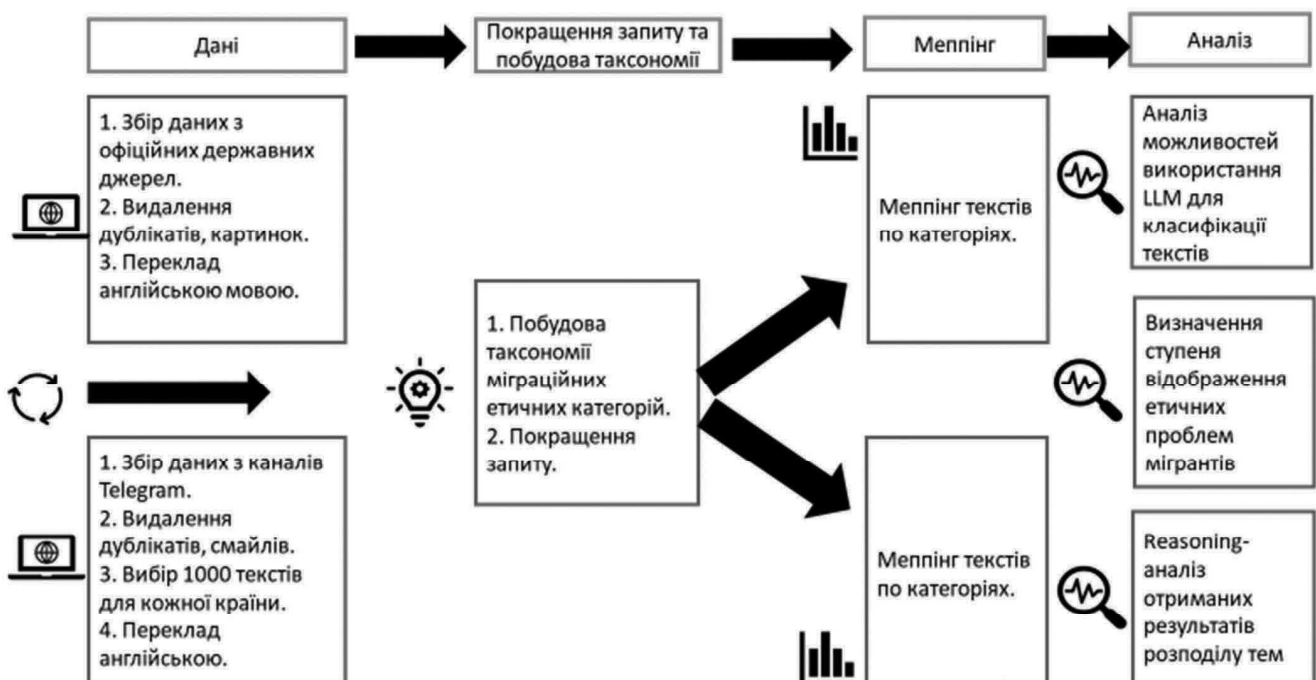


Рис. 1. Чотири етапи методології дослідження

також створюють базу для подальшого порівняльного аналізу між офіційними джерелами різних країн і реальними тенденціями, відображеними у відкритих комунікаційних середовищах, таких як соціальні мережі та месенджери.

Для початку ми вирішили створити 2 датасети, де перший містить публічну інформацію від медіаресурсів держав та міжнародної організації з міграції [15], а другий включає дискурс мігрантів у соціальній мережі Telegram [16].

Для першого датасету ми обробили більше 30 джерел таких як:

- офіційні вебсайти міграційних служб і міністерств закордонних справ різних країн;
- посольства України в країнах перебування;
- YouTube-канали державних установ;
- акаунти урядовців у мережі Twitter;
- державні вебресурси, зокрема сайти міністерств закордонних справ, охорони здоров'я та освіти;
- українські урядові структури, що надають підтримку біженцям за кордоном;
- вебсайт Міжнародної організації з міграції (ІОМ), включно з її українським представництвом.

В результаті було отримано 708 текстів, з яких близько третини припадає на статті з сайту ІОМ.

Формати проаналізованих даних включали тексти вебсайтів, PDF-файли, твіти, а також аудіо- та відеофайли з YouTube. Збирання текстів із вебсайтів здійснювалося за допомогою бібліотеки Python BeautifulSoup. Для обробки PDF-документів використовували бібліотеку PyPDF2, тоді як обробку сирого тексту забезпечувала бібліотека WordNinja, а визначення мови контенту виконувала бібліотека LangDetect. Текстові дані з відеофайлів отримували через вебсервіс Sonix, який надає транскрипцію аудіо у текст на кількох мовах. Для невеликих джерел або вебсайтів із обмеженнями на скрапінг застосовувався ручний парсинг. Кроки попередньої обробки тексту включали очищення (видалення емодзі, символів та іконок), лематизацію та видалення порожніх і дубльованих записів.

Для датасету публічного дискурсу було зібрано великий масив повідомлень із Telegram-груп, де українські біженці обмінюються інформацією та обговорюють актуальні питання. Такі чати стали основою спілкування і взаємної допомоги біженців у приймаючих країнах. Для пошуку відповідних каналів використовували ключові слова. Після початку повномасштабного вторгнення Росії активність у таких групах значно зросла, що перетворило Telegram на важливий канал комунікації для українців [17]. Було зібрано 632 372 повідомлення в соціальній мережі. Для забезпечення рівномірного розподілу текстів по країнах було зроблено обмеження до 1000 випадкових текстів з кожної з країн, що призвело до створення датасету в 12000 повідомлень. Табл. 1 відображає розмір остаточно використаних у дослідженні датасетів.

Таблиця 1

## Об'єм створених датасетів

Назва датасету	Кількість текстів
Датасет офіційних джерел	708
Датасет соціальної мережі	12000

Для вибору мовної моделі було здійснено порівняльний аналіз затримки відповіді (latency) та вартості використання популярних великих мовних моделей, що застосовуються для автоматизованого аналізу текстів. Оцінювання цих параметрів є критично важливим для досліджень, орієнтованих на обробку корпусів текстових даних.

З погляду затримки відповіді, GPT-3.5 Turbo демонструє одну з найнижчих затримок серед великих мовних моделей в відкритих бенчмарках, що пояснюється його оптимізацією для швидкої генерації відповідей. Модель Claude 3 також є конкурентною за швидкістю, однак фактичні показники суттєво залежать від конкретного підтипу (Opus або Sonnet) та платформи розгортання. Для LLaMA-2 швидкодія значною мірою визначається розміром моделі та апаратним забезпеченням: компактні версії (7B) можуть забезпечувати дуже низьку затримку на оптимізованих GPU, тоді як великі моделі часто мають більшу затримку відповіді без пропорційного зростання точності.

Не менш важливим є економічний аспект. GPT-3.5 Turbo має фіксовану ціну \$0.50 за 1 млн вхідних токенів та \$1.50 за 1 млн вихідних токенів, тоді як Claude 3 є значно дорожчим і коштує \$3 та \$15 відповідно. Для LLaMA-2 централізоване ціноутворення відсутнє, а вартість залежить від провайдера. Найнижча зафіксована ціна становила \$0.36 за 1 млн токенів як для входу, так і для виходу. За умовного сценарію аналізу тексту з 250 вхідними та 50 вихідними токенами орієнтовна вартість одного запиту становить \$0.0002 для GPT-3.5 Turbo, \$0.0015 для Claude 3 та \$0.000108 для LLaMA-2. У табл. 2 представлено порівняння трьох великих мовних моделей за показниками, релевантними для застосування запропонованої методики.

З огляду на наведене порівняння, у межах даного дослідження було обрано модель GPT-3.5 Turbo як основний інструмент аналізу. Такий вибір зумовлений оптимальним поєднанням низької затримки відповіді, стабільної якості результатів та високої економічної ефективності, а також легкості використання при масовій обробці текстів. На відміну від дорожчих моделей, ChatGPT забезпечує достатній рівень точності для задач тематичної класифікації на основі таксономії без суттєвого зростання обчислювальних витрат. Порівняно з self-hosted рішеннями на базі LLaMA-2, використання ChatGPT дозволяє уникнути додаткових витрат на інфраструктуру та технічне обслуговування, що є критично важливим для відтворюваності та масштабованості дослідження.

Таблиця 2

## Порівняння великих мовних моделей

Характеристика	GPT-3.5	Claude 3	LLaMA-2
Легкість використання	Найзручніший API на ринку: SDK для більшості мов, просте налаштування, детальна документація.	Повноцінний API від Anthropic. Простий REST/SDK, гарна документація, потокове передавання, стабільний час безперебійної роботи.	Немає офіційного централізованого API. Тільки самостійний хостинг або через провайдерів. Зручність залежить від використання та потребує глибшого дослідження.
Швидкість	Найнижча затримка серед масових моделей, стабільна швидкість навіть під навантаженням.	Високий, але трохи повільніший, ніж GPT-3.5 Turbo	Сильно залежить від хостингу та розміру моделі: 7B – швидкий, 70B – повільніший
Ціна за 1 млн токенів	\$3 вхід/\$15 вихід	\$0.5 вхід/\$1.5 вихід	Від \$0.36 за вихід та вхід
Приблизна вартість аналізу (250 вхід + 50 вихід токенів)	\$0.0002 (0.02 цента)	\$0.0015 (0.15 цента)	\$0.000108 (0.0108 цента)

Таким чином, для визначення етичної складової текстів ми використали OpenAI API та модель “GPT-3.5-turbo-16k”. Ця модель була обрана через її здатність обробляти великі тексти, що було необхідно для визначення тематики довгих документів. Крім того, модель оптимізована для швидшої обробки, що дозволяє ефективно класифікувати тексти за задалегідь визначеними категоріями. При цьому вона зберігає весь спектр можливостей, включно з розумінням мови, генерацією тексту, узагальненням та іншими функціями.

На другому етапі нами було створено таксономію тем етичного дискурсу в текстах, пов’язаних з міграцією. Для цього ми виокремили десять наукових статей, які були визначені академічною пошуковою платформою Google Scholar за період з 2022 по 2025, що стосуються етичних проблем біженців. Після цього з використанням GenAI за допомогою промптів ми виокремили більше 200 сутностей з цих статей. Деякі з них містили дублікати та сутності, які не стосувалися міграції чи етики. Після видалення зайвих сутностей ми отримали більше 66 сутностей, які були розподілені на 7 типів. Табл. 3 вказує на верхній рівень класифікації сутностей, систематично організованих навколо центральної теми міграції та чітко розподілених за етичними аспектами.

Таблиця 3

## Кількість сутностей таксономії

Назва типу	Кількість сутностей
Міграційний статус та наміри	10
Етичні міркування	19
Міграція та емоційні наслідки	8
Етика досліджень	13
Міграційна політика	4
Права людини та соціальна справедливість	8
Методологія дослідження міграції	4
Всього	66

На наступному етапі нашого дослідження ми проаналізували ключові типи, відібрані системою обробки природної мови, представленою моделлю GPT. Оскільки ми прагнули застосовувати механізми штучного інтелекту для підтримки урядових структур, ми додали експертну перевірку правильності інтерпретації концептів, пов’язаних з етикою міграції [18].

1. *Міграційний статус та наміри*. Цей тип охоплює різні статуси міграції, такі як біженці, шукачі притулку, переміщені особи, особи з нестабільним статусом, а також добровільну міграцію, включно з економічними мігрантами, возз’єднанням сімей та процесами інтеграції культурної та економічної. Він дозволяє зрозуміти причини, мотиви та правовий статус мігрантів у новому середовищі.

2. *Етичні міркування*. Цей тип зосереджується на етичних принципах та практиках, включно з етичною відповідальністю, дилемами, викликами та стратегіями пом’якшення. Він також охоплює етичні теорії, такі як догляд, автономія, справедливість, солідарність та довіра, що формують основу для морального аналізу міграційних процесів.

3. *Міграція та емоційні наслідки*. Тут розглядаються емоційні впливи міграції, зокрема емоційні зв’язки, психологічні труднощі, культурний шок, стрес, конфлікти, відчуження та адаптація до повсякденного життя, особливо для малодохідних груп населення. Це допомагає оцінити психосоціальний стан мігрантів і їхню здатність інтегруватися в нове середовище.

4. *Етика досліджень*. Цей тип включає етичні міркування у дослідженнях міграції, такі як інформована добровільна згода, конфіденційність, приватність, оцінка ризиків та переваг, компетентність та асиметрії влади. Також розглядаються специфічні етичні аспекти міграції, наприклад етична трансформація, обмін даними, маскуваність даних та прозорість.

5. *Міграційна політика.* Цей тип аналізує вплив міграційної політики, включно з розробкою політик, їхнім впливом, рекомендаціями та міграційними наративами. Він дозволяє зрозуміти, як політичні рішення формують умови для мігрантів та відображають соціальні пріоритети.

6. *Права людини та соціальна справедливість.* Тут розглядаються питання вразливих груп населення, включно з особами без документів, вразливими групами та громадами прийому, а також проблеми соціальної справедливості, такі як рівність можливостей, культурне різноманіття та суспільні уявлення. Це

підкреслює важливість захисту прав мігрантів і забезпечення рівності в суспільстві.

7. *Методології дослідження міграції.* Цей тип охоплює методи дослідження міграції, включно з моніторингом міграційних потоків, використанням політичних інформаторів, міжурядовим збором даних та статистикою громадянського суспільства. Він спрямований на забезпечення об'єктивного аналізу міграційних процесів для прийняття обґрунтованих рішень.

Ієрархія даної таксономії, що включає в себе 4 рівні, зображена на рис. 2:

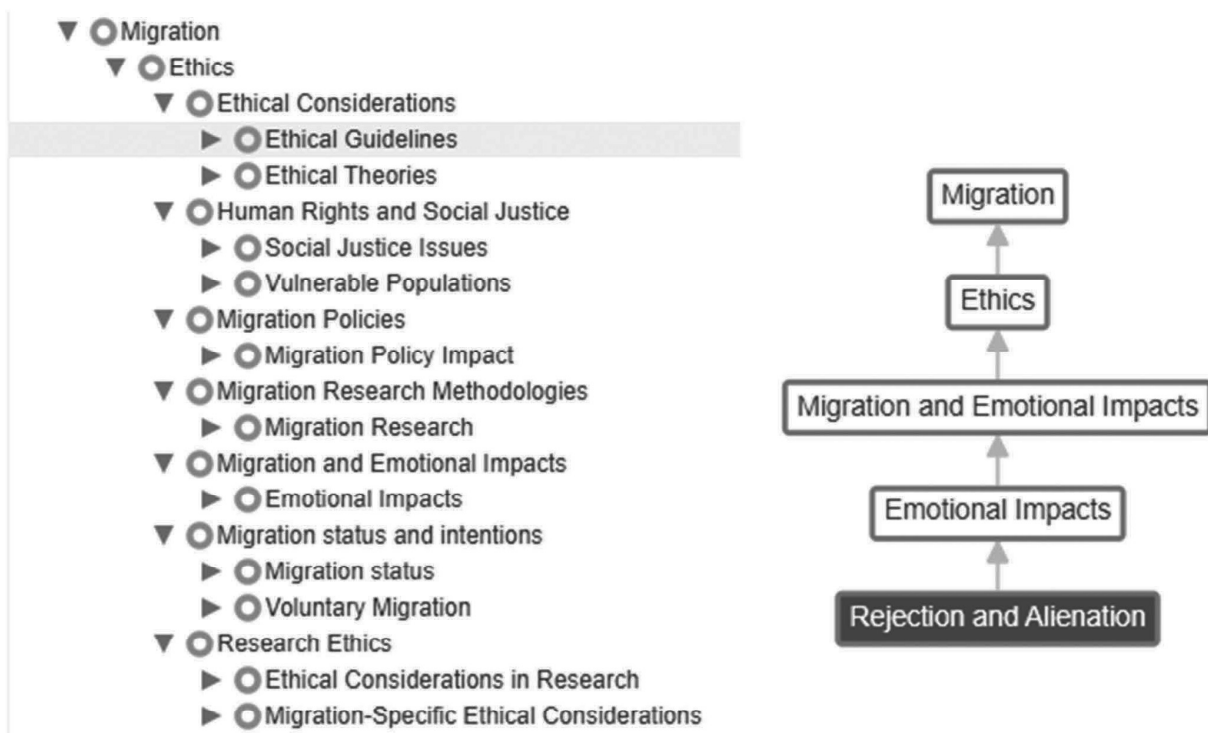


Рис. 2. Ієрархія розробленої таксономії

Наступним кроком було визначено, які тексти з двох датасетів мають відношення до етичних аспектів, а які відносяться до інших питань міграції.

Взаємодія з моделлю GPT організована за допомогою основної функції для визначення типу етичного контексту тексту. Для цього було використано наступний промпт :

*“Here is a taxonomy of types of ethics topics in migration texts. Taxonomy:...*

*I will send you a file with texts and your task is to determine which type every text is and provide reasoning for choosing this type. Types are under numbers 1-7. Don't choose any other options or return None for the type or reason. Always choose between 7 types. Provide me with a CSV file with columns: text\_number, text, type\_name, reason”.*

Створена функція також враховувала можливі обмеження по токенах та дотримувалася правил

використання API. Цінними є також отримані інсайти щодо причин, з яких модель обирає конкретний тип.

Ми провели кількісний аналіз того, як часто публікації офіційних органів влади та дискурси користувачів соціальних мереж спиралися на наративи етики в кожному окремому наборі даних. Такий підхід дозволив оцінити частоту обговорення етичних аспектів міграції у різних типах дискурсів та визначити, наскільки добре сприймається відображення етичних проблем у суспільстві.

Додатково ми здійснили якісний аналіз текстів у цих двох наборах даних, щоб ідентифікувати, які складові елементи етики зустрічаються у публікаціях влади та обговореннях у соціальних мережах у наративах міграції. Це допомогло структурувати матеріал та підготувати його для подальшого порівняльного аналізу. Цей підхід дозволив виявити не лише частоту згадок, а й якісну різницю у фокусі обговорення

етичних питань між офіційними та громадськими дискурсами. Завдяки здатності моделі GPT до аргументування та міркування, ми отримали можливість враховувати контекст, взаємозв'язки між поняттями та тональність висловлювань, що забезпечує більш точний та комплексний аналіз текстового матеріалу.

### 3. Результати

В результаті дослідження було анотовано 2 тексти з офіційних джерел та 37 текстів з повідомлень у соціальній мережі. Що зрозуміло, свідчить про дуже низький рівень висвітлення етичних проблем як в дискурсі політиків, так і в дискурсі мігрантів. Табл. 4 висвітлює отриманий результат з відображення етичних елементів текстів, які пов'язані з міграцією.

Таблиця 4

#### Рівень висвітлення етичних аспектів у суспільстві

Назва датасету	Кількість текстів на тему етики	Загальна кількість текстів	% етичних текстів
Датасет офіційних джерел	2	708	0.27%
Датасет соціальної мережі	37	12000	0.31%

Щодо етичних тем що були відображені у цих текстах, то в офіційних джерел по одному тексту було з тем “Міграційний статус та наміри” та “Міграційна політика”. Розподіл тем, що пов'язані з етичною складовою у чатах неофіційних повідомлень Telegram відображено у табл. 5.

Таблиця 5

#### Розподіл текстів за темами етичного дискурсу

Тема етики	% теми серед текстів, що мають етичну складову
Етичні міркування	70.5%
Міграція та емоційні наслідки	11%
Міграційна політика	8%
Права людини та соціальна справедливість	5.5%
Міграційний статус та наміри	2.5%
Етика досліджень	2.5%
Методологія дослідження міграції	0%

Таким чином, найбільшу кількість текстів система віднесла до загальних етичних та комунікативних орієнтирів, тоді як меншою мірою були представлені теми, пов'язані з емоційним досвідом, міграційною політикою та соціальною справедливістю. Окремі тексти торкалися питань статусу мігрантів, інтеграції та дослідницької етики.

Додатково слід зазначити, що аргументування у межах етичної проблематики міграції є складним,

оскільки значна частина текстів апелює до абстрактних моральних орієнтирів, які не мають чітких меж між категоріями. Такі аргументи, що стосуються справедливості та вразливих груп, часто перетинаються з аргументами, що розглядають етичні та соціальні питання. Аналогічно, аргументи, що стосуються загальних моральних, комунікативних або поведінкових орієнтирів, можуть одночасно включати елементи емоційних переживань, адаптації або психологічні аспекти, що ускладнює їх однозначну інтерпретацію та класифікацію. Табл. 6 відображає аргументування вибору теми, що створено GenAI, з використанням експертної складової

Таблиця 6

#### Аргументація вибору теми

Аргументування	% використання
Загальні моральні, комунікативні або поведінкові орієнтири	59.5%
Розглядає етичні, моральні або питання соціальної справедливості	10.8%
Описує емоційні переживання, адаптацію або психологічні аспекти	10.8%
Згадує міграційну політику, безпечно працевлаштування або інституційні рамки	8.1%
Стосується справедливості, рівності або вразливих груп населення	5.4%
Обговорює міграційний статус, питання біженців або інтеграційні аспекти	2.7%
Розглядає дослідницькі практики, інформовану згоду або етику даних	2.7%

#### Висновки

Використання моделі GPT для аналізу текстів дозволило застосувати підхід, заснований на міркуванні та семантичному виведенні, який виходить за межі простого підрахунку слів і зосереджується на смислових зв'язках та контексті етичних понять із використанням таксономії сутностей у межах етичного аспекту міграції. Завдяки цьому стало можливим виявити, як різні типи дискурсу — офіційний і публічний — репрезентують моральні аспекти міграції та які теми при цьому актуалізуються. Такий підхід не лише поглибив розуміння структури комунікації, а й відкрив перспективи для використання штучного інтелекту як аналітичного інструмента в гуманітарних дослідженнях.

Використання розробленої таксономії відіграє ключову роль у підвищенні точності роботи моделі, орієнтованої на міркування та семантичне виведення, оскільки вона задає чіткі концептуальні межі для класифікації етичних тем. Таксономія допомагає моделі не лише точніше розпізнавати релевантні контексти, а й формувати більш послідовну логіку аргументації під час віднесення текстів до певних категорій. Крім того,

така структурована система полегшує інтерпретацію результатів як для дослідників, так і для потенційних державних застосувань, де важливо забезпечити прозорість та відтворюваність процесів автоматизованого аналізу етичних аспектів у публічному дискурсі.

Крім того, проведене дослідження дозволило виявити особливості того, як етичні питання міграції біженців, спричинені війною, висвітлюються у двох вимірах суспільної комунікації. Аналіз показав, що в обох типах джерел етична тематика трактується мало, а її присутність часто залишається опосередкованою через згадки про справедливість, солідарність, турботу чи соціальну підтримку. Це свідчить про потребу у більш системному врахуванні морального виміру міграційної політики, особливо в умовах гуманітарних криз.

Порівняльний аналіз між публічним та офіційним дискурсами висвітлює значну асиметрію у розподілі етичних тем. У соціальних мережах переважає категорія “Етичні міркування”, що свідчить про підвищений інтерес користувачів до моральних аспектів міграції, зокрема до етичних дилем, відповідальності та справедливості. Меншою мірою обговорювалися теми емоційних наслідків міграції, міграційної політики та прав людини й соціальної справедливості. Натомість у офіційному дискурсі спостерігається обмежене представлення етичних питань: виявлено лише два тексти. Один, що стосується міграційного статусу та намірів, та другий, що стосується міграційної політики. Такий дисбаланс підкреслює, що етичний вимір більше зосереджений на етичних міркуваннях у громадському середовищі, ніж у комунікаціях державних структур, що може вказувати на потребу посилення етичної складової в офіційному обговоренні міграційних процесів.

Результати дослідження можуть бути використані державними інституціями для покращення публічної комунікації щодо міграційної політики, зокрема шляхом формування більш чутливих до етичних контекстів повідомлень. Для суспільства ці результати є важливими у підвищенні рівня усвідомлення моральних аспектів прийняття біженців, формуванні толерантності та розвитку культури солідарності. Такі знання також можуть сприяти вдосконаленню механізмів взаємодії між владою, громадськістю та міжнародними організаціями у питаннях гуманітарної підтримки.

Разом із тим, узагальнення та використання отриманих результатів у практиці державного управління та суспільної комунікації потребують обережності з огляду на низку методологічних і емпіричних обмежень дослідження.

По-перше, емпіричний аналіз був зосереджений виключно на етичних вимірах публічного дискурсу, тоді як правові та соціальні аспекти розглядалися переважно на концептуальному рівні. Такий фокус

звужує можливості комплексного аналізу нормативних взаємозв'язків і потребує розширення в подальших дослідженнях.

По-друге, операціоналізація абстрактних етичних категорій залишається методологічно складним завданням. Формалізація ціннісно навантажених понять у вигляді таксономії є проблематичною не лише для моделей штучного інтелекту, а й для експертів-людей, що може впливати на стабільність класифікації та інтерпретованість результатів. Додатково, окремі категорії мають перехресний характер, що зумовлює неоднозначності під час маркування та потребує подальшого уточнення таксономії й розвитку підходів «людина—ШІ».

