



І.Ю. Шубін¹, Г.Г. Четвериков², В.А. Ляшик³, Н.О. Шанидзе⁴

¹Кандидат технічних наук, професор кафедри програмної інженерії,
Харківський національний університет радіоелектроніки,
igor.shubin@nure.ua, ORCID iD: 0000-0002-1073-023X

²Доктор технічних наук, професор кафедри програмної інженерії,
Харківський національний університет радіоелектроніки,
grirorij.chetverykov@nure.ua, ORCID iD: 0000-0001-5293-5842

³Аспірант кафедри програмної інженерії,
Харківський національний університет радіоелектроніки,
volodymyr.liashyk@nure.ua, ORCID iD: 0000-0001-7326-0813

⁴Кандидат соціологічних наук, доцент кафедри соціології та політології,
Національний технічний університет «Харківський політехнічний інститут»,
nashanidze@ukr.net ORCID iD 0000-0002-9613-186X

МЕТОДИ ШТУЧНИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ АДАПТИВНОГО ТЕСТУВАННЯ ЗНАТЬ

Під адаптивним тестовим контролем розуміють комп'ютеризовану систему науково обґрунтованої перевірки й оцінювання результатів навчання, що має високу ефективність за рахунок оптимізації процедур генерації, пред'явлення й оцінки результатів виконання адаптивних тестів, що заснована на методах побудови та оптимізації логічних мереж. Алгоритми підбору й пред'явлення завдань будуються за принципом зворотного зв'язку, коли при правильній відповіді суб'єкта навчання чергове завдання вибирається більш важким, а невірна відповідь спричиняє пред'явлення наступного більш легкого завдання, ніж те, на яке суб'єктом навчання була дана невірна відповідь. Також є можливість завдання додаткових питань по темах, які суб'єкт навчання знає не дуже добре для більш точного з'ясування рівня знань у даних областях. Вибір алгоритмів тестування наразі фактично обмежений формами представлення тестових завдань і алгоритмами оцінювання результатів тестування. Досягнення більш високих результатів і підвищення мотивації навчання в остаточному підсумку є основною метою тестування знань. Для визначення базового алгоритму, необхідно навести сценарій роботи системи. У його основі лежить модель приймання іспиту викладачем у студента, як модель адаптивного тестування. Такий вибір сценарію роботи системи обумовлений тим, що, по-перше, дана процедура історично добре формалізована, по-друге, при проектуванні тестів, їх розробнику необхідно спиратися на загальноприйняті, відомі й використовувані їм методи з мінімальною модифікацією.

ДИСТАНЦІЙНА ОСВІТА, ЛОГІЧНА МЕРЕЖА, АЛГЕБРА СКІНЧЕННИХ ПРЕДИКАТИВ, ДИСТАНЦІЙНЕ ТЕСТУВАННЯ ЗНАТЬ, МОДЕЛЬ СУБ'ЄКТА НАВЧАННЯ

Шубин И.Ю., Четвериков Г.Г., Ляшик В.А., Шанидзе Н.А. Методы искусственных нейронных сетей для адаптивного тестирования знаний. Под адаптивным тестовым контролем понимают компьютеризированную систему научно обоснованной проверки и оценки результатов обучения, имеет высокую эффективность за счет оптимизации процедур генерации, предъявления и оценки результатов выполнения адаптивных тестов, основанная на методах построения и оптимизации логических сетей. Алгоритмы подбора и предъявления задач строятся по принципу обратной связи, когда при правильном ответе субъекта обучения очередное задание выбирается более тяжелым, а неверный ответ вызывает предъявления следующего более легкого задания, чем то, на которое субъектом обучения была дана неверная ответ. Также имеется возможность задания дополнительных вопросов по темам, которые субъект обучения знает не очень хорошо для более точного выяснения уровня знаний в данных областях. Выбор алгоритмов тестирования пока фактически ограничен формами представления тестовых заданий и алгоритмами оценки результатов тестирования. Достигновения более высоких результатов и повышения мотивации обучения в конечном итоге является основной целью тестирования знаний. Для определения базового алгоритма, необходимо привести сценарий работы системы. В его основе лежит модель принятия экзамена преподавателем у студента, как модель адаптивного тестирования. Такой выбор сценария работы системы обусловлен тем, что, во-первых, данная процедура исторически хорошо формализована, во-вторых, при проектировании тестов, их разработчику необходимо опираться на общепринятое, известные и используемые им методы с минимальной модификацией