По-третє, узагальнюваність результатів обмежується репрезентативністю та доступністю даних соціальних мереж. Аудиторії платформ не є демографічно нейтральними, а доступ до агрегованих даних часто обмежується великими технологічними компаніями, що може призводити до часткового або упередженого відображення публічного дискурсу [19].

Крім того, багатомовність корпусу (українська/російська та англійська мови) могла спричинити семантичні розбіжності, не повністю усунуті на етапі попередньої обробки, зокрема під час перекладу нормативно навантажених висловлювань.

Нарешті, попри залучення матеріалів із кількох країн за період 2022–2024 років, дослідження обмежене часовими рамками та не враховує ширших культурних і політичних чинників, що можуть впливати на зіставлюваність наративів між різними національними контекстами.

Подальші роботи можуть розширити часовий горизонт, застосувати багатомовні моделі та глибше інтегрувати контекстуальні змінні для підвищення валідності та відтворюваності результатів.

Отже, дослідження продемонструвало, що поєднання штучного інтелекту та етичного аналізу створює нові можливості для більш глибокого розуміння суспільних процесів, пов'язаних із міграцією. Врахування етичних вимірів у політичних рішеннях та публічних комунікаціях може стати важливою передумовою підвищення соціальної згуртованості та зміцнення довіри між державою та громадянами.

#### Список літератури:

- [1] Du, C., Fu, K., Wen, B., Sun, Y., Peng, J., Wei, W., Gao, Y., Wang, S., Zhang, C., Li, J. and Qiu, S., (2024), “Human-like object concept representations emerge naturally in multimodal large language models”. arXiv preprint arXiv:2407.01067.
- [2] Khairova, N., Rizun, N., Alexopoulos, C., Ciesielska, M., Lukashevskiy, A. and Redozub, I., (2024), “Understanding the Ukrainian Migrants Challenges in the EU: A Topic Modeling Approach”, In Proceedings of the 25th Annual International Conference on Digital Government Research (pp. 196-205). <https://dl.acm.org/doi/10.1145/3657054.3657252>

- [3] Dignum, V. (2018). Ethics in artificial intelligence: introduction to the special issue. *Ethics and Information Technology*, 20(1), 1-3.
- [4] International Committee of the Red Cross. (2020). AI for humanitarian action: Human rights and ethics. *International Review of the Red Cross*. <https://international-review.icrc.org/articles/ai-humanitarian-action-human-rights-ethics-913>
- [5] ASAT Labs. (2021). 5 Ethical AI Principles Every NGO Should Follow in Humanitarian Projects. <https://asatlabs.org/5-ethical-ai-principles-every-ngo-should-follow-in-humanitarian-projects>
- [6] AP News. (2023). Refugee aid groups turn to AI chatbots, raising ethical questions. <https://apnews.com/article/9c5fae949c429c38019feabaec5fafa>
- [7] United Nations University. (2022). The role of AI agents in humanitarian action: Balancing ethics and innovation. <https://unu.edu/cpr/blog-post/role-ai-agents-humanitarian-action-balancing-ethics-and-innovation>
- [8] Guillén, M., & Teodoro, A. (2023). Embedding Ethical Principles into AI Predictive Tools for Migration Management in Humanitarian Action. *Social Sciences*, 12(2), 53. <https://doi.org/10.3390/socsci12020053>
- [9] Friedman, B., & Hendry, D. G. (2019). *Value Sensitive Design: Shaping Technology with Moral Imagination*. MIT Press.
- [10] Jobin, A., Ienca, M., & Vayena, E. (2022). AI Ethics: An Empirical Study on the Views of Practitioners and Lawmakers. arXiv preprint arXiv:2207.01493. <https://arxiv.org/abs/2207.01493>
- [11] Montreal AI Ethics Institute. (2022). State of AI Ethics Report (Volume 6). <https://arxiv.org/abs/2202.07435>
- [12] Huizinga R. P. et al. Ethical and Methodological Considerations in Research with Asylum-Seeking and Refugee Youth in European Cities // *Social Sciences*. — 2025. — Т. 14. — № 4. — С. 204.
- [13] Ullah A. K. M. A. Ethical framework and inclusivity: research mechanics of difficult-to-reach migrants in civil military context // *Journal of International Humanitarian Action*. — 2024. — Т. 9. — №. 1. — С. 11.
- [14] Martikainen J., Sakki I. (2024) Visual humanization of refugees: A visual rhetorical analysis of media discourse on the war in Ukraine // *British Journal of Social Psychology*. — 2024. — Т. 63. — №. 1. — С. 106-130.
- [15] Redozub, I., and Khairova, N. (2025), “Ukrainian Migrants: Official Sources Discourse (2022–2023)”, Zenodo, available at: <https://doi.org/10.5281/zenodo.17565879> (accessed 9 November, 2025)
- [16] Khairova, N., Lukashchuk, A., & Redozub, I., (2025), “Ukrainian migrants Telegram discourse (2022–2023)”, Zenodo, available at: <https://doi.org/10.5281/zenodo.17566020> (accessed 9 November, 2025)
- [17] Ghasiya, P. and Sasahara., K. (2023) Messaging strategies of ukraine and russia on telegram during the 2022 russian invasion of ukraine. *First Monday*, 2023.
- [18] Dignum, V., Baldoni, M., Baroglio, C., Caon, M., Chatila, R., Dennis, L., ... & De Wildt, T. (2018, December). Ethics by design: Necessity or curse?. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 60-66).
- [19] Lutz, C., (2022), “Inequalities in social media use and their implications for digital methods research”, in A. Quan-Haase & L. Sloan (Eds.), *The SAGE Handbook of Social Media Research Methods*, pp. 679-690, SAGE.

*Received (Надійшла) 28.01.2026*

*Accepted for publication (Прийнята до друку) 26.02.2026*

*Publication date (Дата публікації) 27.03.2026*



О. І. Богатов<sup>1</sup>, Г. А. Плехова<sup>2</sup>, М. В. Костікова<sup>3</sup>, С. М. Неронов<sup>4</sup>, Д. О. Плехов<sup>5</sup>

<sup>1</sup>ХНАДУ, м. Харків, Україна, bogatovolegigor@ukr.net, ORCID iD: 0000-0001-7342-7556

<sup>2</sup>ХНАДУ, м. Харків, Україна, plehovaanna1@gmail.com, ORCID iD: 0000-0002-6912-6520

<sup>3</sup>ХНАДУ, м. Харків, Україна, kmv\_topaz@ukr.net, ORCID iD: 0000-0001-5197-7389

<sup>4</sup>ХНАДУ, м. Харків, Україна, sernikner@gmail.com, ORCID iD: 0000-0003-2381-1271

<sup>5</sup>ХНАДУ, м. Харків, Україна, plehov@gmail.com, ORCID iD: 0009-0004-7873-1716

## ВИЗНАЧЕННЯ ОБМЕЖЕНЬ У ВИКОРИСТАННІ СТАТИСТИЧНИХ ОЦІНОК ПРОФЕСІЙНОГО РИЗИКУ ЗА ОФІЦІЙНИМИ ДАНИМИ В УКРАЇНІ

Робота присвячена аналізу можливих умов і галузей застосування відомих методів оцінки професійних ризиків в державній системі управління охороною праці. Також показано актуальність цього питання, визначено обмеження щодо застосування статистичних оцінок професійного ризику в процесі управління станом небезпечних об'єктів. Набуло подальшого розвитку ризик-орієнтоване управління охороною праці, яке повинно враховувати обмеження у використанні ймовірнісно-статистичних методів визначення професійного ризику, що пов'язані як із конкретними кількісними показниками, які використовуються для розрахунків, так і з різним підходом визначення галузей економіки, який наводиться в даних Держпраці та Держстатистики. Висновки за даними Держпраці або Пенсійного фонду можна застосовувати до показників Держстатистики з певними обмеженнями.

ПРОФЕСІЙНИЙ РИЗИК, СТАТИСТИЧНІ ОЦІНКИ, ОБМЕЖЕННЯ, АНАЛІЗ, НЕЩАСНИЙ ВИПАДОК

**O. I. Bogatov, G. A. Pliekhova, M. V. Kostikova, S. M. Neronov, D. O. Pliekhov. Identifying limitations in the use of statistical estimates of occupational risk based on official data in Ukraine.** The work is devoted to the analysis of possible conditions and areas of application of known methods of occupational risk assessment in the state system of occupational safety management. The relevance of this issue is also demonstrated, and limitations to the application of statistical assessments of occupational risk in the management of hazardous facilities are identified. Risk-based occupational safety management has further developed, which must take into account the limitations in the use of probabilistic and statistical methods for determining occupational risk, which are associated with both specific quantitative indicators used for calculations, and with a different approach to defining economic sectors, which is given in the data of the State Labour Service and the State Statistics Service. Conclusions based on data from the State Labour Service or the Pension Fund can be applied to the indicators of the State Statistics Service, with certain limitations.

OCCUPATIONAL RISK, STATISTICAL ESTIMATES, LIMITATIONS, ANALYSIS, ACCIDENT

### Вступ

Сучасний підхід до організації наглядової діяльності передбачає створення відповідної системи управління, методологічною основою функціонування якої повинна стати оцінка і управління професійними ризиками. Актуальність обраної проблеми зумовлена тим, що забезпечення безпеки населення від різних техногенних джерел у розвинених країнах уже кілька десятиліть здійснюється на основі концепції прийнятного ризику, що вимагає кількісного визначення ризику і порівняння його з прийнятним рівнем.

В останні роки роль і значення управління ризиками як інструменту зниження втрат і підвищення ефективності національних економік у всьому світі постійно зростають. Причина цього в стійкому зростанні різноманіття і масштабів прояву ризику і пов'язаних з цим проблем, підвищення в процесі соціально-економічного розвитку чутливості людини та створених ним організацій вже не до масових, а рідкісним негативним явищам.

Доцільність отримання реальних оцінок ризику в нашій країні регламентується нормативними

документами, серед яких виділяються Розпорядження Кабінету Міністрів України № 37 від 2014 року [1], яке привело Концепцію управління ризиками, та Постанова Кабінету Міністрів України № 1043 від 2017 року [2], в якій у відповідності до Концепції наводяться Критерії, за якими оцінюється ступінь ризику. Проте у 2018 році ця Постанова була скасована. І в Постанові № 715 від 2018 року [3], яка прийшла на заміну, вже кількісні показники ризику відсутні.

Все це підкреслює актуальність отримання реальних оцінок ризику для забезпечення ефективного управління охороною праці, техногенною та промисловою безпекою.

### 1. Контроль ризику – нове законодавче завдання

Нормативно-правова база управління ризиками визначається Кодексом цивільного захисту України [4], розпорядженнями Кабінету Міністрів України від 22 січня 2014 р. № 37-р «Про схвалення Концепції управління ризиками виникнення надзвичайних ситуацій техногенного та природного характеру» [1] та від 25 березня 2015 р. № 419-р «Про затвердження плану заходів щодо реалізації Концепції управління

ризиками виникнення надзвичайних ситуацій техногенного та природного характеру на 2015–2020 роки» [5], а також Постановою Кабінету Міністрів України від 5 вересня 2018 р. № 715 «Про затвердження критеріїв, за якими оцінюється ступінь ризику від провадження господарської діяльності та визначається періодичність здійснення планових заходів державного нагляду (контролю) у сфері техногенної та пожежної безпеки Державною службою з надзвичайних ситуацій» [3].

Аналіз цих документів показує, що стосовно кількісних показників визначення професійного ризику в Концепції управління ризиками виникнення надзвичайних ситуацій техногенного та природного характеру встановлені нормативні рівні ризиків: мінімальний –  $1 \cdot 10^{-8}$ , якщо ризик менше цього значення, то він вважається абсолютно прийнятним, а також гранично допустимий –  $1 \cdot 10^{-5}$ , якщо ризик дорівнює або є більшим цього значення, то він вважається неприйнятним. У відповідності до цього для кожної галузі економіки, окремої території України повинні бути визначені свої нормативи рівнів ризику. Відповідно до Концепції управління ризиками в Постанові Кабінету Міністрів України № 1043 від 27 грудня 2017 року [2] були наведені критерії, за якими оцінюється ступінь ризику. У тому разі визначено, що високим рівнем ризику вважаються показники  $5 \cdot 10^{-6} \div 10^{-5}$ , у відповідності до яких наглядові перевірки здійснюються не частіше 1 разу на 2 роки, середнім –  $10^{-6} \div 5 \cdot 10^{-6}$  (перевірки здійснюються не частіше 1 разу на 4 роки), та низьким – менше  $10^{-6}$  (перевірки здійснюються не частіше 1 разу на 6 років).

Проте ця Постанова була скасована. І в Постанові Кабінету Міністрів України № 715 від 5 вересня 2018 року [3], яка прийшла на заміну, вже кількісні показники ризику відсутні. Все це підкреслює актуальність отримання реальних оцінок ризику для забезпечення ефективного управління охороною праці, техногенною та промисловою безпекою.

Таким чином, аналіз можливих умов і галузей застосування відомих методів оцінки професійних ризиків в державній системі управління охороною праці є актуальною науковою задачею, яка передбачає, у тому разі, й визначення обмежень при застосуванні статистичних оцінок професійного ризику в процесі управління станом небезпечних об'єктів.

## 2. Обґрунтування обмежень у використанні статистичних оцінок професійного ризику

Статистичний метод забезпечує найбільшу точність оцінки (за наявності достатньої статистики). У відповідності зі статистичним методом показник професійного ризику отримання працівником травми в результаті нещасного випадку на виробництві можна оцінити за формулою

$$v_{\text{нв}} = N_{\text{нв}} / N, [1 / (\text{чол} \cdot \text{рік})] \quad (1)$$

або

$$K_{\text{нв}} = 1000 \cdot v_{\text{нв}}, [1 / (1000 \text{ чол} \cdot \text{рік})]$$

де  $N_{\text{нв}}$  – число нещасних випадків на виробництві на рік на підприємстві (в галузі, на території);  $N$  – чисельність персоналу, що піддається оцінці ризику за рік, що розглядається.

У якості вихідних даних для оцінки професійного ризику можна використовувати статистичні дані, що наведені в державних документах. Але аналіз цих документів [6–8] показав, що дані відрізняються залежно від організації, яка їх наводить. Приклади цього добре видно на рис. 1 та рис. 2, аналіз яких показує, що є суттєві відмінності даних, наведених Пенсійним фондом, Держпрацею та Державною службою статистики.

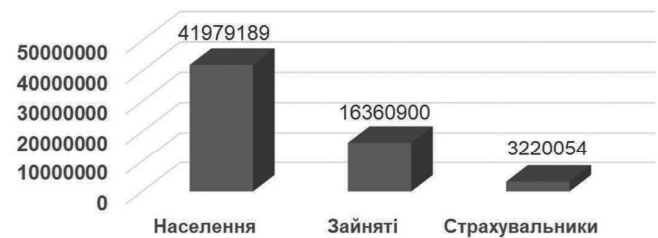


Рис. 1. Кількість населення (2020 рік), до якого розраховуються показники професійного ризику

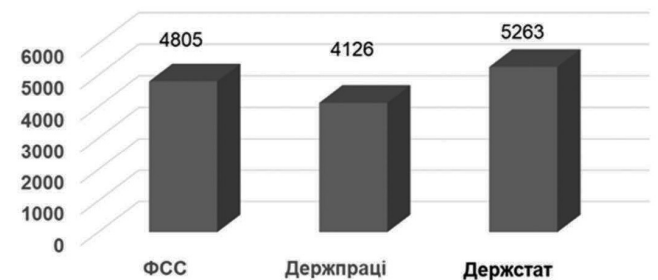


Рис. 2. Виробничий травматизм у 2020 році

Також слід відмітити, що за даними, які наведені на сайті Держпраці, відсутня можливість аналізу за попередні роки (після реорганізації попередніх служб перестали мати місце дані, що наводились на сайтах Пенсійного фонду від нещасних випадків та професійних захворювань, а також Держгірпромнагляду). Дані, що наводяться на сайті Держстату, є однотипними тільки за роки, коли не було війни та коронавірусу.

Виходячи з цього, індивідуальна ймовірність постраждати від нещасного випадку буде приблизно дорівнювати (за даними Держпраці)

$$v_{\text{нв}}(\Delta t) = \frac{n_{\text{нв}}}{N_p} = \frac{4126 \text{ нв} / \text{рік}}{16360900 \text{ чол.}} = 2,522 \cdot 10^{-4} \frac{\text{нв}}{\text{чол} \cdot \text{рік}}, \quad (2)$$

а величина професійного ризику загибелі на виробництві в Україні згідно рис. 3

$$v_{\text{см}}(\Delta t) = \frac{n_{\text{см}}}{N_p} = \frac{409 \text{ см} / \text{рік}}{16360900 \text{ чол.}} = 2,5 \cdot 10^{-5} \frac{\text{см}}{\text{чол} \cdot \text{рік}}. \quad (3)$$

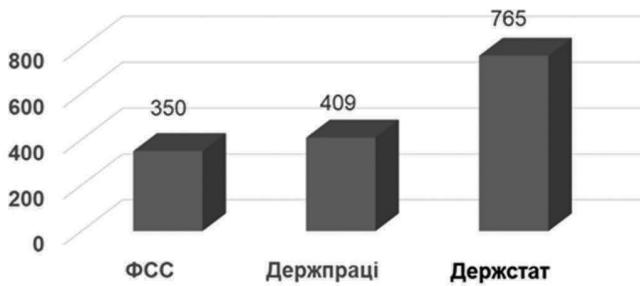


Рис. 3. Виробничий травматизм із смертельними наслідками у 2020 році

Аналіз отриманих у відповідності до (2) та (3) оцінок індивідуальної ймовірності постраждати на виробництві та професійного ризик показує (рис. 4, рис. 5), що рівень професійного ризику у відповідності до вимог Концепції управління ризиками [1] є абсолютно неприйнятним, навіть якщо у якості бази для розрахунку взяти не кількість працюючих або застрахованих, а кількість всього населення України.

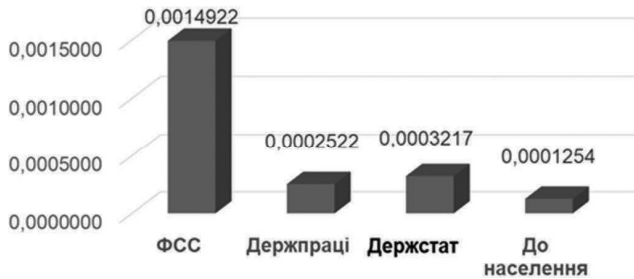


Рис. 4. Ризик виробничого травматизму у 2020 році

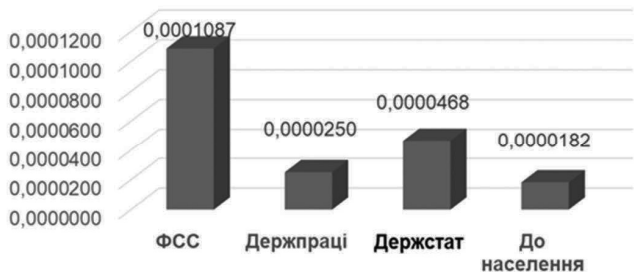


Рис. 5. Ризик виробничого травматизму із смертельними наслідками (професійний ризик) у 2020 році

Крім цього, більше ніж вдвічі відрізняється (рис. 6) доля ризику виробничого травматизму із смертельними наслідками, яка була отримана за даними Держпраці, у порівнянні з відповідною оцінкою Держстатистики. Це свідчить про великий рівень прихованого травматизму, особливо із смертельними наслідками.

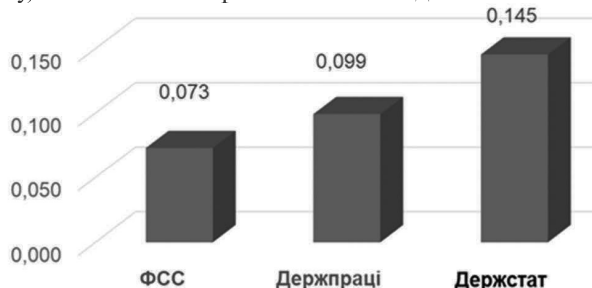


Рис. 6. Доля ризику виробничого травматизму із смертельними наслідками за даними 2020 року

Визначення того, наскільки статистичними оцінками професійного ризику можна користуватися для ризикоорієнтованого управління охороною праці на рівні окремих галузей економіки та підприємств, потребує оцінки мінімального обсягу спостережень.

У припущенні біноміального розподілу числа постраждалих із загального числа, яке підлягало оцінці професійного ризику, статистична невизначеність оцінок (2) та (3) характеризується відносною похибкою

$$\delta_Q(\Delta t) = \frac{Z_\gamma}{\sum v N_p}, \quad (4)$$

де  $Z_\gamma$  – квантиль нормального розподілу рівня  $\gamma$ .

Очевидно, чим менше оцінювана частота події  $v_i$ , наявний обсяг спостережень  $N_p$ , тим більше буде статистична похибка.

Оцінимо, яким повинен бути обсяг спостережень, при якому можлива практична реалізація статистичного методу оцінки професійного ризику отримання виробничої травми працівника на виробництві з похибкою  $\delta_{v_{\text{треб}}} = 10\%$ . Прийнемо довірчу ймовірність  $\gamma = 0,9$ . Їй відповідає значення  $Z_\gamma = 1,282$  [9]. Отримаємо з (4) співвідношення для обсягу спостережень, необхідного для виконання зазначеного обмеження

$$N_p \geq \frac{Z_\gamma^2}{v \delta_{v_{\text{треб}}}^2}. \quad (5)$$

Для умов наведеного прикладу мінімальний обсяг спостережень для оцінки професійного ризику виникнення нещасного випадку на виробництві повинен бути близько

$$N_p \geq \frac{Z_\gamma^2}{v_{\text{нв}} \delta_{v_{\text{треб}}}^2} = \frac{1,282^2}{2,522 \cdot 10^{-4} \cdot 0,1^2} \approx 652 \cdot 10^3, \quad (6)$$

а для оцінки професійного ризику виникнення нещасного випадку зі смертельними наслідками

$$N_p \geq \frac{Z_\gamma^2}{v_{\text{см}} \delta_{v_{\text{треб}}}^2} = \frac{1,282^2}{2,5 \cdot 10^{-5} \cdot 0,1^2} \approx 657 \cdot 10^4. \quad (7)$$

Отримані результати (табл. 1) показали (рис. 7, рис. 8), що для тієї кількості нещасних випадків на виробництві, у тому разі із смертельними наслідками, яка має місце в Україні, показники професійних ризиків, які застосовуються, задовольняються як для нашої держави в цілому, так і для окремих великих галузей економіки.

В той же час, професійні ризики для вибору управлінських рішень в галузі організації харчування та готельного бізнесу не можна використовувати навіть при виборі довірчої ймовірності  $\gamma$  на рівні 0,8, який дозволяється використовувати під час пошукових досліджень.

Таблиця 1

Вихідні дані	Мінімальний обсяг спостережень			
	N		N (см. насл.)	
	$\gamma = 0,9$	$\gamma = 0,8$	$\gamma = 0,9$	$\gamma = 0,8$
ФСС	110140	27535	1512067	378017
Держпраці	651709	162927	6574458	1643614
Держстат	1310922	327730	9018798	2254699

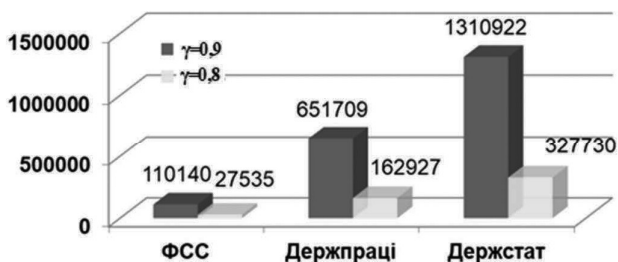


Рис. 7. Мінімальний обсяг спостережень (індивідуальний ризик постраждати на виробництві)

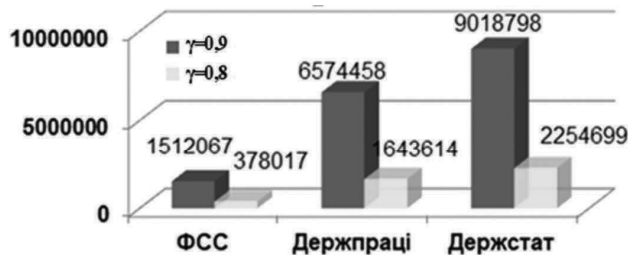


Рис. 8. Мінімальний обсяг спостережень (професійний ризик із смертельними наслідками)

Порівняльний аналіз професійного ризику із смертельними наслідками великих галузей економіки показує (рис. 9–12), що найбільш небезпечна галузь – транспорт та складське господарство. На рисунках

красним кольором показані показники професійного ризику, який характеризує нещасні випадки, що пов'язані з виробництвом, фіолетовим – не пов'язані, а синім – загальний рівень професійного ризику.