ДИСТАНЦІОННЕ ОБРАЗОВАНИЕ, ЛОГІЧЕСКАЯ СЕТЬ, АЛГЕБРА КОНЕЧНЫХ ПРЕДИКАТОВ, ДИСТАНЦИОННОЕ ТЕСТИРОВАНИЕ ЗНАТЬ, МОДЕЛЬ ОБУЧАЕМОГО

Shubin I.Yu., Chetverikov G.G., Liashyk V.A., Shanidze N.A. Methods Of Artificial Neural Networks for Adaptive Knowledge Testing. Adaptive test control is a computerized system of scientifically based verification and evaluation of learning outcomes, which is highly effective by optimizing the procedures for generating, presenting and evaluating the results of adaptive tests, based on methods of building and optimizing logical networks. Algorithms for selection and presentation of tasks are based on the principle of feedback, when the correct answer of the subject of training is the next difficult task, and the wrong answer causes the presentation of the next easier task than that to which the subject

of training the wrong answer was given. It is also possible to ask additional questions on topics that the subject does not know very well to clarify the level of knowledge in these areas. The choice of testing algorithms is currently actually limited by the forms of presentation of test tasks and algorithms for evaluating test results. Achieving higher results and increasing the motivation to learn is ultimately the main goal of testing knowledge. To determine the basic algorithm, it is necessary to provide a scenario of the system. It is based on the model of taking the exam by a teacher as a model of adaptive testing. This choice of the scenario of the system is due to the fact that, firstly, this procedure is historically well formalized, and secondly, when designing tests, their developer must rely on common, known and used methods with minimal modification.

DISTANCE EDUCATION, LOGICAL NETWORK, ALGEBRA OF FINITE PREDICATES, DISTANCE TESTING OF KNOWLEDGE, MODEL OF SUBJECT OF LEARNING

Вступ

Аналіз методів педагогічних вимірювань дозволяє зробити висновок про те, що одним з об'єктивних і ефективних методів контролю якості знань є тестовий метод, заснований на використанні педагогічних тестових матеріалів. У цей час важко назвати дисципліну, процес вивчення якої так чи інакше не містить в собі контроль знань в тестовій формі.

Дослідження показали, що, незважаючи на велику кількість розроблених тестів, вони мають певні недоліки, серед яких необ'єктивність вагових коефіцієнтів тестових завдань, неоптимальна кількість тестових завдань або одновариантність тесту, наявність зв'язку між послідовними завданнями.

На практиці дуже часто виникає ситуація, коли одна й та сама група студентів без особливих зусиль справляється з усіма тестовими завданнями або, навпаки, не може впоратися із більшістю тестових завдань.

Таким чином, існує проблема коректності підбору складності тестових завдань з метою найбільш адекватної оцінки рівня знань студентів. В зв'язку зі збільшенням кількості й недостатньою якістю тестів, що застосовуються при навчанні у вузах, не завжди є можливим якісно визначити рівень навчальних досягнень студента, ґрунтуючись тільки на тестах, в яких кількість завдань є фіксованою (так звані тести фіксованої довжини). Цей недолік можливо усунути за допомогою застосування такого виду тестування, яке здатне «підлаштовуватися» під рівень знань студентів, варіювати складність і кількість завдань в залежності від правильності відповідей на них. Таким чином, процес тестування адаптується до рівня знань студента, що проходить тест. Це дозволяє отримати більш достовірні результати, скоротити час, що потрібен для проходження тесту, не знижує мотивацію студентів до навчання й тестування. Подібні підходи прийнято називати адаптивним тестуванням.

При комп'ютерному адаптивному тестуванні тестові завдання формуються індивідуально для кожного студента, що екзамується, з урахуванням результатів виконання попередніх завдань. Типи завдань, їх кількість та порядок проходження індивідуальні. Таким чином, адаптивне тестування не тільки дає більш об'єктивну оцінку знанням, умінням і навичкам студентів, але й дозволяє виявляти, які знання є

помилковими або неповними, а також дозволяє формувати подальшу траєкторію навчання.

Проблеми вимірів і діагностики є найбільш складними як у теоретичному, так і у практичному відношенні. Під «виміром» зазвичай розуміють виявлення кількісних характеристик досліджуваних явищ або особливу процедуру, за допомогою якої числа або, принаймні, порядкові величини приписуються об'єктам за певними правилами. Сюди слід включати й математичну обробку результатів вимірів. При цьому в якості результату виміру виступає значення величини, що характеризує якість, знайдену шляхом її виміру.

Основна проблема полягає в тому, що часто неможливо безпосередньо виміряти потрібну величину, наприклад рівень знань студентів з певної дисципліни. Такі величини називають латентними. Для визначення латентної змінної використовують вимір інших змінних, на основі яких обчислюють потрібну величину.

Педагогічний вимір має свої рівні (оцінка результатів педагогічних вимірів; аналіз і узагальнення педагогічних явищ; ефективність засобів і методів навчання; оцінка досягнення цілей на рівні глобальних цілей, на рівні навчального предмета, на рівні навчального матеріалу).

До основних проблем педагогічних вимірів слід віднести можливу необ'єктивність, неоднозначність, латентність, забезпечення точності, стійкості та валідності [1].

Основним предметом педагогічних вимірів є розробка якісних тестів для виміру рівня підготовленості студентів, як найбільш точного засобу вимірювань. Ефективність організації та проведення педагогічних тестів підвищується за рахунок використання комп'ютерних технологій.

1. Постановка задачі

Проблеми вимірів і діагностики є найбільш складними як у теоретичному, так і у практичному відношенні. Під «виміром» зазвичай розуміють виявлення кількісних характеристик досліджуваних явищ або особливу процедуру, за допомогою якої числа або, принаймні, порядкові величини приписуються об'єктам за певними правилами. Сюди слід включати й математичну обробку результатів вимірів. При

цьому в якості результату виміру виступає значення величини, що характеризує якість, знайдену шляхом її виміру.

Основною ідеєю класичної теорії тестів є припущення про існування істинного бала (true score). Нерідко в одномірних вимірах дійсний бал називають параметром студента. При цьому передбачається, що кожному студентові можна поставити у відповідність єдине на момент виміру значення параметра, що не залежить від застосовуваного тесту. У цілому істинний бал – це ідеалізована константа випробуваного в гіпотетичній генеральній сукупності завдань нескінченного тесту.

Крім припущення про існування істинного бала в класичній теорії тестів виділяють кілька постулатів, що дозволяють побудувати математико-статистичний апарат для розробки науково обґрунтованих тестів і оцінки якості результатів педагогічних вимірів [2; 3]:

- емпірично отриманий результат виміру (X) є сумою дійсного результату виміру (T) і помилки виміру (E). Величини T і E зазвичай невідомі.
- істинний результат виміру можна виразити як математичне очікування $E(X)$;
- кореляція істинних та помилкових компонентів по групі випробуваних дорівнює нулю;
- помилкові компоненти двох будь-яких тестів не корелюють;
- помилкові компоненти одного тесту не корелюють із дійсними компонентами будь-якого іншого тесту.

Крім цього, основу класичної теорії тестів становлять два визначення – паралельних та еквівалентних тестів [4].

Однак класична теорія має ряд недоліків, що обмежують її використання у сфері педагогічних вимірів.

Оцінка рівня підготовленості студента залежить від складності тесту. Так, якщо тест складається з важких тестових завдань, то частка правильних відповідей у студента буде низкою. Якщо ж тест складається з легких тестових завдань, то частка правильних відповідей у того ж самого студента буде високою.

Оцінка складності тестового завдання залежить від рівня підготовленості студентів. Так, якщо вибірка складається з добре підготовлених студентів, то складність тестових завдань буде невисокою. Якщо ж вибірка складається з погано підготовлених студентів, то труднощі тестових завдань буде високою.