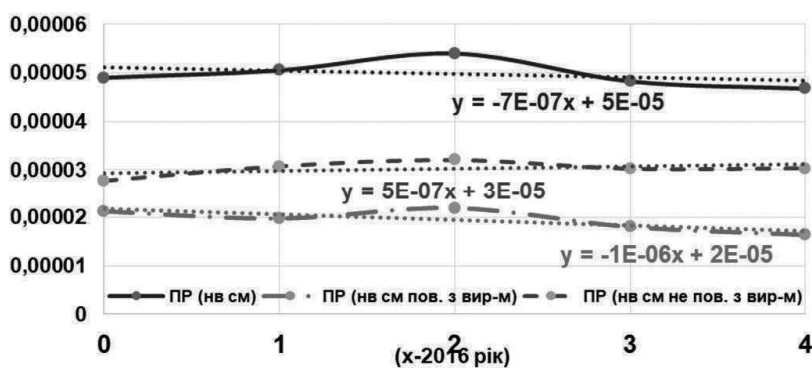


Рис. 9. Професійний ризик (НВ із смертельними наслідками, Держстат)

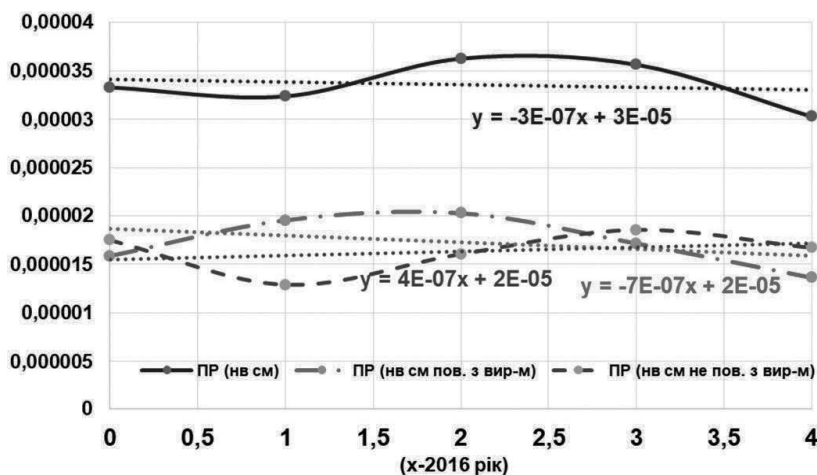


Рис. 10. Професійний ризик (НВ із смертельними наслідками) у сільському, лісовому та рибному господарстві

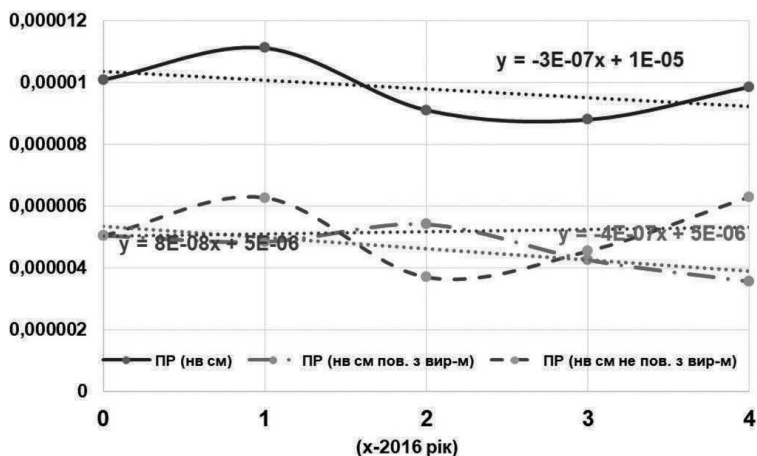


Рис. 11. Професійний ризик (НВ із смертельними наслідками) у промисловості

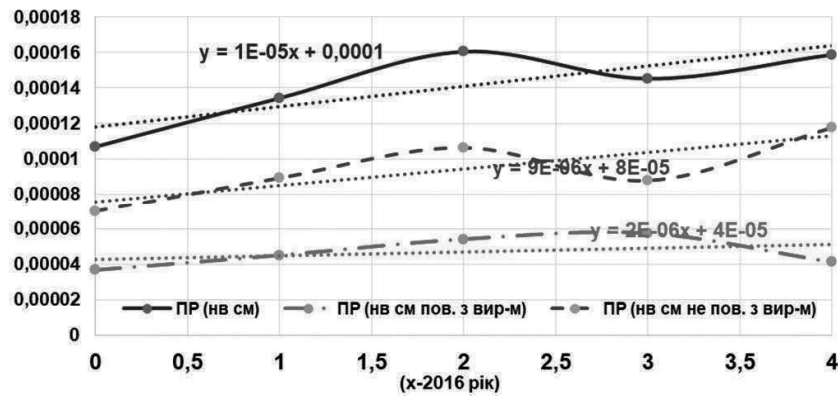


Рис. 12. Професійний ризик (НВ із смертельними наслідками) на транспорті та у складському господарстві

Крім того, видно, що саме в промисловості, де майже всі працівники застраховані, можна очікувати, що саме там найбільш правдивим є співвідношення між професійним ризиком, який пов'язаний з виробництвом, та професійним ризиком, який не пов'язаний з виробництвом. Тобто, в інших галузях штучно не пов'язують нещасні випадки із смертельними наслідками з виробництвом.

### Висновки

Таким чином, набуло подальшого розвитку ризик-орієнтоване управління охороною праці, яке повинно враховувати обмеження у використанні ймовірнісно-статистичних методів визначення професійного ризику, що пов'язані як із конкретними кількісними показниками, які використовуються для розрахунків, так і з різним підходом визначення галузей економіки, який наводиться в даних Держпраці та Держстатистики. Висновки за даними Держпраці або Фонду соціального страхування можна застосовувати до показників Держстатистики з певними обмеженнями.

### Список літератури

[1] Про схвалення Концепції управління ризиками виникнення надзвичайних ситуацій техногенного та природного характеру. Розпорядження Кабінету Міністрів України від 22 січня 2014 р. № 37-р. URL: <https://zakon.rada.gov.ua/laws/show/37-2014-p#Text>.

[2] Про затвердження критеріїв, за якими оцінюється ступінь ризику від провадження господарської діяльності та визначається періодичність здійснення планових заходів державного нагляду (контролю) у сфері техногенної та пожежної безпеки Державною службою з надзвичайних

ситуацій/ Постанова Кабінету Міністрів України від 27 грудня 2017 р. № 1043. URL: <https://zakon.rada.gov.ua/laws/show/1043-2017-p#Text>.

[3] Про затвердження критеріїв, за якими оцінюється ступінь ризику від провадження господарської діяльності та визначається періодичність здійснення планових заходів державного нагляду (контролю) у сфері техногенної та пожежної безпеки Державною службою з надзвичайних ситуацій. Постанова Кабінету Міністрів України від 5 вересня 2018 р. № 715. URL: <https://zakon.rada.gov.ua/laws/show/715-2018-p#Text>.

[4] Кодекс цивільного захисту України: Закон України від 30.11.2012 № 89. URL: <https://zakon.rada.gov.ua/laws/show/5403-17#Text>.

[5] Про затвердження плану заходів щодо реалізації Концепції управління ризиками виникнення надзвичайних ситуацій техногенного та природного характеру на 2015–2020 роки. Розпорядження Кабінету Міністрів України від 25 березня 2015 р. № 419-р. URL: <https://zakon.rada.gov.ua/laws/show/419-2015-p#Text>.

[6] Державна служба статистики України. URL: <https://stat.gov.ua/uk>.

[7] Державна служба України з питань праці. URL: <https://dsp.gov.ua/>.

[8] Пенсійний фонд України. URL: <https://www.pfu.gov.ua>.

[9] Bonchkovskiy, O., Ostapenko, P., Bonchkovskiy, A., & Shvaiko, V. (2025). War-induced soil disturbances in north-eastern Ukraine (Kharkiv region): Physical disturbances, soil contamination and land use change. *Science of The Total Environment*, 964, 178594. <https://doi.org/10.1016/j.scitotenv.2025.178594>.

Received (Надійшла) 18.02.2026

Accepted for publication (Прийнята до друку) 08.03.2026

Publication date (Дата публікації) 27.03.2026

УДК 004.89

DOI: [https://doi.org/10.30837/bi.2026.1\(104\).14](https://doi.org/10.30837/bi.2026.1(104).14)М. С. Широков<sup>1</sup>, В. І. Олевський<sup>2</sup><sup>1</sup>ДНУ, м. Дніпро, Україна, nekit.com37@gmail.com, ORCID iD: 0009-0000-1612-0056<sup>2</sup>ДНУ, м. Дніпро, Україна, olevskiy\_v@365.dnu.edu.ua, ORCID iD: 0000-0003-3824-1013

## РОЗПІЗНАВАННЯ ПІДЗЕМНИХ ПОРОЖНИН У ҐРУНТІ: КЛАСИЧНІ МЕТОДИ ТА СУЧАСНІ ТЕХНОЛОГІЇ

У роботі розглянуто комплексний підхід до виявлення та характеристики підповерхневих порожнин на основі поєднання геофізичних методів та сучасних методів штучного інтелекту. Основну увагу приділено аналізу можливостей електророзвідувальних, сейсмічних та георадіолокаційних методів для формування інформативних вхідних даних, а також застосуванню згорткових і гібридних нейронних мереж для автоматизованого розпізнавання аномалій, пов'язаних з підземними порожнинами. Запропоновано порівняльний аналіз класичних алгоритмів обробки геофізичних даних і моделей глибокого навчання за показниками точності, стійкості до шумів та здатності до узагальнення на різних типах середовищ. Показано, що інтеграція фізично обґрунтованих чисельних моделей із нейромережевими підходами дозволяє підвищити достовірність локалізації та класифікації порожнин, особливо в умовах складної та урбанізованої геологічної будови. Отримані результати підтверджують ефективність використання гібридних геофізично-інтелектуальних підходів для задач інженерної геології, моніторингу техногенних ризиків і безпеки інфраструктури та можуть бути використані як методологічна основа для створення автоматизованих систем підтримки прийняття рішень.

ПІДЗЕМНІ ПОРОЖНИНИ, ГЕОРАДІОЛОКАЦІЯ, ЕЛЕКТРОРОЗВІДКА, СЕЙСМІЧНІ МЕТОДИ, ГЛИБОКЕ НАВЧАННЯ, НЕЙРОННІ МЕРЕЖІ, РОЗПІЗНАВАННЯ АНОМАЛІЙ, ЧИСЕЛЬНЕ МОДЕЛЮВАННЯ

**M.S. Shyrokov, V.I. Oleviskyi. Recognition of underground volatile space: classic methods and modern technologies.**

The paper considers a comprehensive approach to the detection and characterisation of subsurface cavities, combining geophysical methods with modern artificial intelligence. The main attention is paid to the analysis of the capabilities of electrical reconnaissance, seismic, and georadar methods for generating informative input data, as well as to the use of convolutional and hybrid neural networks for automated recognition of anomalies associated with underground cavities. A comparative analysis of classical geophysical data processing algorithms and deep learning models in terms of accuracy, noise resistance, and generalisation ability across different environments is proposed. It is shown that integrating physically based numerical models with neural network approaches increases the reliability of cavity localisation and classification, especially in complex, urbanised geological structures. The results obtained confirm the effectiveness of hybrid geophysical-intellectual approaches for engineering geology tasks, monitoring of technogenic risks, and infrastructure safety, and can serve as a methodological basis for creating automated decision-support systems.

UNDERGROUND VOLTAICS, GEORADIOLOCATION, ELECTRICAL PROSPECTING, SEISMIC METHODS, DEEP LEARNING, NEURAL NETWORKS, ANOMALIES RECOGNITION, NUMERICAL MODELING

### Вступ

Задача розпізнавання підземних порожнин у ґрунті є однією з класичних проблем інженерної геофізики, геотехніки та прикладної механіки суцільних середовищ. Протягом понад двох століть вона розвивалася переважно в межах аналітичних і напіваналітичних підходів, заснованих на розв'язанні хвильових рівнянь для спрощених моделей середовища. Такі методи дозволили закласти фундаментальні уявлення про процеси поширення хвиль, їх відбиття, заломлення та інтерференцію в підповерхневих структурах.

Водночас сучасні умови застосування, зокрема щільна міська забудова, наявність складних техногенних об'єктів, неоднорідність ґрунтів, а також специфіка військових дій, суттєво обмежують можливості класичних підходів. Аналітичні методи виявляються непридатними через складність геометрії, невизначеність фізичних параметрів середовища та неможливість точного задання граничних умов. У таких умовах зростає роль чисельного моделювання та інтелектуальних методів аналізу даних, здатних працювати з зашумленими, неповними та нелінійними сигналами.

Особливого значення набуває поєднання фізично обґрунтованих чисельних моделей поширення хвиль із сучасними методами штучного інтелекту, що дозволяє не лише відтворювати поведінку хвильових полів, а й автоматизувати інтерпретацію результатів. Такий інтегрований підхід є перспективним напрямом розвитку систем розпізнавання підземних порожнин і формує наукове підґрунтя для подальших досліджень у цій галузі.

### 1. Актуальність проблеми

Сучасний етап розвитку суспільства характеризується стрімким зростанням ролі підземного простору в різних сферах людської діяльності — від забезпечення національної безпеки та функціонування критичної інфраструктури до розвитку міського середовища й інженерних комунікацій. Водночас інтенсифікація використання підповерхневих зон супроводжується накопиченням прихованих ризиків, пов'язаних із порушенням структурної цілісності ґрунтового середовища та утворенням підземних порожнин різного походження, зокрема карстові порожнини, покинуті шахти та міська інфраструктура, яку не використовують протягом тривалого часу [1].

У наукових дослідженнях останніх років дедалі частіше наголошується, що підземні порожнини, незалежно від їх походження, становлять суттєву загрозу безпеці населення, стійкості інфраструктури та ефективності систем управління ризиками [2, 3]. Особливу складність створює той факт, що більшість таких об'єктів має прихований характер і не проявляється на поверхні до моменту виникнення аварійних або кризових ситуацій.

Розглянемо більш детально основні види підповерхневих небезпек.

### *1.1. Підземні антропогенні тунелі*

У наукових та аналітичних джерелах останніх років дедалі частіше підкреслюється, що підземні тунелі стали повноцінним елементом сучасних асиметричних конфліктів і гібридних воєн. Дослідники відзначають, що використання підземного простору з військовою та диверсійною метою почало суттєво зростати в середині—кінці ХХ століття, а особливо — у 2010—2020-х роках. Це пов'язують із розвитком міських бойових дій, ущільненням забудови та прагненням сторін конфлікту мінімізувати власні втрати й підвищити прихованість операцій. Активне використання підземного простору дозволяє сторонам конфлікту компенсувати технологічну нерівність і значною мірою нівелювати переваги супротивника у сфері повітряної розвідки та високоточної зброї, як це було під час конфліктів на Близькому Сході [4]. Перенесення частини бойових дій у підповерхневий простір змінює саму логіку конфлікту, роблячи його менш передбачуваним і складнішим для контролю. Яскравим прикладом цього переходу є використання системи підземних тунелів під час В'єтнамської війни в районах Cu Chi та Iron Triangle [5]. Наведене дослідження підтверджує, що американські війська, озброєні значно кращою технікою й вогневыми засобами, виявилися нездатними ефективно знищити ці підземні мережі, зокрема через складність їх ідентифікації. До того ж, такі тунелі стають фактором довготривалої нестабільності навіть після завершення активної фази конфлікту.

Сучасні підземні тунелі суттєво відрізняються від історичних або тимчасових підкопів [6] — мова йде про складні інженерні системи, які включають житлові, складські та командні приміщення і функціонують протягом тривалого часу. Такі мережі часто інтегруються в щільну міську забудову, що не лише ускладнює їх виявлення, а й підвищує ризики для цивільного населення. Це підтверджують останні дослідження сучасної російсько-української війни [7, 8], де підземні маршрути використовуються як для прихованого пересування сил, так і для обходу оборонних ліній супротивника.

Узагальнюючі оглядові дослідження з прикордонної та внутрішньої безпеки також звертають увагу на те, що більшість існуючих систем спостереження орієнтовані на поверхневий або повітряний контроль

і практично не враховують підповерхневий вимір. Також підкреслюється, що дефіцит достовірної інформації про стан ґрунту і підземних структур створює «сліпу зону» в системах національної безпеки, що особливо актуально для прикордонних регіонів і густозаселених територій.

### *1.2. Урбанізація та інфраструктурні ризики, пов'язані з підповерхневими порожнинами*

У галузі інженерної геології, містобудування та управління ризиками значна кількість оглядових публікацій присвячена проблемі деградації ґрунтової основи та появи порожнин унаслідок інтенсивної урбанізації. Відзначається, що сучасні міста дедалі активніше використовують підземний простір для транспортних систем, комунікацій, сховищ і технічних споруд, що істотно змінює природний стан ґрунтів.

Оглядові дослідження у сфері інженерної геології, геотехніки та урбаністики одноставно вказують на те, що інтенсивна урбанізація істотно змінює природний стан ґрунтового середовища [9]. Окремо підкреслюють, що сучасні міста дедалі активніше освоюють підземний простір, що призводить до накопичення прихованих деформацій та локальних зон нестабільності, які часто розвиваються повільно і залишаються непоміченими до моменту аварійного прояву.

Проблема підповерхневих порожнин має комплексний характер і не може бути зведена до окремих помилок проектування або експлуатації. Антропогенне навантаження на ґрунт, застарілі інженерні мережі та зміни гідрогеологічного режиму формують умови для системного зростання ризиків у міському середовищі [10]. Особлива увага приділяється тому, що багато сучасних міст не мають цілісної стратегії моніторингу підповерхневих процесів.

Окремий акцент в літературі робиться на соціально-економічних наслідках таких явищ [11], де зазначається, що раптові просідання та обвали ґрунту можуть призводити до значних людських і матеріальних втрат, порушення транспортної логістики та функціонування критичної інфраструктури. Автори підкреслюють, що в умовах зростаючої щільності забудови навіть локальні порушення ґрунтової стабільності можуть мати непропорційно великі наслідки.

### *1.3. Техногенні порожнини*

Окремий напрям досліджень присвячений проблемі техногенних порожнин, які виникають унаслідок господарської діяльності людини: гірничих робіт, прокладання комунікацій, несанкціонованих підземних споруд або деградації інженерних мереж. У наукових публікаціях підкреслюється, що такі порожнини часто залишаються невідомими або недостатньо задокументованими, особливо в районах зі складною історією забудови [2, 12].

У роботах з інженерної безпеки підкреслюється, що техногенні порожнини є одним із найменш передбачуваних типів підповерхневих дефектів [13, 14].

Значна кількість таких порожнин утворюється внаслідок історичної господарської діяльності — гірничих розробок, старих комунікацій або несанкціонованих підземних споруд, інформація про які є неповною або втраченою. Це створює додаткові складнощі для сучасного міського планування.

Також наголошується, що головна небезпека техногенних порожнин полягає в їх прихованості та відсутності явних поверхневих індикаторів. Такі порожнини можуть залишатися стабільними протягом тривалого часу, але за певних умов призводити до раптових обвалів із катастрофічними наслідками [15]. У цьому контексті дослідники підкреслюють, що традиційні підходи до інженерного обстеження часто є недостатніми для раннього виявлення таких загроз. У свою чергу, деградація інфраструктури без системного моніторингу може призводити до поступового розшлювання ґрунту і втрати його несучої здатності [16]. Додатково підкреслюється, що відсутність єдиних реєстрів і стандартів обліку підземних порожнин значно ускладнює управління ризиками.

Розпізнавання техногенних порожнин варте уваги науковців, є актуальним і відзначається високим ступенем прихованості та непередбачуваності підповерхневих дефектів, зростаючим рівнем зношеності інженерної інфраструктури та відсутністю системного підходу до обліку та моніторингу підземного простору.

## 2. Класичні методи розпізнавання підземної структури

Пошук, опис і розпізнавання підземних структур є однією з найдавніших прикладних задач у природничих науках. Ще в XIX столітті почали формуватися перші геофізичні методи, які дозволяли досліджувати підповерхневі об'єкти без прямого проникнення в товщу ґрунту. У другій половині XIX — на початку XX ст. електричні та сейсмічні підходи поступово розширили свою сферу застосування для виявлення не лише корисних копалин, а й неоднорідностей у підповерхневих шарах [17].

У XX ст. інженерна та прикладна геофізика стала окремою дисципліною, що об'єднала різноманітні підходи до дослідження внутрішньої структури Землі. До традиційних методів відносяться сейморозвідка, електророзвідка, магніторозвідка, гравірознавдя та георадіолокація (GPR) — технології, які дозволяють непрямо оцінювати наявність неоднорідних зон, включаючи порожнини, по зміні фізичних полів.

Інші геофізичні підходи, наприклад, вимірювання природних або збуджених електромагнітних полів, також застосовувалися для вирішення прикладних задач розвідки структури ґрунту та підземних вод.

### 2.1. Електророзвідка

Електророзвідка є одним із найстаріших геофізичних методів, витоки якого сягають експериментів Р. В. Фокса (1829 р.), пов'язаних із вимірюванням

природних електричних потенціалів у ґрунті. Сутність електророзвідки полягає у тому, що різні типи ґрунтів, порід і порожнин мають різну електропровідність [17]. Під час вимірювання штучно створеного або природного електричного поля фіксуються зміни потенціалів, за якими можна судити про підповерхневі неоднорідності. Порожнини, заповнені повітрям або слабопровідними матеріалами, створюють характерні аномалії електричного поля, що дозволяє виявляти їх за відповідної конфігурації електродів.

У прикладній електророзвідці налічується більше п'ятдесяти різних методів і модифікацій, призначених як для глибинних досліджень, так і для вивчення верхньої частини розрізу. Через багатоманітність способів електророзвідку можна поділити на класи, групи, методи й модифікації. Усі методи електророзвідки можна поділити на три групи: методи постійного струму, поляризаційні (електрохімічні) методи та електророзвідка на змінному струмі.

Основними методами постійного струму є методи опорів, які поділяються на електропрофілювання і електричне зондування: вертикальне і дипольне. У даній групі штучних методів електророзвідки використовуються не тільки постійні поля, але й змінні низькочастотні ( $f < 20$  Гц), для яких при малих розмірах установок справедлива теорія постійних полів.

Суть методів опорів зводиться до визначення позірних опорів шляхом виміру різниці потенціалів на приймальних електродах і сили струму в живильній лінії. Із розв'язку прямих задач для стаціонарних електричних полів різних джерел над однорідним провідним півпростором випливають формули для визначення його питомого опору  $\rho$ . У методах опору за цими ж формулами розраховується позірний опір  $\rho_{\text{п}}$ , що характеризує аномальне поле, зумовлене неоднорідною будовою півпростору:

$$\rho_{\text{п}} = K * \Delta U / I, \quad (1)$$

де  $I$  — величина струму, що пропускається в землю через живильні електроди;  $\Delta U$  — різниця потенціалів, яка зареєстрована між вимірювальними електродами у момент проходження струму;  $K$  — коефіцієнт установки.

Методи поляризації ґрунтуються на вимірюванні вторинних електричних полів електрохімічної та електрокінетичної природи, що виникають у гірських породах при протіканні в них електричного струму. Метод викликаної поляризації є одним із провідних у комплексі геофізичних методів, що використовуються при пошуках і розвідці рудних родовищ.

При вимірюванні викликаної поляризації в часовому (нестационарному) режимі в землю гальванічним способом (за допомогою електродів) подається постійний електричний струм, під дією якого протягом певного часу відбувається вимушена поляризація геологічного розрізу. Під час проходження в землі струму вимірюється різниця потенціалів між приймальними

заземленнями й сила струму в живильній лінії, а після вимкнення струму — різниця потенціалів викликані поляризації  $\Delta U_{ВП}$ . Ця різниця потенціалів вимірюється або в строго фіксований момент часу після вимкнення струму, або через певні його проміжки. Швидкість і характер спаду процесу деполіризації залежать від поляризаційних властивостей геологічних утворень. На практиці інформацію про ці властивості отримують шляхом визначення миттєвого значення позірної поляризованості  $\eta_{П}$  для вибраного моменту часу або шляхом вивчення залежності  $\eta_{П}$  від часу. Для визначення позірної поляризованості після подання в живильну лінію постійного електричного струму впродовж певного часу його вимикають і на встановленій часовій затримці (зазвичай 0,5 с) вимірюють різницю потенціалів  $\Delta U_{ВП}$ .