Шкала виміру рівня підготовленості є нелінійною. Наприклад, та сама різниця в балах (5 балів) на краях і в середині шкали відповідає різному збільшенню в рівні підготовленості (100 балів – 95 балів) \neq (50 балів – 45 балів).

Дисперсія оцінки рівня підготовленості найбільша в середині діапазону виміру й найменша по краях,

що суперечить логіці побудови довірчих інтервалів.

Тестовий бал студента нелінійно залежить від рівня його підготовленості.

Аналіз предметної галузі дозволяє зробити декілька висновків:

- задача розробки та використання тестів як засобів реалізації педагогічних вимірів є актуальною та повинна розглядатись у комплексі із іншими задачами підвищення рівня підготовки студентів та оптимізації навчального процесу;
 - на даний час розроблено велика кількість математичних моделей та методів тестування та обробки його результатів;
 - розглянуті методи обробки результатів тестування, включаючи класичні статистичні методи та методи сучасної теорії тестування, дозволяють здійснювати аналіз результатів тестування, але не здатні надавати рекомендації з оптимізації навчального процесу та адаптації змісту тестів, складності завдань та розподілу часу на вивчення різних тем;
 - адаптивні методи комп'ютерного тестування найчастіше використовуються для розв'язання проблеми адаптації процедури пред'явлення тестових завдань;
 - у сучасних дослідженнях недостатньо уваги приділяється питанню адаптації навчальних курсів у залежності від результатів тестування, хоча деякі засоби (наприклад ШНМ) дозволяють її реалізувати.
- Цей аналіз дозволяє сформулювати основні задачі дослідження:
- визначити специфіку функціонування штучних нейронних мереж та можливість їх використання при розробці педагогічних тестів зі змінною складністю завдань;
 - розробити контрольні-вимірювальні матеріали для адаптивного тестування з певної дисципліни;
 - удосконалити розроблені контрольні-вимірювальні матеріали за рахунок використання штучних нейронних мереж;
 - розробити методіку застосування отриманих результатів для адаптації системи навчання студентів.

2. Опис проведених теоретичних досліджень

Одним з шляхів підвищення ефективності комп'ютерних тестів є розробка та реалізація методів адаптивного тестування (АТ). Під АТ зазвичай розуміють комп'ютерну систему науково обґрунтованої перевірки та оцінки результатів тестування, яка має високу ефективність за рахунок реалізації можливості оптимізації процедур генерації, пред'явлення та оцінки результатів виконання адаптивних тестів [4].

Набагато частіше зустрічаються завдання, у яких можливі проміжні варіанти відповідей, або ж взагалі варіантів немає, питання є відкритим, і відповідь оцінюється в якій-небудь шкалі (наприклад,

п'ятибальна). Прикладом моделі, що враховує градацію правильних відповідей, є Partial Credit Model (PCM).

Безсумнівним плюсом застосування IRT моделей є можливість одержувати одночасно з оцінками компетенцій студентів обґрунтовані статистичні оцінки завдань, що може бути основою для поліпшення освітніх програм ВНЗ. Оцінка рівня підготовленості студентів не залежить від набору завдань, а неповнота даних (пропуск деяких комбінацій «випробуваний – завдання») не є критичною. Однак слід відзначити, що досліджувані завдання повинні бути гомогенними («одномірними»), тобто формувати та оцінювати тільки одну компетенцію. Основний шлях збільшення об'єктивності, результатів, а також розширення сфери застосування є використання в якості «сирих» балів не екзаменаційних оцінок, а оцінок, що отримуються під час проміжної атестації за виконання різних завдань. Цей шлях вимагає перегляду й значного розширення банку контрольних завдань по кожному предмету – кожне окремо взяте завдання повинне діагностувати тільки одну компетенцію. Так само це повинне знайти відображення в інформаційній системі ВНЗ – облік і зберігання цих оцінок.

Цю модель доцільно використовувати, коли:

- передбачається оцінювати частково вірні відповіді (наприклад тестові завдання з множинним вибором);
- завдання потребує послідовності кроків в його виконанні наприклад розв'язання задачі з математики). При цьому складність кожного кроку може бути різною.

АТ дозволяє підвищити ефективність педагогічних вимірів, скоротити кількість питань у тесті, знизити затрати часу та вартості тестування та підвищити точність тестування.

Для реалізації технології АТ необхідно [5]:

- визначити цілі реалізації АТ (навіщо потрібна адаптація у конкретному тесті?);
- визначити фактори, які будуть враховуватись у якості вхідної інформації під час прийняття адаптаційних рішень (до чого буде адаптуватися тест?);
- які аспекти будуть змінюватись у процесі адаптації (що буде адаптуватися?);
- які механізми адаптації будуть використовуватись та як вони будуть реалізовані (як буде здійснюватися адаптація?).

Як правило, АТ базується на процедурі оптимізації складності завдань у залежності від припущень щодо рівня підготовки студентів. У найпростішому варіанті загальна процедура виглядає наступним чином. Студент отримує перше завдання тесту на основі початкових припущень. У разі його виконання рівень складності наступного завдання підвищується. У протилежному випадку рівень наступного питання знижується.

Процес тестування може бути закінчено, якщо, наприклад, студент не зміг виконати три завдання поспіль. У загальному випадку можуть бути використані більш складні процедури зупинення.

Тобто АТ навіть у простішій реалізації дозволяє динамічно змінювати кількість та складність тестових завдань конкретного студента. Більш складні алгоритми АТ дозволяють обирати наступне завдання тесту, приймаючи до уваги не тільки складність завдань, але також їх приналежність до певної теми дисципліни, форму представлення та інші фактори.

Таким чином будується індивідуальний тест для кожного студента. Різні студенти отримують різні тести, які розрізняються за складністю та складом завдань, проходячи простір тестів за різними траєкторіями (рис. 1).

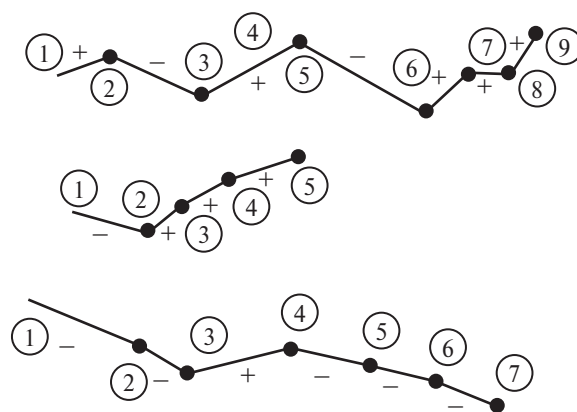


Рис. 1. Візуалізація індивідуальних траєкторій

На цьому рисунку показано три приклади траєкторій трьох студентів. Цифри позначають номери питань, знаки + та – відповідають вірним та невірним відповідям. Для припинення процедури тестування використовується просте правило: тестування закінчується, якщо студент вірно виконав три завдання поспіль, або зробив три помилки поспіль.

Різні алгоритми АТ можуть розрізнятися також стратегією тестування. Розрізняють двокрокові та багатокрокові стратегії. У рамках двокрокової стратегії на першому кроці всі студенти отримують однакові тести. За їх результатами вони розподіляються на осі змінній виміру. На другому кроці вмикається адаптивний режим, та здійснюється підсумкове адаптивне тестування.

Адаптивна навчаюча система (АНС) – це система, яка застосовує у процесі взаємодії зі студентом алгоритми адаптації до певних його характеристик [12]. Найчастіше це рівень знань, цілі, рівень підготовки, досвід роботи з засобами комп'ютерного представлення інформації (гіпермедіа, мультимедіа та ін.).

Використання нечіткої логіки – один з поширених напрямів інтелектуалізації систем контролю знань. Засоби нечіткої логіки використовуються

окремо, або у комбінації з іншими (наприклад з мережами Байєса).

Найбільш поширеним є перехід від завдання складності питань або відповідей у категоріях бінарної логіки до більш загальної та універсальної схеми оцінювання відповідей на основі функцій належності, які визначаються в категоріях нечіткої логіки.

Використання нечіткої оцінки завдань та відповідей передбачає і використання нечітких правил виводу для формування гіпотези про рівень складності наступного завдання [6].