Позірну поляризованість визначають:

$$\eta_{П} = (\Delta U_{ВП} / \Delta U_{ПР}) \cdot 100\%, \quad (2)$$

де  $\Delta U_{ПР}$  — різниця потенціалів на електродах приймальної лінії, виміряна під час пропускання струму перед його вимкненням. Одночасно, як правило, визначають і позірний опір:

$$\rho_{П} = K \cdot \Delta U_{ПР} / I, \quad (3)$$

де  $K$  — коефіцієнт установки.

Вивчення поляризаційних властивостей геологічного розрізу можна виконувати також і на змінному струмі. У цьому випадку викликану поляризацію оцінюють у градусах фазового зсуву виміряного гармонійного сигналу  $E_{MN} \cdot \sin(\omega t + \varphi_{ВП})$  щодо струмового сигналу  $I_{AB} \cdot \sin(\omega t)$ , визначаючи таким чином затримку в реакції середовища щодо процесу дії зовнішнього поля. Потрібно зазначити, що явище викликані поляризації досить інерційне й на високих частотах (>20 Гц) взагалі практично відсутнє.

Якщо середовище однорідне за поляризованістю, то позірна поляризованість співпадає з її істинним значенням. Для неоднорідного середовища  $\eta_{П}$  залежить від просторового розподілу ділянок порід із різною поляризованістю, від їх питомого електричного опору, від розташування живильних і вимірювальних електродів і характеру поляризувального поля.

Польові дослідження останнього десятиріччя, проведені науковцями за допомогою електророзвідки в різних умовах та ґрунтах [18, 19], показують, що реальна точність розпізнавання, якої вдалося досягнути, — близько 80%. Причиною цьому є застосування спрощених моделей середовища, припущень про шаруватість та однорідність, які погано відповідають реальним умовам щільної забудови або зруйнованої інфраструктури.

## 2.2. Сейсмічні та акустичні методи

Сейсмічна розвідка — сукупність методів дослідження геологічного середовища, що ґрунтуються на вивченні розповсюдження пружних хвиль, які збуджуються штучно. Сейсмічні методи ґрунтуються на аналізі поширення пружних хвиль у ґрунті. Акустичні

методи є близькими за фізикою до сейсмічних, але працюють у вищому частотному діапазоні. Джерело (удар, вибух, вібратор) генерує пружні хвилі, які поширюються в різних середовищах з різною швидкістю [20]. На межі, що розділяє ґрунти різного складу, пружні хвилі відбиваються, заломлюються та частково повертаються на поверхню Землі. Вивчаючи час і швидкість поширення хвиль, їх амплітуду, форму сигналу та характер коливань ґрунту, на яких встановлені вібраційні приймачі (сейсмічні приймачі або геофони), можна визначити глибину та форму меж у середовищі, кут їх падіння, напрямок поширення та інші характеристики геологічного середовища (рис. 1).

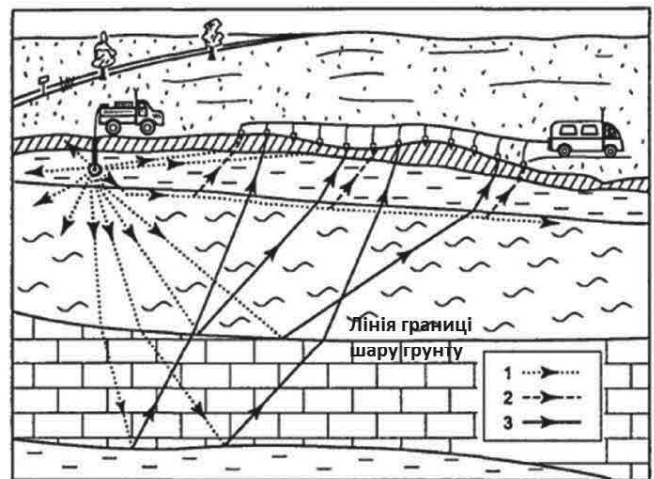


Рис. 1. Схема польових робіт сейсморозвідки

У сейсмічній розвідці використовуються поздовжні хвилі, поперечні та змішані хвилі. Сейсморозвідка включає два основні методи: метод відбитих хвиль (МВХ) і метод заломлених хвиль (МЗХ). Найбільшого поширення набув метод відбитих хвиль, який дозволяє ідентифікувати межі середовищ з точністю понад 95% [20, 21]. Метод заломлених хвиль має нижчу точність і роздільну здатність, що дозволяє вивчати лише шари зі збільшеною швидкістю сейсмічних хвиль.

Хвильове сейсмічне поле, яке виникає при штучному збудженні й реєструється на поверхні Землі, є надзвичайно складним і відображає складну будову і зміну пружних властивостей неоднорідного геологічного середовища. Для дослідження закономірностей розповсюдження пружних хвиль у геологічному середовищі застосовують спрощені моделі. У першому наближенні реальне геологічне середовище розглядають як деяке ідеально пружне ізотропне середовище, в якому під дією об'ємних або поверхневих сил виникають і розповсюджуються пружні хвилі.

В ідеально пружному середовищі процеси, які відбуваються, є повністю зворотними. Це означає, що ніяких незворотних втрат енергії пружної хвилі в ідеально пружному середовищі немає. В моделі ідеально пружного тіла основні закономірності розповсюдження пружних хвиль описуються законами лінійної теорії пружності. Теорія пружності дає можливість знайти

залежність між силами, прикладеними до зовнішньої поверхні тіла, і зміною розмірів і форми тіла, які виникли внаслідок дії цих сил. Цю залежність зручно виразити за допомогою напруг і деформацій, лінійна залежність між якими описується законом Гука, базуючись на якому, можна знайти фундаментальні співвідношення, що визначають умови динамічної рівноваги ідеально пружного середовища.

Рівняння динамічної рівноваги однорідного ідеально пружного ізотропного середовища має такий вигляд (рівняння Ламе):

$$\rho \frac{\partial^2 u}{\partial t^2} = (\lambda + \mu) \nabla \theta + \mu \Delta u + \rho \chi, \quad (4)$$

де  $\rho$  – густина середовища;  $\lambda, \mu$  – пружні постійні Ламе;  $\Delta$  – лапласіан,  $u$  – вектор пружних зміщень, що визначає переміщення частинок середовища відносно їхнього положення рівноваги, який розглядається як функція координат  $x, y, z$  простору та часу  $t$  і характеризує збурення в середовищі внаслідок дії зовнішніх об'ємних сил  $\chi$ ,  $\theta$  – дилатація або відносне об'ємне розширення, яке визначається через компоненти вектора зміщень.

Значна кількість сейсмічних систем працює в умовах, коли інтенсивність корисного сейсмічного сигналу перевищує рівень перешкод через використання вибухових джерел формування пружних хвиль. Однак, при вивченні малих структур, таких як невеликі порожнини в ґрунті, слід використовувати спеціальні методи розпізнавання їх впливу на акустичні хвилі, серед яких є обробка даних з використанням комбінації методів комп'ютерного зору та машинного навчання.

У дослідженнях останніх років [20, 21, 22] сейсмічні та акустичні методи застосовуються переважно для знаходження відповідності реконструйованої аномалії відомій геометрії порожнини або співпадіння локалізації з контрольною інформацією з буріння/архівних даних, де показують непогані результати реконструкції підземних порожнин, але через наведені вище особливості, прямі метрики оцінки якості здебільшого не наводяться. Дані дослідження виступають підтвердженням того, що використання методів аналізу вібрацій і шумів ґрунту із застосуванням сучасних методів машинного навчання є перспективним напрямком досліджень.

### 2.3. Георадіолокація (GPR)

Ще одним широко використовуваним методом для аналізу ґрунтів є георадіолокація. Її фізична сутність ґрунтується електромагнітних полях та законах електродинаміки. За законом Кулона, електричний заряд оточений електричним полем, яке діє на інші заряди, змушує їх пересуватися. Закон Ампера показує, що електричний заряд, що рухається в провіднику, створює магнітне поле, пропорційне швидкості заряду. Якщо електричне поле збільшується, то заряд прискорюється і швидкість його зміни призводить до зміни магнітного поля, що, у свою чергу, індукує інше

електричне поле в провіднику і тим самим впливає на рух прискореного заряду [23].

Якщо два прямі провідники покласти вдовж один одного і з'єднати послідовно, то вони будуть діяти як електричний диполь. Застосування до них змінного електричного поля призведе до коливання диполя, що діятиме як антена для випромінювання електромагнітних хвиль (ЕМ). Такі хвилі складаються з магнітного поля  $H$  та електричного поля  $E$ , які змінюються з частотою осцилятора та орієнтовані під прямим куттом один до одного у площині, перпендикулярній напрямку поширення (рис. 2).

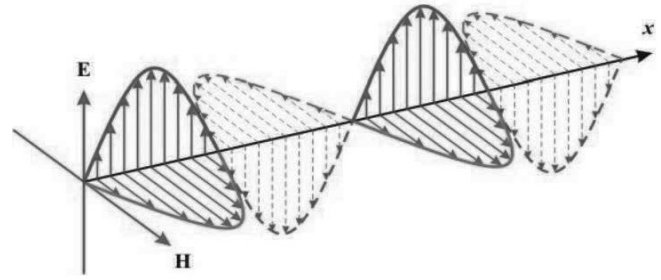


Рис. 2. Поширення електромагнітної хвилі

Поширення електромагнітних хвиль метрового діапазону, характерного для георадарних досліджень, розглядається в рамках законів геометричної оптики відповідно для плоских хвиль на великих відстанях від джерела і для сферичних хвиль на малих відстанях від джерела. Враховуючи контраст діелектричної проникності штучних або природних шарів, їх можна відрізнити один від одного.

При визначенні швидкостей поширення і амплітуд хвиль у рамках допустимих похибок, діють принципи Ферма, Гюйгенса, Френеля і закон Снеліуса [30]. Припускається, що в ізотропному середовищі промінь перпендикулярний фронту хвилі, що дозволяє зробити геометричні побудови шляхів хвиль в середовищі і подальше обчислення параметрів середовища за вимірними кінематичними і динамічними характеристиками записів хвиль. Хвиля, що поширюється, відбивається тільки від границь поділу різних середовищ, в яких спостерігається зміна відносної діелектричної проникності або провідності. Коефіцієнт відбиття при нормальному падінні хвилі на границю поділу між середовищами з різною діелектричною проникністю, буде визначатись за виразом:

$$K_{\text{відб.}} = \left| \frac{\sqrt{\epsilon_1'} - \sqrt{\epsilon_2'}}{\sqrt{\epsilon_1'} + \sqrt{\epsilon_2'}} \right|^2, \quad (5)$$

де  $\epsilon_1', \epsilon_2'$  – дійсні частини відносної проникності двох відповідних середовищ.

Коефіцієнт заломлення (проходження через границю) буде рівний:

$$K_{\text{пр}} = 1 - K_{\text{відб.}} = 1 - \frac{2 \cdot \sqrt{\epsilon_2'}}{\sqrt{\epsilon_1'} + \sqrt{\epsilon_2'}} \quad (6)$$

Також важливим явищем для георадіолокаційних досліджень є дифракція, яка виникає у результаті опромінення об'єкта електромагнітними хвилями,

розміри якого менші за довжину хвилі. Дифракція може вплинути на розпізнавання хвильових картин, так як пошук локальних об'єктів зводиться до пошуку вторинних джерел випромінювання на записах, що впливає на визначення глибини залягання об'єкта та швидкість розповсюдження електромагнітних хвиль в середовищі над об'єктом.

Задачі з пошуку різних об'єктів є одними із основних напрямків георадіолокаційних досліджень. До них відносяться пошук трубопроводів, кабелів, валунів, залишки будівельних конструкцій, ємностей, порожнин, археологічних об'єктів, тощо. Порожнини, особливо заповнені повітрям, створюють різкий контраст діелектричних властивостей, що робить їх добре помітними на радарограмах.

Експерименти, проведені в умовах дорожнього покриття, поблизу великих бетонних споруд (дамби) та у золотих шахтах Єгипту показують точність, близьку до 75-85% за умови низької електропровідності матеріалу заповнення порожнини й однорідності навколишнього ґрунту [24, 25]. У результатах досліджень зазначається, що GPR має істотні обмеження: глибина проникнення сильно залежить від вологості ґрунту, наявності металевих об'єктів і електромагнітних завад, що є типовим для міських умов.

#### 2.4. Обмеження класичних методів

Спільною рисою перелічених методів є їх аналітична природа: інтерпретація результатів ґрунтується на розв'язанні обернених задач, які є погано обумовленими та потребують значних спрощень. У мирних умовах, на відкритих територіях з переважно однорідним ґрунтом це допустимо, оскільки можна збільшити кількість вимірювань і уточнювати модель середовища, проводячи дослідження у декілька етапів. Але в умовах військових дій, обмеженого доступу до території, щільної міської забудови та наявності численних техногенних шумів класичні підходи стають малоприматними, занадто повільними або мають занижену точність. До того ж, аналітичні моделі не здатні повною мірою врахувати складну суперпозицію сигналів, що виникає в реальному середовищі, що є додатковим аргументом для застосування методів, які матимуть більш високу точність.

Наведені вище обмеження класичних методів зумовлюють перехід до використання систем штучного інтелекту. На відміну від класичних методів, використання ШІ не потребують явного задання фізичної моделі середовища і здатні навчатися як на реальних, так і змодельованих даних, враховуючи складні нелінійні залежності між формою акустичного сигналу та структурою ґрунту.

Перспективним є аналіз акустичних і сейсмоакустичних сигналів, оскільки такі сигнали містять інформацію про резонансні властивості порожнин, локальні відбиття та особливості затухання хвиль. Системи штучного інтелекту можуть виділяти ці ознаки навіть у зашумлених даних, що робить їх

ефективними в умовах, де аналітичні методи виявляються непридатними.

Сучасні огляди [25, 26] підкреслюють, що поєднання сенсорних систем із методами машинного навчання є одним із ключових напрямів розвитку підземного зондування.

### 3. Інтелектуальні рішення для розпізнавання підземних порожнин

На практиці існуючі інтелектуальні рішення можна умовно поділити на дві основні групи: підходи на основі традиційних алгоритмів машинного навчання та підходи на основі глибокого навчання і нейронних мереж.

#### 3.1. Підходи на основі традиційного машинного навчання

Алгоритми класичного машинного навчання (Support Vector Machines, Random Forest, k-Nearest Neighbors, Logistic Regression, тощо) широко застосовуються для задач класифікації та виявлення аномалій у геофізичних і акустичних даних. Типовий підхід полягає в тому, що первинний сигнал попередньо обробляється — виконується фільтрація, спектральний аналіз, розрахунок статистичних характеристик або ручне виділення ознак (features), після чого сформований вектор ознак подається на вхід класифікатора [27].

До переваг традиційних ML-підходів можна віднести відносно простоту реалізації та навчання моделей, менші вимоги до обчислювальних ресурсів, можливість часткової інтерпретації результатів та придатність для невеликих навчальних вибірок. В якості недоліків традиційних ML-підходів можна виділити критичну залежність від якості ручного виділення ознак, слабку здатність працювати з високорозмірними сигналами, обмежену ефективність у разі нелінійних і складних взаємозв'язків у даних та значне падіння точності за наявності сильних шумів і варіативних умов вимірювання.

У задачах розпізнавання підземних порожнин ці недоліки стають особливо критичними, оскільки акустичні сигнали мають складну часово-частотну структуру, а їх форма істотно змінюється залежно від типу ґрунту, геометрії порожнини та зовнішніх завад. У таких умовах ручне формування ознак часто призводить до втрати значної частини корисної інформації.

#### 3.2. Підходи на основі глибокого навчання

Глибоке навчання принципово змінює підхід до аналізу підземних сигналів, оскільки нейронні мережі здатні автоматично виділяти інформативні ознаки без явного задання фізичної моделі середовища [28]. Це особливо важливо для задач, де класичні припущення про однорідність або лінійність є неприйнятними.

Серед архітектур глибоких нейронних мереж використовуються: повнозв'язані (fully connected) мережі, рекурентні мережі (RNN, LSTM), згорткові нейронні мережі (CNN) та гібридні архітектури.

Вибір архітектури нейронної мережі є критично важливим етапом при побудові інтелектуальної системи розпізнавання підземних порожнин, оскільки тип архітектури безпосередньо визначає здатність моделі виявляти інформативні закономірності в акустичних сигналах за умов високого рівня шумів, неоднорідності середовища та обмеженості апріорної інформації.

Згорткові нейронні мережі історично створювалися для аналізу зображень, проте їхня архітектура виявилася надзвичайно ефективною і для аналізу акустичних сигналів, поданих у вигляді спектрограм, скалограм або багатовимірних часових матриць [29].

Згорткові нейронні мережі мають низку переваг у задачах розпізнавання підземних порожнин. Вони здатні використовувати фільтри, які виявляють характерні локальні патерни в акустичних сигналах, що відповідають відбиттям, резонансам і зонам ослаблення хвиль, пов'язаних із наявністю порожнин. Завдяки локальним згорткам і багаторівневій ієрархії ознак CNN ігнорують випадкові шуми та концентруються на стабільних структурних елементах сигналу, що є критично важливим в умовах міської забудови та бойових дій. Також згорткові нейронні мережі здатні розпізнавати характерні ознаки незалежно від їх точного положення у часовому просторі сигналу та добре оптимізуються для паралельних обчислень, що відкриває можливість їх використання в мобільних або польових системах для оперативного прийняття рішень.

Наведені переваги підтверджуються сучасними дослідженнями [29, 30] у роботах з використанням CNN для ідентифікації підземних штучних порожнин та інверсії георадіолокаційних даних, де продемонстровано перевищення точності порівняно з традиційними ML-методами.

### 3.3. Поєднання згорткових нейронних мереж з акустичним методом

Акустичний метод є особливо привабливим для задач підземного зондування, оскільки акустичні хвилі легко генеруються, поширюються в різних типах ґрунтів та чутливо реагують на геометрію та механічні властивості порожнин.

Через те, що акустичні сигнали є надзвичайно складними для аналітичної інтерпретації, поєднання акустичного методу з CNN створюватиме синергетичний ефект: сигнали забезпечуватимуть багатий інформаційний вміст, а згорткові мережі — ефективно його вилучення та інтерпретацію. CNN дозволяють аналізувати часові характеристики сигналів та їх представлення, що значно підвищує ймовірність коректного виявлення підземних порожнин навіть за несприятливих умов.

## 4. Використання чисельного моделювання для задачі розпізнавання підземних порожнин

Через складність аналітичного розв'язання хвильових рівнянь у реальних умовах, з урахуванням неоднорідності середовища, зміни фізичних параметрів ґрунту, складних граничних умов та шумових впливів, аналітичні розв'язки стають практично

недосяжними або непридатними для використання в задачах практичного прогнозування поведінки хвильових полів. Тому в сучасних дослідженнях широко застосовується чисельне моделювання, яке дозволяє ефективно отримувати точні прогнози поведінки хвильових полів у складних середовищах та при складних граничних умовах, коли прямі математичні формули виявляються неефективними або неможливими для явного розв'язання.

Чисельне моделювання дає змогу не лише оцінювати розповсюдження хвиль у неоднорідних підповерхневих структурах, але й апроксимувати вплив складних геометрій і змінних властивостей матеріалів. Підхід базується на дискретизації простору та часу і систематичному розрахунку змін поля у кожній точці дискретної сітки.

Найпоширенішими методами чисельного моделювання є метод кінцевих різниць (FDM), метод кінцевих елементів (FEM) та метод кінцевих об'ємів (FVM).

### 4.1. Метод кінцевих різниць

Метод кінцевих різниць є одним із найдавніших і широко застосовуваних чисельних методів для моделей хвильових рівнянь. Він замінює похідні у диференціальних рівняннях їхніми апроксимованими різницевидами аналогами. Для простих задач FDM забезпечує відносно просту реалізацію та ефективну реалізацію на структурованих сітках.

Приклади його використання є в задачах обчислення сейсмічних хвиль з інверсією даних для реконструкції підповерхневих структур та роботі з сейсмічними хвилями у середовищах із аномаліями [31]. Дані дослідження також відмічають відносно просту реалізацію та низькі обчислювальні вимоги цього методу.

За допомогою метода кінцевих різниць можна сформулювати реалістичні хвильові сигнали із пустотами та урахуваннями підповерхневих аномалій, має він має труднощі з моделями складної геометрії та помітно високу чутливість до неоднорідностей та дисперсії, що не так відчутно в методах кінцевих елементів та кінцевих об'ємів. В свою чергу, дані недоліки можуть мати суттєвий негативний вплив під час навчання моделей нейронних мереж, оскільки вона буде вчитися на свідомо не точних даних, через що кінцева точність може бути недостатньо високою.

### 4.2. Метод кінцевих елементів

Метод кінцевих елементів розбиває обчислювану область на елементарні підобласті (елементи), де рішення апроксимується наближеними функціями. Він широко застосовується для розрахунків механічних і хвильових проблем, особливо у складних геометріях, оскільки легко адаптується до нерівномірних сіток та складних граничних умов.

Прикладами використання методу кінцевих елементів є задачі чисельного моделювання поширення сейсмічних хвиль із подальшою інверсією даних для реконструкції підповерхневих структур, а також дослідження поширення хвиль у середовищах зі

складними аномаліями та порожнинами різної геометрії [32]. У наведених роботах зазначається, що метод кінцевих елементів дозволяє точно враховувати складну геометрію середовища, нерівні межі розділу та неоднорідні матеріальні властивості, що є важливим для задач моделювання реальних ґрунтових умов.

Разом із цим, у наведених вище публікаціях відзначається, що попри високу точність і гнучкість геометричного опису, метод кінцевих елементів характеризується значно вищими обчислювальними витратами у порівнянні з методами кінцевих різниць та кінцевих об'ємів, а також підвищеними вимогами до генерації якісної сітки та стабільності чисельного розв'язання. Це може створювати додаткові труднощі при формуванні великих масивів синтетичних даних для навчання нейронних мереж, через що обмежується масштабованість та швидкість створення навчальних вибірок. Крім того, висока складність реалізації підвищує ризик накопичення чисельних похибок на етапі попереднього моделювання, що може призвести до формування частково спотворених навчальних даних і мати негативний вплив на кінцеву точність моделей.

#### 4.3. Метод кінцевих об'ємів

Як і попередні методи, метод кінцевих об'ємів широко застосовується в задачах чисельного моделювання хвильових процесів у неоднорідних середовищах для розрахунку поширення акустичних та сейсмічних хвиль з метою відтворення підповерхневої структури [33]. Приклади його використання наведені в роботах, присвячених моделюванню хвильових полів у середовищах зі складною геометрією та різко вираженими фізичними контрастами, такими як порожнини, тріщини та техногенні аномалії [34, 35]. У цих дослідженнях підкреслюється, що метод кінцевих об'ємів забезпечує строге збереження фізичних законів (мас, імпульсу, енергії) на дискретному рівні, що є важливим фактором для формування фізично коректних хвильових сигналів.

На відміну від методу кінцевих різниць, метод кінцевих об'ємів значно краще працює зі складною геометрією області, що робить його придатним для моделювання реальних ґрунтових середовищ із підземними порожнинами різної форми [33]. Ця властивість є важливою для задач навчання нейронних мереж, оскільки модель буде отримувати фізично узгоджені репрезентативні навчальні вибірки, що матиме позитивний вплив на підвищення кінцевої точності розпізнавання.

Також у дослідженнях [34, 35] відзначається, що метод кінцевих об'ємів демонструє вищу чисельну стійкість у порівнянні з методом кінцевих різниць, особливо при моделюванні хвиль у середовищах із високими контрастами густини. Цей аспект є важливим, оскільки він показуватиме більш стабільні результати в місцях, де присутні різні типи ґрунтів.

Один з основних недоліків методу кінцевих об'ємів — це дещо вищі обчислювальні витрати у

порівнянні з методом кінцевих різниць. Але сучасні реалізації та можливість паралельних обчислень нівелюють цей недолік, що підтверджується практичними застосуваннями [35]. До того ж, метод кінцевих об'ємів все одно менш ресурсомісткий у порівнянні з методом кінцевих елементів.

Виходячи з наведеного вище, метод кінцевих об'ємів виглядає оптимальним рішенням, оскільки він поєднує високу точність, коректність та гнучкість щодо складної геометрії, що робить його використання доцільним для генерації навчальних даних у задачах розпізнавання підземних порожнин. Сформовані за його допомогою хвильові поля створюють надійну відправну точку для побудови стійких і високоточних систем аналізу підповерхневих структур.

### Висновки

У результаті проведеного аналізу сучасного стану проблеми розпізнавання підземних порожнин можна зробити такі узагальнені висновки.

По-перше, традиційні аналітичні методи, попри їх фундаментальну значущість, мають обмежену практичну застосовність у реальних умовах через складність геометрії підповерхневих структур, неоднорідність ґрунтових середовищ і неможливість точного задання параметрів моделі. Це зумовлює необхідність переходу до чисельних методів, які забезпечують гнучкість і універсальність при моделюванні хвильових процесів.

По-друге, чисельне моделювання є ключовим інструментом для дослідження поведінки акустичних і сейсмічних хвиль у складних середовищах. Серед розглянутих чисельних підходів метод кінцевих об'ємів вирізняється як найбільш придатний для задач хвильової динаміки завдяки своїй консервативності, стійкості до різких змін параметрів середовища, здатності працювати зі складними геометріями та можливості розширення на мультифізичні моделі.

По-третє, аналіз інтелектуальних методів обробки даних показав, що традиційні алгоритми машинного навчання є обмеженими через залежність від ручного формування ознак і недостатню здатність працювати з високорозмірними сирими сигналами. Натомість методи глибокого навчання, зокрема згорткові нейронні мережі, забезпечують автоматичне виділення інформативних ознак, високу стійкість до шумів і кращу узагальнювальну здатність.

Поєднання акустичного методу зондування зі згортковими нейронними мережами створює ефективний інструментарій для розпізнавання підземних порожнин у складних умовах, де класичні методи є непридатними. Такий підхід дозволяє інтегрувати фізичну природу хвильових процесів із адаптивними можливостями штучного інтелекту, що відкриває перспективи створення точних, стійких і масштабованих систем підповерхневого моніторингу.

Отримані висновки формують науково-методологічну основу для подальших досліджень, спрямованих на розробку комплексних систем розпізнавання

підземних порожнин із використанням чисельного моделювання та нейронних мереж, а також підтверджують доцільність обраного напрямку дослідження.

#### Список літератури:

- [1] Rahimi, M., Wood, C. M., & Kallivokas, L. F. (2024). A comparative study of using geophysical methods for imaging subsurface voids of various sizes and at different depths. *Engineering Geology*, *341*, 107711. <https://doi.org/10.1016/j.enggeo.2024.107711>
- [2] Zhou, J., Feng, W., Guan, P., Liu, J., Zhang, H., & Wang, Z. (2025). Comprehensive detection of groundwater-affected ancient underground voids during old town renewal: A case study from Wuhan, China. *Water*, *17*(23), 3356. <https://doi.org/10.3390/w17233356>
- [3] Mevoli, F. A., Fazio, N. L., Perrotti, M., et al. (2024). Assessing the stability of underground caves through iSUMM (innovative, straightforward, user-friendly, mechanically-based method). *Geoenvironmental Disasters*, *11*, 10. <https://doi.org/10.1186/s40677-023-00264-3>
- [4] Steenkamp, C. (2022). The impact of tunnels on conflicts in the Middle East. *International Affairs*, *98*(2), 689–706. <https://academic.oup.com/ia/article/98/2/689/6530475>
- [5] Olson, K., & Morton, L. (2017). Why were the soil tunnels of Cu Chi and Iron Triangle in Vietnam so resilient? *Open Journal of Soil Science*, *7*, 34–51. <https://doi.org/10.4236/ojss.2017.72003>
- [6] Abuamer, M. (2024). Gaza's subterranean warfare: Palestinian resistance tunnels vs. Israel's military strategy. *Terrorism and Political Violence*. <https://doi.org/10.1080/1057610X.2024.2347843>
- [7] Dorsel, D. B. (n.d.). *Use of subterranean infrastructure in contemporary conflicts*. <https://oe.tradoc.army.mil/product/use-of-subterranean-infrastructure-in-contemporary-conflicts/>
- [8] Richemond-Barak, D. (n.d.). *From Gaza to Ukraine: Three principles of underground warfare*. <https://www.miryaminstitute.org/commentary-blog/from-gaza-to-ukraine-three-principles-of-underground-warfare>
- [9] Huning, L. S., Love, C. A., Anjileli, H., Vahedifard, F., Zhao, Y., Chaffe, P., Cooper, K., Alborzi, A., Pleitez, E., & Martinez, A. (2024). Global land subsidence: Impact of climate extremes and human activities. *Reviews of Geophysics*. <https://doi.org/10.1029/2023RG000817>
- [10] Ikuemonisan, F. E. (2025). An integrative review of land subsidence dynamics and monitoring in Nigeria. *Discover Environment*, *3*, Article 143.
- [11] Lee, S., Kang, J., & Kim, J. (2024). Prediction of ground subsidence risk in urban centers using underground characteristics information. *Applied Sciences*, *14*(23), 11044. <https://www.mdpi.com/2076-3417/14/23/11044>
- [12] Jacob, T., Pannet, P., Beaubois, F., Baltassat, J. M., & Hannion, Y. (2020). Cavity detection using microgravity in a highly urbanized setting: A case study from Reims, France. *Journal of Applied Geophysics*, *174*, 104113. <https://doi.org/10.1016/j.jappgeo.2020.104113>
- [13] Pringle, J. K., Ruffell, A. H., Styles, P., et al. (2024). Forensic geoscience non-invasive detection and characterisation of underground clandestine complexes, bunkers, tunnels and firing ranges. *Forensic Science International*. <https://www.sciencedirect.com/science/article/pii/S0379073824001142>
- [14] Paul, A., Jayapal, U. M., & Ayothiraman, R. (2025). A review of ground penetrating radar for underground utility detection and subsurface profiling: Challenges, strategies and a future-oriented framework. *Indian Geotechnical Journal*.
- [15] Hall, C. M. (n.d.). *An ominous effervescence: The Bayou Corne sinkhole*. <https://64parishes.org/an-ominous-effervescence>
- [16] Noshahri, H., Meijde, M., & Scholtenhuis, L. (2023). AI-assisted characterization of voids around sewer pipes using in-pipe GPR surveys. *SSRN*. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4506855](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4506855)
- [17] Вижда, С. А., Рева, М. В., Онищук, І. І., & Онищук, В. І. (2013). *Електрометрія*. Київський університет.
- [18] Sutter, E. M., & Barounis, N. (2021). Underground void detection by applying electrical resistivity tomography (ERT) in a limestone quarry. *NZGS Symposium*.
- [19] Negri, S., & Barbolla, D. F. (2025). Challenges in the detection of water-filled cavities in karst environments using electrical resistivity tomography. *Geosciences*, *15*(9), 349. <https://doi.org/10.3390/geosciences15090349>
- [20] Liu, L., Shi, Z., Tsofilias, G. P., et al. (2022). Detection of karst voids by full-waveform inversion. *Soil Dynamics and Earthquake Engineering*, *152*, 107048. <https://doi.org/10.1016/j.soildyn.2021.107048>
- [21] Jiang, W., Zelt, C. A., & Zhang, J. (2020). Detecting an underground tunnel using waveform inversion. *Journal of Applied Geophysics*, *174*, 103957. <https://doi.org/10.1016/j.jappgeo.2020.103957>
- [22] Torrese, P. (2025). Near-surface cavity detection using ambient noise analysis. *Journal of Applied Geophysics*, *241*, 105858. <https://doi.org/10.1016/j.jappgeo.2025.105858>
- [23] Ковальчук, В. В., Кравець, І. Б., & Лучко, Й. Й. (2019). Георадіолокація як неруйнівний метод моніторингу. *Дороги і мости*, 19–20, 117–137.
- [24] Allam, H., Metwally, A., Ismael, M., et al. (2025). Subsurface voids detection using GPR and ERT. *Pure and Applied Geophysics*, *182*, 4137–4167. <https://doi.org/10.1007/s00024-025-03805-z>
- [25] Lytvyn, V., Peleshchak, I., Peleshchak, R., et al. (2023). Hybrid neural network model for mine detection. *Eastern-European Journal of Enterprise Technologies*, *3*(9), 78–85.
- [26] Dai, Q., Lee, Y. H., Sun, H., et al. (2023). 3DInvNet: A deep learning-based GPR data inversion. *arXiv*. <https://arxiv.org/abs/2305.05425>
- [27] Zeng, A., Yan, L., Huang, Y., et al. (2021). Intelligent detection of small faults using SVM. *Energies*, *14*(19), 6242.
- [28] Yu, S., & Ma, J. (2021). Deep learning for geophysics. *Reviews of Geophysics*, *59*(3).
- [29] Xia, J., Peng, R., Li, Z., et al. (2023). Identification of underground artificial cavities. *Sensors*, *23*(19), 8169.
- [30] Vesala, G. T., Ghali, V. S., Subhani, S., et al. (2022). CNN-based subsurface anomaly detection. *SN Computer Science*, *3*, 219.
- [31] Vanga, M., Ojha, M. (2025). Modeling the seismic wave equation using a staggered grid finite-difference method optimized with a genetic algorithm. *Journal of Seismic Exploration*, *34*(2), 1–13. <https://doi.org/10.36922/JSE025290035>
- [32] Li, J., Wan, J., Wang, T., et al. (2022). Leakage simulation and acoustic characteristics. *Advances in Geo-Energy Research*, *6*(3), 181–191. <https://doi.org/10.46690/ager.2022.03.02>
- [33] Praditia, T., Karlbauer, M., Otte, S., et al. (2021). Finite volume neural network. *ICLR SimDL Workshop*. <https://doi.org/10.48550/arXiv.2104.06010>
- [34] Cen, J., & Zou, O. (2024). Deep finite volume method for PDEs. *arXiv*. <https://doi.org/10.48550/arXiv.2305.06863>
- [35] Li, T., Zou, Y., Zou, S., et al. (2025). Finite volume-informed neural networks. *Journal of Computational Physics*, *530*, 113919. <https://doi.org/10.1016/j.jcp.2025.113919>

Received (Надійшла) 10.12.2025

Accepted for publication (Прийнята до друку) 09.01.2026

Publication date (Дата публікації) 27.03.2026

УДК 519.85:004.94:61

DOI: [https://doi.org/10.30837/bi.2026.1\(104\).15](https://doi.org/10.30837/bi.2026.1(104).15)А. М. Чугай<sup>1,2</sup>, Г. М. Яськов<sup>1,3</sup>, О. В. Старкова<sup>2</sup>, Є. Г. Яськова<sup>4</sup><sup>1</sup>ІЕМС ім. А.М. Підгорного, м. Харків, Україна, [chugay.andrey80@gmail.com](mailto:chugay.andrey80@gmail.com),  
ORCID iD: 0000-0002-4079-5632<sup>2</sup>ХНЕУ імені С. Кузнеця, м. Харків, Україна, [olha.starkova@hneu.net](mailto:olha.starkova@hneu.net),  
ORCID iD: 0000-0002-9034-8830<sup>3</sup>ХНУРЕ, м. Харків, Україна, [heorhii.iaskov@nure.ua](mailto:heorhii.iaskov@nure.ua),  
ORCID iD: 0000-0002-1476-1818<sup>4</sup>ХНУ ім. В.Н. Каразіна, м. Харків, Україна, [yelyzaveta.yaskova@student.karazin.ua](mailto:yelyzaveta.yaskova@student.karazin.ua),  
ORCID iD: 0009-0007-6306-3366

## ГІБРИДНИЙ АЛГОРИТМ ПРОСТОРОВОГО ПАКУВАННЯ ДЛЯ МОДЕЛЮВАННЯ БІОМЕДИЧНИХ СИСТЕМ

Просторове пакування є ключовим завданням у моделюванні багатьох біомедичних систем, де ефективне розміщення гетерогенних елементів у складних геометріях має вирішальне значення для точності та функціональності моделей. У роботі запропоновано гібридний алгоритм, що поєднує жадібний вибір, послідовну інкрементальну оптимізацію та локальне переставлення для оптимізації пакування структурних компонентів у обмежених просторових областях. Метод враховує варіації розмірів об'єктів, допускає біологічно обґрунтовані перекриття та адаптується до нестандартних конфігурацій. Ефективність алгоритму продемонстровано на прикладі кількох прикладних варіантів моделювання біомедичних систем.

БІОМЕДИЧНА СИСТЕМА, МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ, КУЛЯ, ЦИЛІНДР, ПРЯМОКУТНИЙ ПАРАЛЕЛЕПІПЕД, ПАКУВАННЯ, ІНКРЕМЕНТАЛЬНА ОПТИМІЗАЦІЯ, АЛГОРИТМ

**A. M. Chuhai, G. M. Yaskov, O. V. Starkova, Y. G. Yaskova. Hybrid spatial packing algorithm for biomedical system modelling.** Spatial packing is a key task in modelling many biomedical systems, where the efficient placement of heterogeneous elements within complex geometries is critical for the accuracy and functionality of models. This paper proposes a hybrid algorithm that combines greedy selection, sequential incremental optimisation, and local rearrangement to optimise the packing of structural components in constrained spatial regions. The method accounts for variations in object sizes, allows biologically justified overlaps, and adapts to non-standard configurations. The efficiency of the algorithm is demonstrated through several applied examples of biomedical system modelling.

BIOMEDICAL SYSTEM, MATHEMATICAL MODELING, SPHERE, CYLINDER, RECTANGULAR PARALLELEPIPED, PACKING, INCREMENTAL OPTIMIZATION, GREEDY ALGORITHM

### Вступ

Просторова організація є фундаментальним принципом як у природних, так і в інженерних системах, впливаючи на все — від клітинної архітектури до проектування медичних пристроїв. У біологічних тканинах розташування клітин визначає ріст, сигнальні процеси та механічні властивості, тоді як у медичних застосуваннях позиціонування терапевтичних агентів або структурних компонентів може впливати на ефективність і безпеку лікування. Розуміння та моделювання того, як різноманітні об'єкти займають простір за геометричних і функціональних обмежень, є критично важливим завданням для багатьох дисциплін.

Біологічні системи демонструють високоефективні стратегії пакування, коли клітини, органи та позаклітинні матриці самоорганізуються у складні функціональні геометрії. Наприклад, щільне пакування нейронів у мозку, орієнтація колагенових волокон у сполучних тканинах та розподіл наночастинок, що доставляють ліки, у цільовій терапії ґрунтуються на ефективних просторових розташуваннях. Аналогічно, у біомедичній інженерії проектування каркасів для регенерації тканин, виготовлення стентів

із контрольованим вивільненням ліків та оптимізація 3D-бідруктованих структур потребують точного контролю над розміщенням компонентів.

Попри важливість цього завдання, комп'ютерне моделювання таких систем залишається складним через гетерогенність форм, динамічні взаємодії та багатомасштабні залежності. Традиційні алгоритми пакування, такі як методи Монте-Карло, пакування сфер або ґратчасті підходи, часто не здатні врахувати нерегулярність і деформованість біологічних об'єктів. Крім того, багато існуючих методів віддають перевагу або швидкості, або точності, але рідко досягають обох показників одночасно, що обмежує їх застосування для великомасштабних або високоточних симуляцій.

У цій статті ми пропонуємо гібридний алгоритм пакування, розроблений для передових завдань моделювання в біології та медицині, з особливим акцентом на планування радіотерапевтичного лікування, такого як брахітерапія та радіохірургія. Точне просторове розташування терапевтичних агентів, наприклад радіоактивних джерел у брахітерапії, або точна доставка доз опромінення в радіохірургічних процедурах потребують складних обчислювальних методів,

здатних враховувати комплексні біологічні та фізичні обмеження. Наш гібридний алгоритм пакування інтегрує дискретні та неперервні методи оптимізації, використовуючи машинне навчання та евристичні стратегії для ефективного моделювання та оптимізації розміщення терапевтичних елементів у біологічних тканинах. Цей підхід дозволяє враховувати багатомасштабні дані — від молекулярного та клітинного рівнів до рівня тканин і органів, що є необхідним для точного моделювання динаміки росту пухлин і реакцій на лікування. Розв'язуючи складні обмеження, пов'язані з розташуванням джерел і формуванням доз — аналогічні задачам складного пакування — алгоритм підтримує ітераційні процеси уточнення, що відповідають клінічним вимогам адаптивного та персоналізованого планування лікування. Застосування цього гібридного алгоритму сприяє точному моделюванню розподілу доз і конфігурацій джерел, підвищуючи ефективність брахітерапії та радіохірургічних втручань і мінімізуючи побічне ураження здорових тканин. Крім того, структура алгоритму адаптована для інтеграції даних реального часу з медичної візуалізації, моделей біологічної відповіді та клінічних параметрів пацієнта, що робить його перспективним інструментом для покращення прийняття рішень у радіотерапії. У цій роботі ми демонструємо, як гібридні алгоритми пакування можуть поєднати комп'ютерне моделювання та клінічну практику, пропонуючи надійне та гнучке рішення для оптимізації просторових і дозиметричних задач, притаманних сучасним біологічним і медичним технологіям.

Аналіз останніх досліджень і публікацій. Класичні моделі пакування використовуються для моделювання того, як клітини, органели або молекули організуються в обмежених середовищах. Ці моделі допомагають пояснити просторову ефективність у таких системах, як архітектура тканин, розподіл пилкових зерен та ущільнення ДНК. Робота [1] є базовим джерелом, що вводить математичне моделювання в біологічному контексті. Моделі пакування циліндрів застосовуються в системах, таких як масиви мікротрубочок, судинні мережі або інженерні каркаси з видовженими порами. Ці моделі допомагають оптимізувати просторову ефективність, механічну стабільність та транспортні властивості як у природних, так і в синтетичних середовищах. Наприклад, просторове розташування мікротрубочок (порожнистих циліндричних полімерів) досліджувалося за допомогою математичних і комп'ютерних моделей для розуміння того, як їхня динамічна збірка та механічні властивості сприяють внутрішньоклітинній організації та генерації сил [2]. У [3] автори досліджують поведінку неорієнтованих циліндрів за умов зростання щільності пакування. Вони використовують чисельні симуляції для моделювання поступового процесу

ущільнення, аналізуючи, як неупорядковані структури еволюціонують у більш впорядковані, кристалічні. Методологія включає відстеження коефіцієнта пакування, локальної координації та параметрів орієнтаційного порядку для характеристики структурних переходів. Окрім структурної біології, моделі пакування знаходять застосування в терапевтичних технологіях. У брахітерапії радіоактивні джерела — часто у формі невеликих циліндрів або сфер — імплантуються в пухлину або поблизу неї. Просторовий розподіл цих джерел має бути ретельно оптимізований для забезпечення рівномірної доставки дози та мінімізації опромінення навколишніх здорових тканин. Класичні системи, такі як системи Парижа та Манчестера, а також сучасні алгоритми 3D-планування ґрунтуються на геометричних принципах, тісно пов'язаних із теорією пакування [4]. У стереотаксичній радіохірургії, особливо при метастазах у мозку, планування кількох променів, що сходяться, можна концептуалізувати як задачу оптимізації просторового розташування сфер. Мета полягає в тому, щоб «упакувати» області високої дози всередині об'єму пухлини, уникаючи критичних структур — завдання, яке дедалі частіше вирішується за допомогою комп'ютерного моделювання та AI-методів формування дози [5]. Останніми роками попит на ефективні та адаптивні стратегії розміщення об'єктів значно зріс у медичних застосуваннях. Ці задачі характеризуються складними геометричними та комбінаторними обмеженнями, де традиційні методи глобальної оптимізації часто не справляються через обчислювальні обмеження або відсутність гнучкості. Тому ми пропонуємо єдину оптимізаційну структуру, що інтегрує інкрементальну оптимізацію, жадібні алгоритми та локальне переставлення. Рішення будується поступово: об'єкти додаються індивідуально або невеликими групами, керуючись жадібною стратегією, яка пріоритизує корисність або просторову ефективність. На відміну від класичних жадібних алгоритмів [6], після кожного додавання виконується неперервне локальне переставлення раніше розміщених об'єктів для збереження допустимості та покращення загальної якості пакування. Ця гібридна методологія ґрунтується на концепціях інкрементальних моделей «рюкзака» [7], багатокритеріальних жадібних евристик [8] та адаптивних структур планування в оптимізації охорони здоров'я [9]. Жадібний компонент забезпечує вибір на кожному кроці найбільш перспективного об'єкта на основі критерію корисності. Цей принцип широко застосовується в задачах типу «рюкзака», де жадібні евристики пропонують швидкі та ефективні наближення, особливо за умов обмежених часу та ресурсів. Останніми роками ці підходи були вдосконалені за допомогою гібридних стратегій, що поєднують жадібний вибір із локальною оптимізацією та адаптивним переставленням.

Такі методи виявилися особливо цінними в галузях зі складними просторовими та функціональними обмеженнями, включаючи медичне планування та проектування терапевтичних пристроїв [10].

Мета статті полягає у розробці гібридного алгоритму просторового пакування, що поєднує жадібний вибір, послідовну інкрементальну оптимізацію та локальну оптимізацію для моделювання біомедичних систем із урахуванням складних геометричних та функціональних обмежень. Завданням дослідження є побудова математичної моделі, яка дозволяє оптимізувати розміщення об'єктів різної форми (куль, циліндрів, багатогранників) у заданій області, забезпечуючи адаптивність до біологічних умов та можливість інтеграції в системи планування лікування.

### 1. Постановка задачі та математична модель

Нехай  $S_i \in \mathbf{R}^3$ ,  $S_i \in \mathbf{R}^3$  є тілом із заданими розмірами  $\mathbf{a}_i$ ,  $i \in I_N = \{1, 2, \dots, N\}$ . Ми визначимо положення тіл як  $\mathbf{u}_i = (\mathbf{v}_i, \Theta_i)$ ,  $\mathbf{v}_i = (x_i, y_i, z_i)$  — координати початку координат,  $\Theta_i$  — кути повороту  $S_i$ ,  $i \in I_N$ . Позначимо тіло  $S_i$  з параметрами розміщення  $\mathbf{u}_i$  як  $S_i(\mathbf{u}_i)$ ,  $i \in I_N$ .

Тіла  $S_i$ ,  $i \in I_N = \{1, 2, \dots, N\}$ , розмістити в області розміщення  $C$ , яка має специфічну просторову форму та метричні характеристики, максимізації кількість розміщених тіл  $n^* \leq N$ .

Основою моделі є клас аналітичних інструментів, відомих як  $\Phi$ -функції (phi-functions) Стояна [11]. Ці функції забезпечують гнучкий і точний спосіб опису просторових взаємозв'язків між об'єктами. Зокрема,  $\Phi$ -функції можуть моделювати: умови контакту (коли об'єкти лише торкаються один одного), задані відстані розділення та контрольовані перекриття (які часто є біологічно прийнятними або необхідними, наприклад, у середовищах м'яких тканин).

Математичну модель задачі можна сформулювати як

$$n^* = \max_{n \in I_N} n \quad (1)$$

для якого існує вектор  $\mathbf{u} = (\mathbf{u}_1, \dots, \mathbf{u}_n) \in W_n$ ,

$$W_n = \{\mathbf{u} \in \mathbf{R}^{3n}, \Phi_i(\mathbf{u}_i) \geq d_i, i \in I_n,$$

$$\Phi_{ij}(\mathbf{u}_i, \mathbf{u}_j) \geq d_{ij}, i < j \in I_n\},$$

$$I_n = \{1, 2, \dots, n\}.$$

Тут,  $\Phi_i(\mathbf{u}_i)$  та  $\Phi_{ij}(\mathbf{u}_i, \mathbf{u}_j)$  — нормалізовані  $\Phi$ -функції. Значення  $d_{ij}$  задає мінімальну відстань між тілами  $S_i$  та  $S_j$  для  $d_{ij} \geq 0$  та глибину накладання  $S_i$  та  $S_j$  для  $d_{ij} < 0$ . Відповідно, значення  $d_i$  задає відстань між тілом  $S_i$  та межею області розміщення для  $d_i \geq 0$  та виступ тіла  $S_i$  за цю межу для  $d_i < 0$ .

Хоча побудова нормалізованих  $\Phi$ -функцій загалом є складним завданням, у попередніх роботах [12] наведено такі функції для ряду стандартних геометричних тіл, що полегшує практичну реалізацію моделі.

Розв'язання задачі (1) зводиться до послідовного пошуку допустимої точки  $\mathbf{u} = (\mathbf{u}_1, \dots, \mathbf{u}_n) \in W_n$  для інкрементальних значень  $n$ , а отже, і для інкрементальної кількості змінних  $3n$ . Алгоритм пошуку таких точок розглядається в наступному розділі.