Тестування починається з припущення, що студент має певний рівень знань (S), який має нечіткий характер. Обчислюється група $G=G(S)$. Студент отримує завдання відповідного рівня складності. При вірній відповіді збільшується кількість вірних відповідей, відсоток вірних відповідей та змінюється складність. На базі цього визначається новий рівень підготовки ($S1$) як функція поточного рівня підготовки S , відсотка вірних відповідей p , рівня складності завдання T , та часом на відповідь t :

Якщо кількість виданих завдань менше певної кількості, яка є обов'язковою для відповіді у межах одного рівня складності, то $G=G1$ та студент отримує нове завдання.

Якщо кількість критичних завдань перевищена, то рівень знань вважається рівним $S1$, та студент отримує відповідну оцінку $R = R(G(S1))$. Якщо змінилася група, то встановлюється $S=S1$, $G=G1$, обнуляється кількість завдань, та здійснюється повернення до видачі завдань.

В цієї процедурі АТ складність завдань змінюється у залежності від вірності відповідей студента, а поняття рівня підготовки, вірності відповіді, складності завдань та ін. є нечіткими та задаються за допомогою нечітких оцінок.

Марківські процеси дуже активно використовуються для реалізації алгоритмів АТ. Розглянемо функціонування АТ на основі моделі, яка використовує марковські процеси з дискретними станами та дискретним часом (ланцюг Маркова). Особливістю цього підходу є те, що складність завдань визначається на основі граничних розподілів ймовірностей перебування у станах, отриманих за допомогою матриць ймовірностей переходу. Тут оцінка враховує всю історію виконання тестових завдань, яка містить розподіл успішних та неуспішних виконань завдань та їх порядок, а також час виконання тестових завдань. Ймовірності переходу між станами є параметрами моделі.

Мережі Петрі широко використовуються не тільки для реалізації АТ, але і для моделювання прогресу студентів під час вивчення деякого курсу у адаптивних навчаючих системах [7]. Моделювання здійснюється на рівні подій. Визначається, які дії відбуваються

у системі, які стани передували цим діям та як і у яких станах опиниться система після виконання дії. Виконання моделі подій описує поведінку системи. Аналіз результатів виконання може розповісти про те, у яких станах перебувала, або не перебувала система, які стани не є досяжними. Під час тестування алгоритм обирає переходи до завдань, які мають більш високий або більш низький рівень складності.

Рівень знань є найбільш важливою характеристикою студента (користувача) та найбільше часто використовуваним джерелом інформації для адаптації. Рівень знань є змінною величиною для кожного конкретного студента. Тому будь-яка система, що адаптується до знань користувача, повинна мати його модель, фіксувати зміни рівня знань і відповідним чином корегувати цю модель.

Найпоширенішою в даний час є оверлейна модель студента, яка заснована на структурній моделі дисципліни, що вивчається. Структурна модель зазвичай представляється у вигляді семантичної мережі, що відбиває взаємозв'язки між окремими елементами знань із цієї дисципліни (поняттями, діями, завданнями й ін.). Для кожного елемента знань, представленого в мережі модель зберігає деяке значення, що представляє собою оцінку знання студентом даного елемента. Іншим підходом до побудови моделі користувача є поділ студентів на певні категорії й наступне використання для адаптації інформації не про окремого студента, а узагальненої інформації про категорію, до якої він віднесений. У цьому випадку модель виходить трохи більш спрощеної й надає менше можливостей для адаптації дій системи.

На основі аналізу зазначених параметрів моделі користувача адаптивний навчальний курс повинен дозволяти змінювати зміст матеріалу, що пропонується для вивчення, його форму представлення, послідовність проходження за курсом і можливості навігації на сторінках курсу (якщо це дисципліна, що представлена в гіпермедійному форматі). Адаптація змісту може реалізовуватися шляхом представлення навчального матеріалу з різним ступенем деталізації, з урахуванням рівня знань про суміжні теми та поняття, а також моделі користувача.