### 2. Гібридний алгоритм

Ми застосовуємо евристичний підхід, який послідовно додає тіла, використовуючи жадібну стратегію. Для підвищення ефективності традиційних жадібних алгоритмів ми вводимо техніку інкрементальної пакетної оптимізації, що дозволяє локально переставляти вже розміщені тіла, створюючи додатковий простір для наступних вставок. На відміну від класичних жадібних алгоритмів, де на кожному кроці оптимізується лише пакет, що відповідає поточному тілу, наш підхід поступово збільшує розмір пакета, забезпечуючи більш гнучке та ефективне використання простору. Розміщення кожного тіла формулюється як задача нелінійної оптимізації, що включає його гометичне перетворення. Мета полягає в максимізації коефіцієнта гомететії, починаючи з  $0 < \lambda_0 \leq 1$ . Тіло вважається розміщеним, якщо оптимізація дає коефіцієнт  $\lambda = 1$ . Для цього ми послідовно розв'язуємо такі задачі нелінійного програмування:

$$(\mathbf{u}^*, \lambda^*) = \arg \max_{\mathbf{u}=(\mathbf{u}, \lambda)=(\mathbf{u}_1, \lambda) \in G_1} \lambda, \quad (2)$$

$$G_n = \{(\mathbf{u}, \lambda) \in \mathbf{R}^4, \Phi_1(\mathbf{u}_1, \lambda) \geq d_1, 0 \leq \lambda \leq 1\},$$

$$(\mathbf{v}^*, \lambda^*) = \arg \max_{\mathbf{v}=(\mathbf{u}_1, \dots, \mathbf{u}_n, \lambda_n) \in G_n} \lambda, \quad n = 2, \dots, n^*, \quad (3)$$

$$G_n = \{(\mathbf{u}, \lambda) \in \mathbf{R}^{3n+1}, \Phi_i(\mathbf{u}_i) \geq d_i, i \in I_{n-1},$$

$$\Phi_n(\mathbf{u}_n, \lambda) \geq d_n, \Phi_{ij}(\mathbf{u}_i, \mathbf{u}_j) \geq d_{ij}, i < j \in I_{n-1},$$

$$\Phi_{in}(\mathbf{u}_i, \mathbf{u}_n, \lambda) \geq d_{in}, i \in I_{n-1}, 0 \leq \lambda \leq 1\}.$$

Якщо  $\lambda^* = 1$ , то знайдено точку  $\mathbf{u} \in W_n$ . Якщо  $\lambda^* < 1$ , то ми вважаємо, що  $n^* = n - 1$ , та точка  $\mathbf{U} \in W_{n-1}$ , яку отримано при розв'язанні попередньої задачі (2), (3), розглядається як наближення до розв'язку задачі (1).

Початкові значення  $u_1$  для задачі (2) вибираються випадково, але гарантуючи допустимість. Для задач (3),  $\mathbf{u}_n$  також вибирається випадково, водночас  $\mathbf{u}_i$ ,  $i \in I_{n-1}$ , складають компоненти вектора  $\mathbf{u}_i$ , отриманого для попередньої задачі. Початкове значення коефіцієнту гомететії  $\lambda_0 = 1$ . Якщо після декількох спроб для цього значення не можна визначити початкову точку  $(\mathbf{u}, \lambda_0) \in G_n$ , значення  $\lambda_0$  зменшується до нуля, доки не буде знайдено допустиму точку задачі (3).

На кожному наступному кроці використовуються результати попередньої задачі, але розташування всіх раніше розміщених тіл коригується зі збільшенням коефіцієнта гомететії. Така техніка забезпечує допустимість на кожному етапі завдяки динамічному

налаштуванню коефіцієнта гомотетії, що дозволяє здійснювати контрольоване та стабільне розміщення тіл без порушення просторових обмежень.

### 3. Числові результати

Розглянуто декілька прикладних прикладів.

Для оцінки запропонованої моделі пакування в біологічно мотивованому середовищі ми розглядаємо задачу розміщення витягнутих циліндричних об'єктів у обмеженій тривимірній області. Область визначається як прямокутний паралелепіпед, що представляє спрощений сегмент внутрішньоклітинного простору. Кожен об'єкт наближено моделює мікротрубочку та описується як опукла багатогранна призма, що максимально відповідає геометрії циліндра. Така багатогранна репрезентація використовується для застосування попередньо побудованих нормалізованих  $\Phi$ -функцій [12], які описують просторові взаємозв'язки між твердими тілами та між тілами і межами області.

Кожен об'єкт визначається трьома параметрами трансляції та двома параметрами обертання, що дозволяє довільне розташування та орієнтацію в просторі. Мета полягає у визначенні максимальної кількості таких багатограних циліндрів, які можна розмістити в області без порушення геометричних обмежень. У деяких прикладах вводиться мінімальна відстань між центрами для відображення біологічно правдоподібного інтервалу, що спостерігається в щільних масивах мікротрубочок, де об'єкти розташовані близько, але не перетинаються, часто розділені білками, асоційованими з мікротрубочками [13]; в інших випадках дозволяється прямий контакт між об'єктами, моделюючи щільно упаковані пучки, як це спостерігається в експериментах зі зв'язування мікротрубочок, індукованих білками, асоційованими з мікротрубочками (MAP) [14]. Така формалізація відображає суттєву геометричну складність організації мікротрубочок у клітинному середовищі, де філаменти щільно упаковані, мають різну орієнтацію та обмежені нерегулярними об'ємами.

Приклад 1. Числові дані:  $n^* = 30$ ;  $S_i$  – прями циліндри з напіввисотами  $h_i = 6$  та радіусами  $r_i = 0.5$ ,  $i \in I_n$  (див. Таблицю 1);  $C$  – прямокутний паралелепіпед з розмірами  $W = 30.23$ ,  $L = 21.64$ ,  $H = 11.19$ ; мінімально допустимі відстані – відповідно  $d_{ij} = 2$ ,  $d_i = 0$ . Центри циліндрів знаходяться на одній відстані від основ. Координати центрів циліндрів та їхні кути повороту  $\varphi_i$  та  $\omega_i$  наведені в таблиці 1. Час обчислення був приблизно 5 хвилин. Ілюстрацію розмішених циліндрів показано на Рис. 1.

Приклад 2. Числові дані:  $n^* = 30$ ;  $W = 21.56$ ,  $L = 18.91$ ,  $H = 7.01$ . Мінімально допустимі відстані – відповідно  $d_{ij} = 0$ ,  $d_i = 0$ . Час обчислення – біля 3 хвилин. Ілюстрацію наведено на Рис. 2.

Таблиця 1

Параметри розміщення

№	$x_i$	$y_i$	$z_i$	$\varphi_i$	$\omega_i$
1	6.917	20.459	35.883	-0.112	2.601
2	24.592	1.335	33.005	-3.002	-2.001
3	29.113	4.639	36.186	2.5	-0.125
4	25.098	17.692	31.168	-2.512	-1.447
5	2.806	16.01	31.728	1.12	2.033
6	13.226	10.916	37.237	-0.874	0.346
7	15.414	16.01	35.829	1.12	-1.907
8	10.964	11.731	31.168	2.022	-4.481
9	19.372	5.508	40.158	-1.07	1.36
10	8.043	5.623	36.444	1.12	2.666
11	25.006	16.609	39.998	-0.593	-1.417
12	1.66	2.871	35.665	2.729	0.214
13	12.777	5.58	36.825	-0.695	-0.491
14	16.313	4.008	31.606	-0.644	-1.696
15	7.687	1.102	33.069	-3.041	1.13
16	3	14.405	36.02	-2.022	-1.333
17	21.697	5.623	35.764	2.022	-1.175
18	5.623	6.559	32.47	0.3	-1.228
19	9.395	17.132	37.124	0.76	0.534
20	18.954	17.091	40.161	0.77	1.431
21	24.088	16.399	35.039	0.969	-1.031
22	24.94	11.968	39.124	2.64	-1.891
23	5.866	6.871	40.161	-0.438	-1.681
24	26.842	7.845	37.867	0.804	2.36
25	17.15	5.144	36.553	-0.94	-0.608
26	19.027	15.492	32.9	-2.022	-0.742
27	12.346	15.672	40.161	2.357	-1.429
28	25.045	9.262	31.168	-1.12	1.339
29	6.917	20.459	35.883	-0.112	2.601
30	24.592	1.335	33.005	-3.002	-2.001

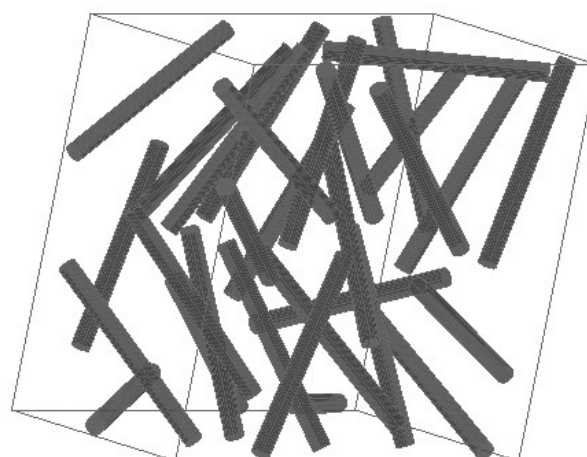
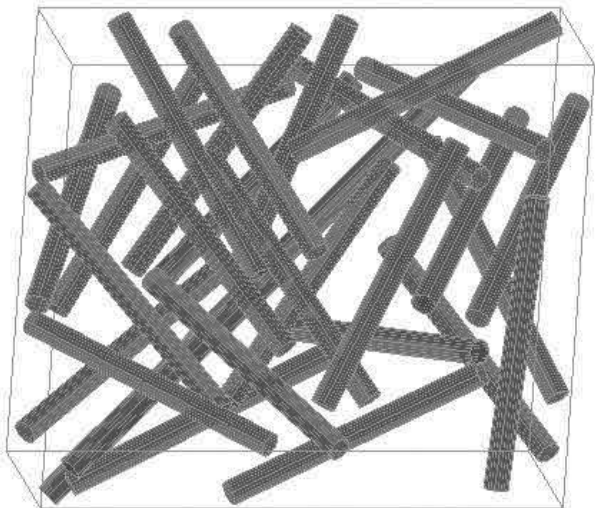


Рис. 1. Ілюстрація розміщення циліндрів з мінімально допустимими відстанями



**Рис. 2.** Ілюстрація розміщення циліндрів без мінімально допустимих відстаней

Приклад 2. Числові дані:  $n^* = 30$ ;  $W = 21.56$ ,  $L = 18.91$ ,  $H = 7.01$ . Мінімально допустимі відстані – відповідно  $d_{ij} = 0$ ,  $d_i = 0$ . Час обчислення – біля 3 хвилин. Ілюстрацію наведено на Рис. 3.

Також ми розглядаємо задачу просторового розподілу сферичних високодозових джерел випромінювання у пухлині під час стереотаксичної радіохірургії. Кожне джерело представляє фокусну точку, де збігається кілька променів випромінювання, утворюючи сферичну ізодозну область. Мета полягає у визначенні оптимального розташування цих куль для забезпечення повного покриття пухлини при мінімізації опромінення навколишніх здорових тканин та критичних структур. Область визначається як сегментований 3D-об’єм, змодельований як опуклий багатогранник, що відображає анатомічні межі зони, яка містить пухлину. Сама пухлина розглядається як цільова область, водночас навколишні органи-мішені моделюються як зони виключення зі строгими просторовими обмеженнями. Кожне джерело випромінювання моделюється кулею зі змінним радіусом, що відповідає різним інтенсивностям дози або конфігураціям променів. Ключовою особливістю цієї моделі є допущення часткового перекриття між кулями. Це відображає клінічну практику, де незначні перекриття необхідні для уникнення недостатнього дозування на межах між сусідніми «шотами». Якщо кулі розташовувати в строгому контакті без перекриття, можуть залишитися прогалини в розподілі дози, що призведе до неповного покриття пухлини. Для вирішення цієї проблеми обмеження неперекриття в моделях (1) та (3) модифікуються, щоб дозволити мінімально необхідне перекриття, забезпечуючи безперервність дози в об’ємі.

Приклад 3. Числові дані:  $n^* = 15$ ;  $S_i$  – кулями з радіусами  $r_i$ ,  $i \in I_n$  (див. Таблицю 2);  $C$  – опуклий багатогранник, заданий вершинами та гранями (див. Таблицю 3). Мінімально допустимі

відстані –  $d_i = d_{ij} = -3.5$ . Координати центрів куль наведено в Таблиці 3. Час обчислення – приблизно 1 година. Ілюстрацію розміщення куль наведено на Рис. 3.

**Таблиця 2**

**Радіуси та координати куль**

№	$r_i$	$x_i$	$y_i$	$z_i$
1	6	14.456	13.888	6.511
2	6	6.086	9.449	13.121
3	4	11.170	3.240	9.510
4	4	15.477	7.747	12.781
5	4	8.764	18.069	11.761
6	4	5.857	15.631	5.831
7	4	7.910	8.966	3.264
8	4	13.325	13.984	15.234
9	3	13.553	6.743	2.186
10	3	1.856	15.618	10.515
11	3	7.879	15.843	17.437
12	3	15.930	6.114	6.859
13	3	12.126	7.740	17.950
14	3	2.372	9.844	6.050
15	3	5.726	4.523	6.929

**Таблиця 3**

**Координати вершин багатогранника**

№	$x$	$y$	$z$
1	12.14	7.5	8.82
2	4.64	-7.5	14.63
3	-4.64	-7.5	14.63
4	12.14	-7.5	8.82
5	-15	7.5	0
6	-12.14	-7.5	-8.82
7	-4.64	7.5	-14.63
8	4.64	7.5	-14.63
9	12.14	7.5	-8.82
10	15	7.5	0
11	0	16.8	0
12	0	-16.8	0

Багатогранник  $P$  має 20 граней, які визначаються такими трійками вершин: 1-11-3, 3-11-5, 5-11-7, 7-11-9, 9-11-1, 1-3-2, 2-3-4, 3-5-4, 4-5-6, 5-7-6, 6-7-8, 7-9-8, 8-9-10, 9-1-10, 10-1-2, 2-12-10, 4-12-2, 6-12-4, 8-12-6, 10-12-8.

Процес оптимізації реалізує інкрементальну пакетну стратегію, описану раніше. На кожному кроці додається нова сфера, а позиції раніше розміщених сфер локально коригуються для підвищення просторової ефективності та збереження допустимості. Розміщення на основі гомотетії гарантує, що кожна

сфера приймається лише тоді, коли вона задовольняє геометричні обмеження, включно з необхідним перекриттям та відступом від меж.

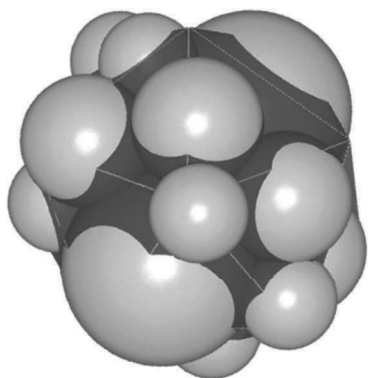


Рис. 3. Розміщення куль з накладеннями

Числові експерименти, проведені в цьому дослідженні, демонструють надійність та ефективність запропонованого гібридного алгоритму пакування при моделюванні складних біологічних і медичних сценаріїв. Результати підкреслюють здатність алгоритму точно оптимізувати просторові конфігурації за множиною обмежень, що відображає реалістичні умови, характерні для брахітерапії та планування радіохірургічного лікування. У різних тестових випадках алгоритм стабільно досягав високої точності у розміщенні джерел та розподілі дози, зберігаючи обчислювальну ефективність, придатну для клінічних застосувань.

Експерименти підтверджують потенціал алгоритму для поліпшення персоналізованого планування лікування шляхом інтеграції багатомасштабних біологічних даних та адаптації до анатомічних особливостей пацієнта. Загалом числові результати підтверджують, що гібридний підхід до пакування є перспективним і практичним інструментом для підвищення точності та ефективності радіотерапевтичних втручань.

### Висновки

Для вирішення проблеми моделювання складних біологічних і медичних систем гібридні алгоритми пакування стали потужними інструментами, що поєднують кілька обчислювальних стратегій. Ці підходи інтегрують методи машинного навчання, симуляційні методи та алгоритми оптимізації для розв'язання складних задач пакування у структурній біології, медичному прийнятті рішень та моделюванні біохімічних систем.

У цьому дослідженні ми представили гібридний алгоритм просторового пакування, який поєднує жадібний вибір, інкрементальне розміщення та локальне переставлення для вирішення складних задач моделювання в біології та медицині. Підхід розроблений для роботи з широким спектром форм об'єктів, допускає біологічно значущі перекриття та добре адаптується до нерегулярних просторових областей.

Це робить його сильним кандидатом для застосувань, таких як моделювання тканин, планування лікування та проектування біомедичних структур.

Однією з ключових переваг методу є використання жадібних алгоритмів — не ізольовано, а як частини ширшої гібридної стратегії. Жадібний компонент допомагає швидко визначити перспективні розміщення на основі корисності, тоді як етапи локального переставлення та оптимізації на основі гомотетії уточнюють конфігурацію для покращення загальної якості пакування. Це поєднання допомагає подолати типові слабкі сторони суто жадібних методів, особливо в щільних або геометрично складних середовищах.

Ще однією важливою особливістю є використання нормалізованих  $\Phi$ -функцій Стояна, які надають моделі тонкий контроль над тим, як об'єкти взаємодіють у просторі: торкаються, не перетинаються або частково перекриваються. Такий рівень контролю особливо цінний у біомедичних контекстах, де просторові обмеження часто суворі та безпосередньо пов'язані з функціональними результатами.

Загалом алгоритм пропонує гнучку та ефективну основу для просторового моделювання. Його модульний дизайн та обчислювальна продуктивність роблять його придатним не лише для конкретних біомедичних прикладів, розглянутих у статті, але й для набагато ширшого спектра задач у комп'ютерній біології, медичній фізиці та інженерії. Хоча ми показали лише кілька конкретних прикладів, таких як пакування мікротрубочок та планування радіохірургії, сам метод є високопродатним і може застосовуватися в багатьох інших сферах, де важлива розумна просторова організація.

Запропонований гібридний алгоритм пакування продемонстрував значний потенціал у вирішенні складних задач моделювання в біологічній та медичній галузях, зокрема в плануванні радіотерапевтичного лікування, такого як брахітерапія та радіохірургія. Поєднуючи дискретні та неперервні методи оптимізації разом із сучасними методами машинного навчання та евристичними підходами, алгоритм ефективно враховує складні біофізичні обмеження та забезпечує точне розміщення терапевтичних елементів у тканинах. Інтеграція багатомасштабних даних дозволяє більш реалістично моделювати ріст пухлини та реакцію на лікування, сприяючи розробці персоналізованих і адаптивних планів терапії. Реалізація цього алгоритму підвищує точність дозиметрії та знижує опромінення здорових тканин, що є критичними факторами для покращення ефективності та безпеки радіотерапевтичних втручань. Крім того, гнучкість і масштабованість цього підходу відкривають перспективні можливості для майбутньої інтеграції з системами візуалізації та біологічними моделями, розширюючи його застосування в клінічній практиці

та підтримуючи обґрунтований вибір рішень. Загалом розроблений гібридний алгоритм пакування є цінним інструментом із значним потенціалом для вдосконалення планування та виконання сучасних біомедичних процедур.

#### Список літератури

- [1] Garfinkel A., Shevtsov J., Guo, Y. (2017). Modeling life: the mathematics of biological systems. Springer. <https://doi.org/10.1007/978-3-319-59731-7>.
- [2] Gutmann G., Inoue D., Kakugo A., Konagaya A. (2014). Real-time 3D microtubule gliding simulation. Ma S., Jia L., Li X., Wang L., Zhou H., Sun X. Life system modeling and simulation. ICSEE LSMS 2014, Communications in Computer and Information Science, vol. 461., Berlin, Heidelberg : Springer. [https://doi.org/10.1007/978-3-662-45283-7\\_2](https://doi.org/10.1007/978-3-662-45283-7_2).
- [3] Gutmann G., Inoue D., Kakugo A., Konagaya A. (2014). Real-time 3D microtubule gliding simulation. Ma S., Jia L., Li X., Wang L., Zhou H., Sun X. Life system modeling and simulation. ICSEE LSMS 2014, Communications in Computer and Information Science, vol. 461., Berlin, Heidelberg : Springer. [https://doi.org/10.1007/978-3-662-45283-7\\_2](https://doi.org/10.1007/978-3-662-45283-7_2).
- [4] Thomadsen B., Venselaar J., Li Z. (2011). Physics and clinical aspects of brachytherapy. In Levitt S., Purdy J., Perez C., Poortmans P. (eds.). Technical Basis of Radiation Therapy. Medical Radiology. Berlin, Heidelberg. Springer, pp. 401–434. [https://doi.org/10.1007/174\\_2011\\_361](https://doi.org/10.1007/174_2011_361).
- [5] Lee C.-C., Yang H.-C., Wu H.-M., et al. (2024). Computational modeling and AI in radiation neuro-oncology and radiosurgery. In A. Di Ieva, E. Suero Molina, S. Liu and C. Russo (eds.) Computational Neurosurgery, Advances in Experimental Medicine and Biology. Vol. 1462, pp. 307–322, Springer, Cham. [https://doi.org/10.1007/978-3-031-64892-2\\_18](https://doi.org/10.1007/978-3-031-64892-2_18).
- [6] Yaskov G., Chugay A. (2020). Packing equal spheres by means of the block coordinate descent method. CEUR Workshop Proceedings. Vol. 2608, pp. 150–160. <https://doi.org/10.32782/cmis/2608-13>.
- [7] Faenza Y., Segev D., Zhang L. (2023). Approximation algorithms for the generalized incremental knapsack problem. Math. Program. 198, pp. 27–83. <https://doi.org/10.1007/s10107-021-01755-7>.
- [8] Yuan J., Li Y. (2021). Solving binary multi-objective knapsack problems with novel greedy strategy. Memetic Comp. 13, pp. 447–458. <https://doi.org/10.1007/s12293-021-00344-7>.
- [9] Bernardelli A. M. (2025). Methods for combinatorial optimization and their applications. 4OR-Q J Oper Res. <https://doi.org/10.1007/s10288-025-00589-9>.
- [10] Wilbaut C., Hanafi S., Coelho I. M., Lucena A. (2022). The knapsack problem and its variants: formulations and solution methods. In Salhi S., Boylan J. (eds.) The Palgrave Handbook of Operations Research. Palgrave Macmillan. Cham, pp. 105–151. [https://doi.org/10.1007/978-3-030-96935-6\\_4](https://doi.org/10.1007/978-3-030-96935-6_4).
- [11] Bennell J., Scheithauer G., Stoyan Y., Romanova T. (2010). Tools of mathematical modeling of arbitrary object packing problems. Ann Oper Res 179, pp. 343–368. <https://doi.org/10.1007/s10479-008-0456-5>.
- [12] Chugay A.M., Zhuravka A.V. (2020). Packing Optimization Problems and Their Application in 3D Printing. In Hu Z., Petoukhov S., Dychka I., He M. (eds.) Advances in Computer Science for Engineering and Education III. ICCSEE 2020, Advances in Intelligent Systems and Computing, vol 1247, Springer, Cham, pp. 75–85. [https://doi.org/10.1007/978-3-030-55506-1\\_7](https://doi.org/10.1007/978-3-030-55506-1_7).
- [13] Yeow Z. Y., Sarju S., Chang F.-C., et al. (2025). Mesoscale regulation of microtubule-organizing centers by the E3 ligase TRIM37. Nat Struct Mol Biol. <https://doi.org/10.1038/s41594-025-01540-6>.
- [14] Schmidt-Marcec S., Ross A., Smertenko A. (2023). Quantification of microtubule-bundling activity of MAPs using TIRF microscopy. In Hussey P. J., Wang P. (eds.) The Plant Cytoskeleton, Methods in Molecular Biology. New York, NY : Humana, vol. 2604, pp. 1–12. [https://doi.org/10.1007/978-1-0716-2867-6\\_1](https://doi.org/10.1007/978-1-0716-2867-6_1).

*Received (Надійшла) 15.12.2025*

*Accepted for publication (Прийнята до друку) 07.01.2026*

*Publication date (Дата публікації) 27.03.2026*

УДК 519.85:004.94:61

DOI: [https://doi.org/10.30837/bi.2026.1\(104\).16](https://doi.org/10.30837/bi.2026.1(104).16)К. О. Веретельник<sup>1</sup><sup>1</sup>ХНЕУ ім. С. Кузнеця, м. Харків, Україна, kostiantyn.veretelnik@hneu.net,  
ORCID iD: 0009-0006-1114-8846

## ПОСЛІДОВНИЙ МУЛЬТІСТАРТОВИЙ АЛГОРИТМ РОЗМІЩЕННЯ БАГАТОВИМІРНИХ КУЛЬ

У роботі розглядається задача розміщення конгруентних багатовимірних куль у евклідовому просторі в кулі мінімального радіуса, яка належать до класу NP-складних. Запропоновано швидкий послідовний мультістартовий алгоритм розміщення куль, який ґрунтується на поетапному додаванні елементів до конфігурації з локальним уточненням положення кожної нової кулі. На кожному кроці для чергової кулі генерується початкове допустиме положення, після чого здійснюється спрямований пошук у напрямку до початку координат. За допомогою дихотомії локалізації визначається перший контакт із уже розміщеними кулями, що дає змогу сформулювати обмежену множину активних геометричних обмежень. Подальше уточнення положення нової кулі виконується в локальному околі точки контакту. На заключному етапі оптимізаційного процесу для зменшення залежності результату від вибору стартової точки використовується мультістартова стратегія для останніх куль, після чого обирається найліпше з отриманих локально допустимих розміщень. Це дає змогу зменшити накопичення локальних геометричних дефектів без виконання повної глобальної оптимізації. Запропонований алгоритм характеризується низькою обчислювальною складністю, добре масштабується з ростом розмірності простору та зберігає можливість розпаралелювання обчислень. Результати числових експериментів для розмірностей аж до  $d=64$  підтверджують ефективність методу та поліпшення якості отриманих конфігурацій порівняно з одностартовими послідовними підходами.

БАГАТОВИМІРНА КУЛЯ, NP-СКЛАДНА ЗАДАЧА, МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ, ПОСЛІДОВНА ОПТИМІЗАЦІЯ, МУЛЬТІСТАРТОВИЙ АЛГОРИТМ

**K. O. Veretelnik. Sequential multistart algorithm for placing multidimensional spheres.** The paper addresses the problem of packing congruent multidimensional spheres into a Euclidean sphere of minimal radius, a problem in NP-hard. A fast sequential multi-start packing algorithm is proposed, based on the incremental addition of elements to the configuration, followed by local refinement of each new sphere's position. At each step, an initial feasible position for the current sphere is generated, followed by a directed search toward the origin. Bisection localisation is employed to identify the first contact with previously placed spheres, enabling the formation of a restricted set of active geometric constraints. Further refinement of the sphere's position is conducted within the local neighbourhood of the contact point. To reduce the impact of the starting point selection on the final result, a multi-start strategy is applied during the final stages of the packing process, after which the best locally feasible configuration is selected. This approach minimises the accumulation of local geometric defects without necessitating full global optimisation. The proposed algorithm is characterised by low computational complexity, scales effectively with increasing space dimensionality, and maintains the potential for parallelisation. Numerical experiments for dimensions up to  $d=64$  confirm the efficiency of the method and demonstrate improved configuration quality compared to single-start sequential approaches

MULTIDIMENSIONAL BALL, NP-HARD PROBLEM, MATHEMATICAL MODELING, SEQUENTIAL OPTIMIZATION, MULTISTART ALGORITHM

### Вступ

Багатовимірні геометричні моделі відіграють важливу роль у сучасних підходах до опрацювання та передавання інформації. У теорії зв'язку точки високої розмірності асоціюються із сигналами, а їх околиці можна розглядати як області коректного прийому в шумових каналах. Форма та взаємне розташування цих областей визначають мінімальні відстані між кодовими словами, енергетичну ефективність системи та середню ймовірність помилки. Зокрема, у гаусових каналах кожне сигнальне слово природно співвідноситься з кулею у  $d$ -вимірному просторі, а проблема побудови сигнальної конфігурації трансформується у задачу розміщення багатовимірних куль без їх взаємного перетину [1]. Тому задачі геометричного проектування, зокрема задачі пакування, виступають фундаментальним інструментом для розроблення ефективних методів кодування та оптимізації сигнальних конфігурацій [2].

У високих розмірностях геометричні властивості простору змінюються кардинально: об'єми різко концентруються біля границь, взаємні відстані між точками стають майже однаковими, а оптимальні структури демонструють складні симетрії, які важко відтворити конструкторськими методами. Через це навіть базові задачі, такі як розміщення множини конгруентних куль стають обчислювально надзвичайно складними. Загальна постановка задачі пакування належить до класу NP-складних, що унеможливує використання глобальних методів оптимізації у просторах розмірності, характерних для прикладних задач зв'язку та інформаційної безпеки [1].

У багатьох практичних задачах метою є не побудова глобально оптимального пакування, а отримання допустимої високовимірної конфігурації, яка задовольняє геометричні обмеження й водночас може бути створена швидкими обчислювальними методами [3]. Такі конструктивні схеми особливо актуальні тоді,

коли розміщення відбувається у контейнері обмеженого радіуса, а кількість об'єктів є великою. За таких умов методи глобальної оптимізації стають практично невикористовуваними, а тому ключову роль відіграють ефективні послідовні алгоритми, здатні будувати конфігурацію крок за кроком, опираючись лише на локальну інформацію про взаємодію нової кулі з уже розміщеними. Вони відомі як алгоритми RSA (випадкового послідовного додавання) або онлайн-алгоритми [4,5].

У теорії кодування та модулювання значна увага приділяється пошуку оптимальних решіткових структур у багатовимірних просторах, оскільки решітки забезпечують найліпші відомі асимптотичні властивості для багатьох задач, зокрема, мінімізацію середнього квадрату відхилення, максимізацію мінімальної відстані та побудову ефективних сигнальних сузір'їв. У низці вимірностей такі структури відомі: наприклад, решітки D4, E8 або решітка Ліча забезпечують оптимальні сферичні пакування і часто розглядаються як еталонні моделі для високовимірного модулювання [6]. Проте, у практичних сигнальних просторах завжди присутні додаткові обмеження, які роблять використання класичних решіток у чистому вигляді неможливим або неефективним. Зокрема, амплітудні обмеження передавача приводять до того, що решітка фактично руйнується біля межі контейнера, а велика частина її структури (особливо точки з великими нормами) не може бути використана без порушення вимог до потужності сигналу або через обмеження характеристики підсилювача. У багатовимірному модулюванні саме ця проблема чітко простежується у роботах з геометричного шейпінгу, де автори наголошують, що реалізовані сузір'я мусять задовольняти обмеження на максимальну амплітуду, і тому класичні решіткові пакування потребують модифікації, урізання або повного переформулювання задачі розміщення точок [7, 8].

Через ці практичні обмеження сучасні системи модулювання нерідко відходять від строгих решіткових структур на користь гібридних, нерешіткових або оптимізованих методом локального пошуку конфігурацій, які ліпше узгоджуються з амплітудними, енергетичними й технологічними обмеженнями реальних каналів зв'язку [9]. Фактично це знову приводить до задачі розміщення куль у скінченному контейнері, де необхідно знайти якомога щільніше або енергетично вигідніше розташування центрів.

Класичні послідовні методи, попри низькі обчислювальні витрати, мають суттєвий недолік: якість результуючої конфігурації сильно залежить від початкових положень [4, 10]. Це зумовлює потребу у методах, здатних комбінувати простоту послідовних схем із гнучкішими механізмами локальної адаптації.

У цій роботі пропонується підхід, спрямований

саме на розв'язання цієї проблеми. Ми розробляємо послідовний мультістартовий алгоритм, у якому кожна нова куля не лише додається до конфігурації, а й проходить процес локалізованого уточнення свого положення. Такий підхід поєднує масштабованість простих конструктивних методів із поліпшеною стійкістю до вибору стартової точки, що є критично важливим для задач моделювання багатовимірних сигнальних конфігурацій.

## 1. Постановка задачі та математична модель

Розглядається задача розміщення множини конгруентних гіперкуль у  $d$ -вимірному евклідовому просторі. Нехай задано  $N$  куль радіуса  $r > 0$ , центри яких позначимо через  $x_1, x_2, \dots, x_N \in \mathbb{R}^d$ .

Усі кулі мають бути розташовані всередині більшої кулі мінімального можливого радіуса  $R$  з центром у початку координат. Будь-яка точка кожної розміщеної кулі не повинна виходити за межі області  $S(R) = \{x \in \mathbb{R}^d : x \leq R\}$ .

Тобто має виконуватися умова:

$$\|x_i\| + r \leq R, i = 1, 2, \dots, N. \quad (1)$$

Крім того, кулі не повинні накладатися. Це означає, що відстань між центрами будь-яких двох куль має бути не меншою за  $2r$ :

$$\|x_i - x_j\| \geq 2r, i, j = \{1, \dots, N\}, i < j. \quad (2)$$

Метою є побудова конфігурації

$$X = (x_1, x_2, \dots, x_N) \in \mathbb{R}^{Nd},$$

яка задовольняє всі геометричні обмеження та мінімізує радіус контейнера.

Математична модель задачі має такий вигляд:

$$R^* = \min_{(R, X) \in W} R. \quad (3)$$

де множина допустимих розв'язків  $W \subset \mathbb{R}^{Nd+1}$  визначається нерівностями (1), (2).

Це задача нелінійного програмування з великою кількістю обмежень та лінійною функцією цілі. Відомо, що її розв'язання є NP-складним, а глобальна оптимізація для великих  $d$  і  $N$  практично нездійсненна

## 2. Послідовна модель розміщення

Оскільки контейнерна область  $S(R)$  є кулею, а кожна розмішувана куля має відомий радіус  $r$ , то центр будь-якої кулі не може розташовуватися далі ніж на відстані  $R - r$  від початку координат. З цього безпосередньо випливає, що оптимальний радіус контейнера визначається найвіддаленішим центром серед усіх розміщених куль, тобто

$$R^*(x_1, x_2, \dots, x_N) = \max_i \|x_i\| + r. \quad (4)$$

Ця рівність означає, що змінна  $R$  не є незалежною оптимізаційною змінною: її оптимальне значення задається як функція від координат центрів. У термінах

математичної оптимізації це відповідає аналітичній згортці початкової задачі за параметром  $R$  (variable elimination). Внаслідок такої згортки відбувається проекція задачі на простір змінних  $X$ , а розмірність задачі зменшується на одиницю.

Таким чином, задача мінімізації радіуса контейнера еквівалентно перетворюється на мінімаксну задачу:

$$\min_{x_1, \dots, x_N} \max_i \|x_i\| \quad (5)$$

за умов ненакладання (2).

Такий перехід є типовою згорткою параметризованої задачі в теорії геометричної оптимізації, який дає змогу уникнути окремої оптимізації за  $R$  та працювати лише у просторі конфігурацій центрів.

Розглянемо послідовну конструктивну схему. Нехай на кроці  $k$  уже побудовано допустимі центри  $x_1, \dots, x_{k-1}$  і потрібно знайти положення нової кулі  $k$ .

Локальна задача для наступної кулі набуває вигляду

$$x_k^* = \arg \min_{x \in \mathbb{R}^d} \left\{ \|x\|^2 : \|x - x_i\|^2 \geq 4r^2, i = 1, \dots, k-1 \right\}. \quad (6)$$

Таким чином, локальна послідовна мінімізація норми кожної кулі є прямим наслідком згортки параметра (4) та еквівалентного переходу до мінімаксної постановки (5), (2).

Функція цілі

$$f(x) = \|x\|^2 \quad (7)$$

є опуклою. Множина допустимих розв'язків визначається перетином зовнішніх областей куль

$$D^{(k)} = \bigcap_{i=1}^{k-1} \{x : \|x - x_i\|^2 \geq 4r^2\}, \quad (8)$$

що є неопуклою із складною геометрією для довільного  $d$ . Ілюстрацію множини допустимих рішень для  $d = 2$  наведено на рис. 1.

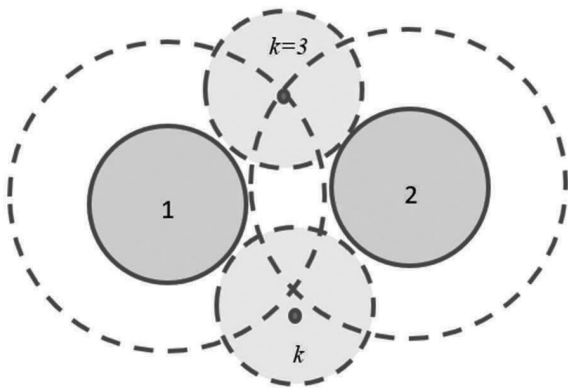


Рис. 1. Ілюстрація множини допустимих розв'язків

Мінімум задачі (6) знаходиться у крайніх точках множини допустимих рішень, які визначаються системами відповідних рівнянь. Для отримання такої точки необхідно розв'язати систему  $d$  сумісних квадратичних рівнянь методом виключення змінних. Такі точки визначаються системами рівнянь вигляду

$$\|x - x_j\|^2 \geq 4r^2, j = 1, \dots, d \quad (9)$$

Розв'язання кожної системи (9) зводиться до послідовного виключення змінних у системі із  $d$  квадратичних рівнянь. Однак далеко не всі розв'язки таких систем є припустимими: частина з них не належить множині  $D^{(k)}$  через порушення інших нерівностей у (8). Гарантований пошук глобального мінімуму вимагає повного перебору всіх можливих комбінацій активних обмежень, тобто всіх підмножин сусідніх куль. Кількість таких комбінацій зростає зі збільшенням  $d$  і  $k$ , що робить повний перебір обчислювально неприйнятним.

Зважаючи на те, що метою є отримання швидкого локального (а не глобально оптимального) розв'язку на кожному кроці, компромісом є дослідження лише частини крайніх точок. Ефективним способом реалізації цього підходу є використання мультістартової стратегії, коли задача (6) розв'язується засобами локальної оптимізації з декількох стартових точок. Це істотно зменшує ймовірність потрапляння у неякісну конфігурацію.

Слід також відзначити важливу особливість задачі (6): множини крайніх точок для послідовних куль  $k$  та  $k+1$  відрізняються лише кількома новими обмеженнями (додається умова відносно шойно розміщеної кулі). Ілюстрацію трьох положень кулі  $k=4$  для  $d=2$ , які відповідають крайнім точкам множини  $D^{(4)}$ , наведено на рис. 2. Зрозуміло, що якщо четверту кулю буде розміщено у будь-якому з положень, то для п'ятої дві крайні точки залишаться.

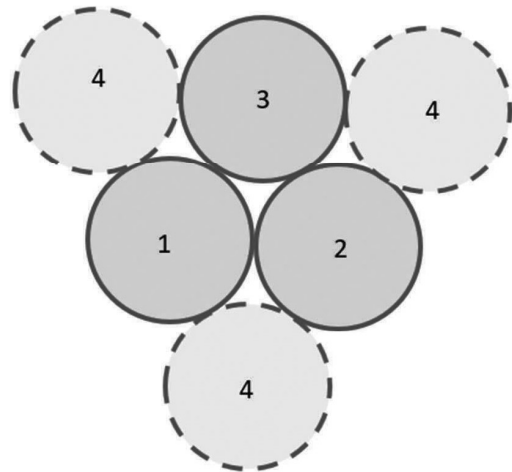


Рис. 2. Ілюстрація положень кулі  $k=4$  для  $d=2$ , які відповідають крайнім точкам множини  $D^{(4)}$

Тому для перших куль немає необхідності використовувати багато стартових точок, оскільки геометрія задачі ще відносно проста. Натомість для останніх куль, коли структура обмежень значно ускладнюється, роль мультістарту зростає: більша кількість стартових позицій дозволяє уникати небажаних локальних мінімумів та підвищити якість кінцевої конфігурації без значного зростання обчислювальних витрат.

### 3. Локалізація пошуку

Зі зростанням  $d$  і  $k$  кількість обмежень в локальній задачі (6) збільшується і повний перебір крайніх точок множини допустимих рішень  $D^{(k)}$  є обчислювально надто дорогим. Тому розроблено спеціальну процедуру, яка дає змогу уникнути розгляду великої кількості неактивних нерівностей.

Першу кулю розміщуємо на початку координат — це безумовний мінімум цільової функції цілі (7). Для вибору стартового положення решти куль використовується параметричний опис точок у евклідовому просторі  $\mathbb{R}^d$  у гіперсферичних (полярних, сферичних) координатах, який дає змогу рівномірно породжувати початкові точки центрів куль у кулі радіуса  $\rho^{(k)} = R^{(k)} - r$ ,

$$\begin{aligned} \rho^{(k)} &= R^{(k)} - r, \\ R^{(k)} &= \max_{i \in \{1, \dots, k-1\}} \|x_i\| + 4r. \end{aligned} \quad (10)$$

Загальна формула для координат кулі:

$$\begin{aligned} x_l^{(k)} &= \rho \cos \varphi_l \sin \varphi_1 \dots \sin \varphi_{l-1}, l = 1, \dots, d-1, \\ x_d^{(k)} &= \rho \sin \varphi_1 \dots \sin \varphi_{d-1}, \end{aligned} \quad (11)$$

де значення радіус-вектора та кутових координат вибирається випадково на таких сегментах:

$$\rho \in [r, \rho^{(k)}], \varphi_1, \dots, \varphi_{d-2} \in [0, \pi], \varphi_{d-1} \in [0, 2\pi]. \quad (12)$$

Якщо для згенерованої за формулами (10)–(12) стартової точки для  $k = 2, 3, \dots, N$ , порушується хоча б одна з умов задачі (6), процедура побудови початкового наближення повторюється. Враховуючи, що  $\rho^{(k)}$  обирається таким, щоб було вільне місце для розміщуваної кулі, така процедура не створює суттєвих обчислювальних витрат.

В результаті отримуємо точку  $x_k^0 \in D^{(k)}$ .

Наступним етапом є пошук межової точки  $D^{(k)}$ . У будь-якій точці простору  $x_k^{(0)} \in \mathbb{R}^d$ , окрім початку координат, антиградієнт цільової функції (7) буде спрямований до початку координат. Тому виконуємо оптимізацію уздовж вектора, який з'єднує центр розміщуваної кулі  $x_k^{(0)}$  та початок координат, тобто уздовж вектора  $-x_k^{(0)}$ . Пошук межової точки  $x_k^{(1)}$  множини допустимих розв'язків  $D^{(k)}$  здійснюється методом дихотомії.

Далі формується множина індексів куль  $I^{(k)}$ , які мають не порожній перетин з кулею

$$B(x_k^{(1)}, \tau) = \{x : \|x - x_k^{(1)}\| \leq \tau\}, \quad (12)$$

$$I^{(k)} = \{i \in \{1, \dots, k-1\} : \|x_i^* - x_k^{(1)}\| \leq \tau\}. \quad (13)$$

Тепер замість задачі (6) розв'язуємо задачу

$$x_k^* = \arg \min_{x \in \mathbb{R}^d} \left\{ \|x\|^2 : \|x - x_i\|^2 \geq 4r^2, i \in I^{(k)} \right\}, \quad (14)$$

яка з урахуванням (12), (13) враховує обмеження на накладання лише з кулями в безпосередній близькості до розміщуваної кулі.

Ілюстрацію оптимізаційного процесу для  $d = 2$  показано на рис. 3. Тут активними будуть обмеження для накладання розміщуваної кулі 8 лише з кулями 1, 3, 4, 5, 6, тобто  $I^8 = \{1, 3, 4, 5, 6\}$ .

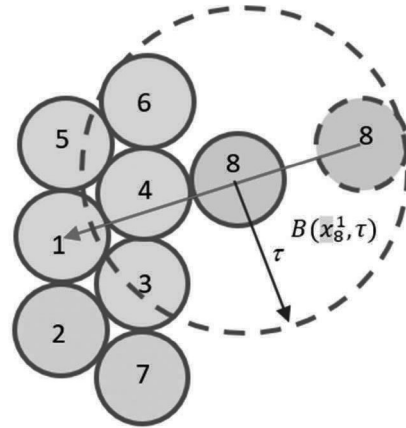


Рис. 3. Ілюстрація локального пошуку

### 4. Мультистартова стратегія

Задачі (6) та (14) є багатоекстремальними задачами нелінійного програмування з неопуклою множиною допустимих розв'язків. Тому локальний розв'язок значною мірою залежить від стартової точки. Щоб уникнути неякісних локальних екстремумів і не виконувати перебір усіх крайніх точок, доцільно застосувати мультистартову стратегію, яка виконує локальну оптимізацію для кулі  $k$  для набору стартових точок  $x_{k1}^{(0)}, x_{k2}^{(0)}, \dots, x_{kM_k}^{(0)}$ .

В результаті розв'язання задач (14) отримуємо відповідно точки  $x_{k1}^*, x_{k2}^*, \dots, x_{kM_k}^*$ . Вибираємо серед них той розв'язок, який має найменшу норму:

$$x_k^* = \arg \min_{l=1, \dots, M_k} \|x_{kl}^*\|. \quad (15)$$

Важливо зазначити, що множини активних обмежень для двох сусідніх кроків  $k$  та  $k+1$  відрізняються лише додаванням декількох обмежень. Це означає, що на ранніх етапах геометрія задачі є простою, і невелика кількість стартових точок достатня для знаходження якісного розв'язку. Натомість, на пізніх етапах, коли кількість крайніх точок значно збільшується, слід збільшувати і кількість стартових точок, щоб зменшити ризик потрапляння в поганий локальний мінімум та забезпечити краще положення нової кулі (15).

Такий адаптивний підхід зберігає обчислювальну ефективність та дає змогу суттєво поліпшити якість конфігурації без значного збільшення витрат.

### 5. Числові результати

У цьому розділі наведено результати числових експериментів, проведених для оцінювання роботи запропонованої мультистартової послідовної стратегії локалізованого пошуку. Мета експериментів — дослідити, наскільки мультистартова стратегія дає змогу

поліпшити якість отримуваних розміщень порівняно з одностартовим алгоритмом та як змінюється поведінка методу для різних розмірностей простору.

Для кожної розмірності  $d \in \{2, 3, 4, 8, 24, 32, 64\}$  було згенеровано конфігурації з  $N = 100$  куль однакового радіуса у кулі мінімального радіуса. Для кожного значення  $d$  обчислювалися такі характеристики: мінімальний контейнерний радіус  $R^*$ , визначений за формулою згортки (4); час виконання алгоритму; поліпшення  $R^*$  порівняно з одностартовим алгоритмом.

Для числової оптимізації локальних задач використовувався солвер IPOPT [11], який добре зарекомендував себе для задач великої розмірності з нелінійними обмеженнями.

Для  $d = 2$  розміщення куль можна візуально перевірити. На рисунку 4 наведено ілюстрацію розміщення 100 кругів за мультістартовою процедурою. На відміну від одностартового варіанта, конфігурація є більш симетричною, а кількість “випадкових вироджень” значно менша, що підтверджує доцільність мультістарту.

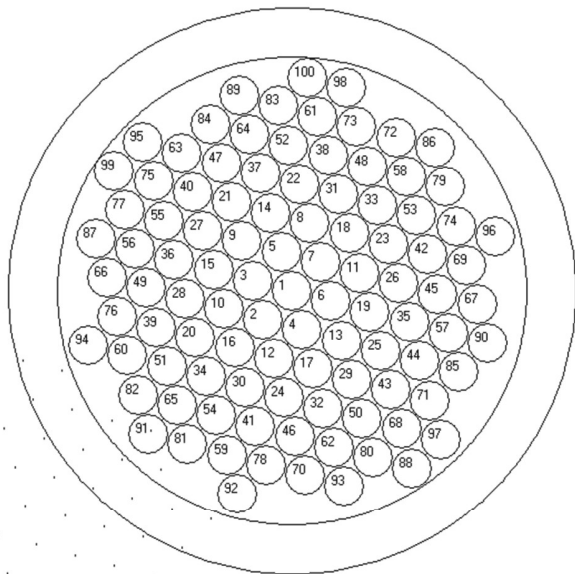


Рис. 4. Ілюстрація розміщення кругів за мультістартовою стратегією

Для  $d = 2, 3$  структура пакування наближається до фрагментів FCC/GCP-решіток, що можна спостерігати візуально.

Для більших розмірностей мультістарт дозволяє компенсувати частину геометричної деградації, притаманної високим вимірностям. Добре відомо, що для цих розмірностей існують щільні решіткові пакування D4, E8 та решітка Ліча L24. Хоча контейнерне обмеження не дає можливості відтворити решітку у первісному вигляді, мультістартова стратегія наближає розміщення до структур, подібних до фрагментів оптимальних решіток.

Для  $d = 32, 64$  експерименти демонструють робастність та масштабованість методу. контейнерний

радіус змінюється незначно, а час виконання зростає помірно завдяки локальній природі пошуку й адаптивній кількості стартових точок.

У таблиці 1 наведено порівняння результатів одностартового жадібного алгоритму  $R_{single}^*$  та запропонованої мультістартової стратегії  $R_{multy}^*$  для всіх тестованих розмірностей.

Таблиця 1

Характеристики конфігурацій багатовимірних куль

$d$	$R_{single}^*$	час, с	$R_{multy}^*$	час, с
2	13,1642	12	11,5829	55
3	6,7564	15	6,3265	71
4	4,8360	18	4,7654	89
8	3,0019	30	3,0011	154
24	2,9976	59	2,9972	295
32	2,9972	86	2,9970	437
64	3,0080	180	2,9966	907

Мультістартова стратегія дає менший радіус контейнера для більшості протестованих значень  $d$ . Розміщення стають геометрично більш регулярними; структура має більше «центрування» та менше локальних артефактів. Часу виконання алгоритму оцінюється як  $O(N^2 d M_k)$ .

## Висновки

У роботі запропоновано та досліджено мультістартову послідовну стратегію розміщення неперетинних гіперкуль у багатовимірних просторах усередині контейнерної кулі мінімального можливого радіуса. Запропонований підхід ґрунтується на редукції вихідної min–max задачі до послідовності локальних задач мінімізації норми, що природним чином впливає зі згортки параметра контейнера. Така конструкція дозволяє уникнути високої розмірності початкової постановки та зберегти її геометричну інтерпретацію.

Введено поліпшений механізм локалізації пошуку, який обмежує оптимізацію лише тими обмеженнями, що можуть стати активними, та використовує адаптивну оцінку максимально можливої області розміщення для нової кулі. Це суттєво зменшує розмірність локальних задач та дозволяє ефективно працювати у просторах великої розмірності.

Ключовим елементом методу є мультістартова стратегія, що передбачає розв’язання локальної задачі з кількох стартових точок. Такий підхід компенсує залежність від початкових умов, характерну для одностартових жадібних методів, і дає змогу значно знизити ймовірність потрапляння у небажаний локальний мінімум. Особливо це актуально в пізніх кроках алгоритму, коли кількість крайніх точок значно збільшується. Адаптивне збільшення кількості стартових точок для останніх куль забезпечує поліпшення

розташування конфігурації без суттєвого зростання обчислювальної складності.

Числові експерименти підтверджують ефективність мультістартового підходу порівняно з одностартовим алгоритмом. Крім того, мультістарт дозволяє формувати геометрично більш регулярні конфігурації та демонструє стійкість результатів навіть у високих розмірностях. При цьому час виконання алгоритму зростає помірно завдяки локальному характеру оптимізації та обмеженню кількості активних обмежень.

Таким чином, запропонована мультістартова стратегія є ефективним, масштабованим та обчислювально недорогим підходом для побудови допустимих розміщень багатовимірних куль у задачах великої розмірності. Метод може використовуватися як самостійний інструмент формування конфігурацій або як початковий етап для подальших локальних чи глобальних методів удосконалення.

#### Список літератури.

- [1] Conway J. H., Sloane N. J. A. (1999). Sphere Packings, Lattices and Groups, Springer New York, NY. <https://doi.org/10.1007/978-1-4757-6568-7>. Стоян Ю.Г., Яськов Г.М., Романова Т.Є., Яковлев С.В. (2021). Пакування сферичних об'єктів: моделі, методи, застосування. Київ: Наукова думка, 279 с.
- [2] Litvinchev I., Fischer A., Romanova T., Stetsyuk P. (2024). A New Class of Irregular Packing Problems Reducible to Sphere Packing in Arbitrary Norms. *Mathematics* 12(7), 935. <https://doi.org/10.3390/math12070935>.
- [3] Torquato S., Uche O. U., Stillinger F. H. (2006). Random sequential addition of hard spheres in high
- [4] Euclidean dimensions. *Phys. Rev. E* 74(6), 061308. <https://doi.org/10.1103/PhysRevE.74.061308>.
- [5] Ryan B. Jadrich, Beth A. Lindquist, Thomas M. Truskett (2022). Treating random sequential addition via the replica method. *J. Chem. Phys.* 157 (8), 084116. <https://doi.org/10.1063/5.0096276>.
- [6] Cohn H., Kumar A., Miller S., Radchenko D., Viazovska M. (2022). Universal optimality of the  $E_8$  and Leech lattices and interpolation formulas. *Ann. of Math. (2)* 196 (3), pp. 983–1082. <https://doi.org/10.4007/annals.2022.196.3.3>.
- [7] Chen B., et al. (2023). Geometrically-Shaped Multi-Dimensional Modulation Formats in Coherent Optical Transmission Systems. *Journal of Lightwave Technolog.* 41(3), pp. 897–910. <https://doi.org/10.1109/JLT.2022.3204101>.
- [8] Favano A., Ferrari M., Magarini M., Barletta L. (2023). A Sphere Packing Bound for Vector Gaussian Fading Channels Under Peak Amplitude Constraints. *IEEE Transactions on Information Theory.* 69(1), pp. 238–250. <https://doi.org/10.1109/TIT.2022.3203293>.
- [9] Chen B. et al. (2025). On Shaping Gain of Multidimensional Constellations in Linear and Nonlinear Optical Fiber Channel, *IEEE Journal on Selected Areas in Communications* 43(5), pp. 1455–1468. <https://doi.org/10.1109/JSAC.2025.3543507>.
- [10] Wang X., Fujisawa M., Mikawa M. (2023). XProtoSphere: an eXtended multi-sized sphere packing algorithm driven by particle size distribution. *Vis Comput* 39, pp. 3333–3346. <https://doi.org/10.1007/s00371-023-02977-w>.
- [11] Andrei, N. (2022). Interior-Point Filter Line-Search. In: *Modern Numerical Nonlinear Optimization*. Springer Optimization and Its Applications, vol 195, pp. 661–678. Springer, Cham. [https://doi.org/10.1007/978-3-031-08720-2\\_19](https://doi.org/10.1007/978-3-031-08720-2_19).

*Received (Надійшла) 22.12.2025*

*Accepted for publication (Прийнята до друку) 20.01.2026*

*Publication date (Дата публікації) 27.03.2026*

# ПРАВИЛА оформлення рукописів для авторів науково-технічного журналу «БІОНІКА ІНТЕЛЕКТУ»

Науково-технічний журнал «Біоніка інтелекту» приймає до друку написані спеціально для нього оригінальні рукописи, які раніше ніде не друкувались. Структура рукопису повинна бути такою: індекс УДК, відомості про авторів, заголовок, анотації (на трьох мовах), ключові слова, вступ, основний текст статті, висновки, список використаної літератури, резюме.

Відповідно до Постанови ВАК України від 15.01.2003 №7-05/1 (Бюлетень ВАК, №1, 2003, с. 2), стаття повинна мати такі необхідні елементи: постановка проблеми у загальному вигляді та її зв'язок із важливими науковими чи практичними завданнями; аналіз останніх досліджень і публікацій і виділення не вирішених раніше частин загальної проблеми в даній області; формулювання цілей та завдань дослідження; виклад основного матеріалу досліджень з повним обґрунтуванням отриманих наукових результатів; висновки з даного дослідження та перспективи подальших досліджень у даному напрямку.

Статті мають бути виконані в редакторі Microsoft Word. Формат сторінки — А4 (210×297 мм), поля: верхнє — 25 мм, нижнє — 20 мм, ліве, праве — 17 мм. Кількість колонок — 2, з інтервалом між ними 5 мм, основний шрифт Times New Roman, кегль основного тексту — 10 пунктів, міжрядковий інтервал — множник (1,1), абзацний відступ — 6 мм. Обсяг рукопису — від 6 до 12 сторінок (мови: українська, англійська, російська та мовою оригіналу).

*УДК* друкується з першого рядка, без відступів, вирівнювання по лівому краю.

*ПІБ автора* (-ів), назва статті, назва та адреса учбового закладу необхідно надати повністю російською, українською та англійською мовами.

*Назва статті* друкується прописними літерами; шрифт прямий, напівжирний, кегль 12.

*Назви розділів* нумерують арабськими цифрами, виділяють жирним шрифтом. Відступи для назви статті, ініціалів та прізвищ авторів, відомостей про авторів, назв розділів, вступу та висновків, списку літератури: зверху — 6 пт, знизу — 3 пт.

*Анотації* (мовою статті, абзац 6—12 рядків, кегль 9) розміщують на початку статті, в ній має бути розміщена інформація про очікувані результати описаних досліджень (на трьох мовах).

*Ключові слова* (4—10 слів з тексту статті, які з точки зору інформаційного пошуку несуть змістовне навантаження) наводять мовою рукопису, через кому в називному відмінку, кегль 9.

*Рисунки та таблиці* (чорно-білі, контрастні) розміщуються у тексті після першого посилання у вигляді окремих об'єктів і нумерують арабськими цифрами наскрізною нумерацією за наявності більше ніж одного об'єкта. Невеликі схеми, що складаються з 3—4 елементів виконують, використовуючи вставку об'єкта Рисунок Microsoft Word. Більш складні виконують у графічних редакторах у вигляді чорно-білих графічних файлів форматів .tif, .jpg, .wmf, .cdr із

розділенням 300 dpi. Рисунки мають міститися у текстовому файлі й обов'язково подаватися окремими файлами з відповідними назвами (наприклад, рис1.jpg).

Усі елементи рисунка, включаючи написи, повинні бути згруповані. Усі написи в рисунках і таблицях мають бути виконані шрифтом Times New Roman, кегль у рисунках — 10, у таблицях — 9.

Рисунок повинен мати центрований підпис (поза рисунком), шрифт 9, відступи зверху і знизу по 6 пт. Ширина рисунка має відповідати ширині колонки (або ширині сторінки).

*Формули, символи, змінні* повинні бути набрані в редакторі формул **MathType**. Формули розміщують посередині рядка й нумерують за наявності посилань на них у рукописі. Шрифт — Times New Roman. Висота змінної — 10 пунктів, великих і малих індексів — 8 пт, основний математичний символ — 12 (10) пт. Змінні, позначені латинськими літерами, набирають курсивом, грецькі літери, скорочення російських слів і цифри — прямим написанням. Змінні, які є в тексті, також набирають у редакторі формул.

*Список літератури* вміщує опубліковані джерела, на які є посилання в тексті, укладені у квадратні дужки, друкують без абзацного відступу, кегль 9 пт, відступ зверху — 6 пт.

Після списку літератури з відступом зверху 6 пт зазначають *дату подання статті до редколегії*. Число та місяць задають двозначними числами через крапку. Розмір шрифту — 9 пт, курсив, вирівнювання по правому краю.

*Резюме* (Times New Roman, кегль — 10 пунктів,) подають англійською мовою: обсяг резюме до 2000 знаків (бажаний переклад). *Структура резюме: Background, Materials and methods, Results, Conclusion.*

Разом із рукописом (на аркушах білого паперу формату А4 щільністю 80-90 г/м<sup>2</sup>, надрукований на лазерному принтері) необхідно подати такі документи:

1. Заяву, яку повинні підписати всі автори.
2. Акт експертизи про можливість опублікування матеріалів у відкритому друці (якщо потрібно).
3. Рецензію, підписану доктором чи кандидатом наук.
4. Відомості про авторів.
5. Електронний варіант рукопису, резюме та відомостей про авторів.
6. Зробити оплату публікації.

Необхідно також зазначити один з наступних тематичних розділів, якому відповідає рукопис:

1. Теоретичні основи інформатики та кібернетики. Теорія інтелекту.
2. Математичне моделювання. Системний аналіз. Прийняття рішень.
3. Інтелектуальна обробка інформації. Розпізнавання образів.
4. Інформаційні технології та програмно-технічні комплекси.
5. Структурна, прикладна та математична лінгвістика.
6. Дискусійні повідомлення.

## ЗМІСТ

### **СТРУКТУРНА, ПРИКЛАДНА ТА МАТЕМАТИЧНА ЛІНГВІСТИКА**

<i>Воржевітіна А. І., Шанідзе В. Д.</i> Формування системи ознак для моделювання лінгвістичних структур на основі методу компарації за засобів теорії інтелекту .....	3
<i>Юрченко О. М., Повесьма Ю. Ю.</i> Семантичний аналіз контекстів різної довжини за метрикою лексичної різноманітності MTLD.....	10

### **ІНЖЕНЕРІЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ**

<i>Гринченко М. А., Лісова Д. В., Куценко Д. О.</i> Проектування рекомендаційної системи підтримки прийняття клінічних рішень для сімейного лікаря.....	19
<i>Шапошніков М. І.</i> Інформаційна система планування розподілу ресурсів для підвищення рейтингу закладу вищої освіти .....	30

### **ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ. МАШИННЕ НАВЧАННЯ. БАЗИ ДАНИХ**

<i>Vukhanovskyi V., Ryabova N.</i> Fuzzy-adapters: Integration of fuzzy membership functions for Vision Transformers' effective transfer learning.....	39
<i>Федоричко А. С., Висоцька В. А., Линник Р. О.</i> Інформаційна технологія класифікації текстів для виявлення ознак ПТСР на основі технологій глибинного навчання та NLP.....	45
<i>Кириченко І. В., Терещенко Г. Ю., Лазаренко О. В.</i> Дослідження інструментарію статистичного моделювання у контексті цифрової маркетингової аналітики .....	56
<i>Гулів Н. Б., Назаров О. С.</i> Дослідження способів усунення дисбалансу класів серед медичних та психологічних даних у побудові алгоритму випадковий ліс.....	69
<i>Данієль Я., Максименко Д., Панченко Д., Калиниченко О., Турута О.</i> Набір даних українських новин як бенчмарк для класифікації текстів.....	75

### **ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ ТА ОБРОБКА ДАНИХ**

<i>Лециньська І. О.</i> Метод побудови нейросимвольного представлення ментальної моделі рішення інтелектуальної системи.....	81
<i>Чалій С. Ф., Кравченко Р. В.</i> Метод побудови темпорально узгоджених пояснень в інтелектуальних системах на основі темпоральних графових нейронних мереж.....	88
<i>Редозуб І. А., Хайрова Н. Ф.</i> Оцінювання здатності моделей генеративного штучного інтелекту виявляти етичні виміри публічного дискурсу: кейс міграції українських біженців .....	95

### **МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ. СИСТЕМНИЙ АНАЛІЗ ТА УПРАВЛІННЯ**

<i>Богатов О. І., Плехова Г. А., Костікова М. В., Неронов С. М., Плехов Д. О.</i> Визначення обмежень у використанні статистичних оцінок професійного ризику за офіційними даними в Україні.....	104
<i>Широков М. С., Олевський В. І.</i> Розпізнавання підземних порожнин у ґрунті: класичні методи та сучасні технології .....	109
<i>Чугай А. М., Яськов Г. М., Старкова О. В., Яськова Є. Г.</i> Гібридний алгоритм просторового пакування для моделювання біомедичних систем.....	118
<i>Веретельник К. О.</i> Послідовний мультистартовий алгоритм розміщення багатовимірних куль.....	125

### **ПРАВИЛА**

оформлення рукописів для авторів науково-технічного журналу «БІОНІКА ІНТЕЛЕКТУ» .....	131
---	-----

# CONTENTS

## **STRUCTURAL, APPLIED AND MATHEMATICAL LINGUISTICS**

<i>Vorzhevitina A. I., Shanidze V. D.</i> Formation of a system of features for modelling linguistic structures based on the method of comparison using the tools of the theory of intelligence .....	3
<i>Yurchenko O. M., Povesma Yu. Yu.</i> Semantic analysis of contexts of varying lengths using the MTLD lexical diversity metric .....	10

## **SOFTWARE ENGINEERING**

<i>Grinchenko M. A., Lisova D. V., Kutsenko D. O.</i> Conceptual design of a clinical decision support recommender system for family physicians .....	19
<i>Shaposhnikov M. I.</i> An information system for resource allocation planning to increase the ranking of a higher education institution .....	30

## **INFORMATION TECHNOLOGIES. MACHINE LEARNING. DATABASES**

<i>Bukhanovskiy V., Ryabova N.</i> Fuzzy-adapters: Integration of fuzzy membership functions for Vision Transformers' effective transfer learning .....	39
<i>Fedorychko A. S., Vysotska V. A., Lynnyk R. O.</i> Information technology for text classification of PTSD signs detection based on deep learning and NLP .....	45
<i>Kyrychenko I.V., Tereshchenko G. Yu., Lazarenko O.V.</i> A Study of Statistical Modelling Tools in the Context of Digital Marketing Analytics.....	56
<i>Huliiev N. B., Nazarov O. S.</i> Research on ways to eliminate class imbalance among medical and psychological data in building a random forest algorithm.....	69
<i>Daniil Y., Maksymenko D., Panchenko D., Kalynychenko O., Turuta O.</i> Text Classification Benchmark of Ukrainian News Dataset .....	75

## **INTELLIGENT SYSTEMS AND DATA PROCESSING**

<i>Leshchynska I. O.</i> Generalised symbolic mental model of an intelligent system decision for external users .....	81
<i>Chalyi S. F., Kravchenko R. V.</i> Method for constructing temporally consistent explanations in intelligent systems based on temporal graph neural networks .....	88
<i>Redozub I. A., Khairova N. F.</i> Evaluating the Capability of GenAI Models to Identify Ethical Dimensions of Public Discourse: The Case of Ukrainian Refugees.....	95

## **MATHEMATICAL MODELLING. SYSTEM ANALYSIS AND CONTROL**

<i>Bogatov O. I., Pliekhova G. A., Kostikova M. V., Neronov S. M., Pliekhov D. O.</i> Identifying limitations in the use of statistical estimates of occupational risk based on official data in Ukraine .....	104
<i>Shyrovkov M.S., Olevskiy V.I.</i> Recognition of underground volatile space: classic methods and modern technologies .....	109
<i>Chuhai A. M., Yaskov G. M., Starkova O. V., Yaskova Y. G.</i> Hybrid spatial packing algorithm for biomedical system modelling. ....	118
<i>Veretelnyk K. O.</i> Sequential multistart algorithm for placing multidimensional spheres .....	125

## **AUTHOR GUIDELINES**

Manuscript Preparation for the Scientific and Technical Journal “Bionics of Intelligence” .....	131
---	-----



**SCIENTIFIC AND TECHNICAL JOURNAL  
"BIONICS OF INTELLIGENCE"**

**Chief Editor**

*G.G. Chetverikov* — Doctor of Technical Sciences, Professor; Director of the Interregional Center for Applied and Mathematical Linguistics of the Ukrainian Language-Information Foundation of the National Academy of Sciences of Ukraine (Kyiv) and Kharkiv National University of Radio Electronics (KhNURE).

**Editorial Board**

1. Igor Aizenberg (USA) — Professor, Chair of the Department of Computer Science, M.Sc. in Mathematics, Ph.D. in Computer Science, Manhattan College (Riverdale, NY, USA).
2. A.A. Bezsonov — Doctor of Technical Sciences, Professor.
3. V.O. Gorokhovatsky — Doctor of Technical Sciences, Professor, KhNURE.
4. I.V. Grebennik — Doctor of Technical Sciences, Professor, Professor, KhNURE.
5. Z.V. Dudar — Candidate of Technical Sciences, Professor, KhNURE.
6. A.L. Erokhin — Doctor of Technical Sciences, Professor, KhNURE.
7. K.E. Petrov — Doctor of Technical Sciences, Professor, KhNURE.
8. I.A. Revenchuk — Candidate of Technical Sciences, Associate Professor, KhNURE.
9. O.G. Rudenko — Doctor of Technical Sciences, Professor, KhNURE.
10. S.G. Udovenko — Doctor of Technical Sciences, Professor, Simon Kuznets Kharkiv National University of Economics.
11. Shafagat Mahmudova (Azerbaijan) — Ph.D. in Technical Sciences, Associate Professor, Institute of Information Technology, Azerbaijan National Academy of Sciences, Department of Biometric Technology and Software Engineering (Baku, Azerbaijan).
12. N.V. Sharonova — Doctor of Technical Sciences, Professor, National Technical University "Kharkiv Polytechnic Institute".
13. V.A. Shyrovok — Doctor of Technical Sciences, Professor; Director of the Ukrainian Language-Information Foundation of the National Academy of Sciences of Ukraine; Academician of the National Academy of Sciences of Ukraine.
14. O.A. Kobylin — Head of the Department of Informatics, Candidate of Technical Sciences, Associate Professor, KhNURE.
15. S.V. Mashtalir — Doctor of Technical Sciences, Professor, KhNURE.
16. O.V. Nechiporenko — Doctor of Technical Sciences, Professor, KhNURE.

**НАУКОВО-ТЕХНІЧНИЙ ЖУРНАЛ  
«БІОНІКА ІНТЕЛЕКТУ»**

**Головний редактор**

*Г. Г. Четвериков* — доктор технічних наук, професор; директор Межрегіонального центру прикладної і математичної лінгвістики Українського мовно-інформаційного фонду НАН України (м. Київ) і Харківського національного університету радіоелектроніки (ХНУРЕ).

**Редакційна колегія**

1. Ігор Айзенберг (США) — професор, завідувач кафедри комп'ютерних наук, магістр математики, PhD з комп'ютерних наук, Manhattan College (Рівердейл, Нью-Йорк, США).
2. А. А. Безсонов — доктор технічних наук, професор.
3. В. О. Гороховатський — доктор технічних наук, професор, ХНУРЕ.
4. І. В. Гребеннік — доктор технічних наук, професор, ХНУРЕ.
5. З. В. Дудар — кандидат технічних наук, професор, ХНУРЕ.
6. А. Л. Єрохін — доктор технічних наук, професор, ХНУРЕ.
7. К. Е. Петров — доктор технічних наук, професор, ХНУРЕ.
8. І. А. Ревенчук — кандидат технічних наук, доцент, ХНУРЕ.
9. О. Г. Руденко — доктор технічних наук, професор, ХНУРЕ.
10. С. Г. Удовенко — доктор технічних наук, професор, Харківський національний економічний університет імені Семена Кузнеця.
11. Шагафат Махмудова (Азербайджан) — кандидат технічних наук, доцент, Інститут інформаційних технологій Національної академії наук Азербайджану, відділ біометричних технологій та інженерії програмного забезпечення (м. Баку, Азербайджан).
12. Н. В. Шаронова — доктор технічних наук, професор, Національний технічний університет "Харківський політехнічний інститут".
13. В. А. Широков — доктор технічних наук, професор; директор Українського мовно-інформаційного фонду НАН України; академік НАН України.
14. О. А. Кобилін — завідувач кафедри інформатики, кандидат технічних наук, доцент, ХНУРЕ.
15. С. В. Машталір — доктор технічних наук, професор, ХНУРЕ.
16. О. В. Нечипоренко — доктор технічних наук, професор, ХНУРЕ.

Журнал включено до Переліку наукових фахових видань України, в яких можуть публікуватися результати дисертаційних робіт на здобуття наукових ступенів доктора наук, кандидата наук та ступеня доктора філософії, згідно з наказом Міністерства освіти і науки України № 820 від 11.07.2016 (внесено зміни згідно з наказом МОНУ № 920 від 26.06.2024)

## MAGAZINE TOPICS

---

- algebraization of logic
- database and knowledge base
- genetic engineering and molecular biology
- games and machine creativity
- hierarchical and dynamic systems
- intelligent agents and robots
- linguistic technologies
- logical networks
- mathematical psychology
- mathematical models  
and computational methods
- machine learning and evolution
- informatization methodology
- brain-like structures
- nanotechnology
- neural networks and neuromathematics
- object modeling
- ontological systems
- parallel information processing
- information search
- multi-criteria decision  
making
- speech recognition and synthesis
- распознавание образов
- system analysis
- structural, applied  
and mathematical linguistics
- intelligence theory
- informatization philosophy
- expert systems
- wavelet theory
- management

## ТЕМАТИКА ЖУРНАЛА

---

- алгебраїзація логіки
- бази даних та знання
- генна інженерія та молекулярна біологія
- ігри та машинна творчість
- ієрархічні та динамічні системи
- інтелектуальні агенти та роботи
- лінгвістичні технології
- логічні мережі
- математична психологія
- математичні моделі  
та обчислювальні методи
- машинне навчання та еволюція
- методологія інформатизації
- мозкоподібні структури
- нанотехнології
- нейронні мережі та нейроматематика
- об'єктне моделювання
- онтологічні системи
- паралельна обробка інформації
- пошук інформації
- прийняття багатокритеріальних  
рішень
- розпізнавання та синтез мови
- розпізнавання образів
- системний аналіз
- структурна, прикладна  
та математична лінгвістика
- теорія інтелекту
- філософія інформатизації
- експертні системи
- теорія вейвлетів
- менеджмент

*Наукове видання*

**БІОНІКА ІНТЕЛЕКТУ**  
**інформація, мова, інтелект**

**Науково-технічний журнал**

**№ 1 (104)**  
**2026**

Головний редактор — *Г. Г. Четвериков*  
Відповідальний редактор — *І. В. Кириченко*

Комп'ютерна верстка — *О. Б. Ісаєва*

Ідентифікатор медіа R30-03875

Витяг з реєстру суб'єктів у сфері медіа — реєстрантів від 25.04.2024 No 1410

Затверджений до друку Науково-технічною радою  
Харківського національного університету радіоелектроніки  
(протокол No 4 від 19 березня 2026 року).

Адреса редакції:  
Україна, 61166, Харків-166, просп. Науки, 14,  
Харківський національний університет радіоелектроніки, к. 127  
тел. 702-14-77, факс 702-10-13,  
e-mail: [bionics@nure.ua](mailto:bionics@nure.ua)  
Інформаційний сайт: <https://bionics.nure.ua>

---

Підписано до друку 25.03.2026. Формат 60 × 84 1/8. Друк ризографічний.  
Папір офсетний. Гарнітура Newton. Умов. друк. арк. 15,6. Обл.-вид. арк. 15,4.  
Тираж 20 прим.

Віддруковано в редакційно-видавничому відділі ХНУРЕ  
61166, Харків, просп. Науки, 14.