Багатокрокові стратегії поділяються на фіксовано-розгалужені та змінно-розгалужені. У фіксовано-розгалужених стратегіях для всіх студентів використовується однакові тестові набори завдань з фіксованим розташуванням на осі складності. Але шлях кожного студента є різним. Як правило, всі завдання на цієї осі розташовані на одній відстані, або шаг зменшується відповідно складності. Це дає можливість змінювати швидкість тестування у залежності від стану втомленості студентів.

У змінно-розгалужених тестах завдання обираються з бази завдань на основі певних алгоритмів,

яки прогнозують оптимальну складність наступного завдання. З цих окремих завдань складається адаптивний тест. Ця стратегія реалізує покрокову переоцінку рівня знань студентів, яка здійснюється після кожного виконання завдання тесту.

Вибір початкових оцінок для входу в процедуру АТ може ґрунтуватися на результатах попереднього тестування, на результатах тестів, які проводилися протягом певного періоду, або на основі моделі суб'єкта навчання, якщо вона використовується у системі.

Необхідно відзначити, що коректність роботи ШНМ прямо залежить від правильності вибору моделі ШНМ і алгоритму навчання для поставленого завдання, що важливо. Аналіз можливостей та властивостей різних ШНМ дозволяє зробити висновок, що з усіх моделей ШНМ найкращою для вирішення задач адаптивного тестування є перцептрон Розенблатта [8].

Для подальшого дослідження необхідно розглянути класифікацію способів навчання ШНМ і визначитися з алгоритмом навчання ШНМ. Якщо алгоритм буде обраний неправильно, то ШНМ не зможе коректно навчитися на результатах тестування, а відповідно, не можна буде говорити ні про яке поліпшення адаптивного тесту.

Алгоритм передбачає виконання наступної послідовності дій:

- підготовляється навчальна вибірка вхідних та вихідних значень;

- на початку роботи ШНМ усім зв'язкам мережі присвоюються невеликі випадкові значення;

- з навчальної вибірки береться поточний приклад і його вхідні параметри, (що у сукупності складають вектор вхідних сигналів) подаються на вхідні синапси початкової ШНМ. Зазвичай кожний вхідний параметр прикладу подається на один певний вхідний синапс;

- ШНМ робить задану кількість епох функціонування, вектор вхідних сигналів поширюється по зв'язках між нейронами (пряме функціонування);

- визначення дійсного виходу мережі, виконання її перерахування;

- обчислення різниці між дійсним і бажаним виходом мережі. Чим менше ця різниця, тем краще розпізнаний приклад, і тем ближче відповідь до необхідного. Якщо різниця рівна 0, то ніяких дій не вживається;

- зміна ваги мережі для мінімізації помилки (у цьому і є зміст навчання);

- повторення пунктів доти, поки загальна помилка мережі не буде мінімальною. Прохід по всіх прикладах навчальної вибірки з першого до останнього вважається одним циклом навчання.

Як можна бачити з опису алгоритму навчання ШНМ, цей варіант найкраще підходить для аналізу

комп'ютерних інструктивних матеріалів на предмет їх відповідності заявленій тематиці, тобто іноді тестові завдання по певній темі не завжди підходять для визначення рівня знань саме для цієї теми. Для рішення цієї проблеми необхідно:

- описати банк термінів, які будуть розподілятися по класах, тобто потрібно виділити ознаки, по яких терміни будуть поєднуватися в класи. Той самий термін може належати різним класам. Для цього необхідна строга ієрархія термінів;

- після того, як будуть виділені необхідні ознаки термінів для кожного конкретного класу з наступним їхнім розподілом, необхідно навчити мережу на готових прикладах, у яких заздалегідь відомо до якого класу буде віднесений кожен термін;

- після навчання ШНМ (добору коефіцієнтів) можна пред'являти їй завдання. У процесі кластеризації термінів по класах ШНМ буде виносити рішення про відповідність або невідповідність даного завдання заявленій темі.

При навчанні без вчителя ШНМ не має потреби в цільовому векторі для виходів і, отже, не потребує порівнянь із визначеними ідеальними відповідями. Навчальна множина складається лише із вхідних векторів. Навчальний алгоритм налаштовує ваги мережі так, щоб виходили погоджені вихідні вектори, тобто щоб пред'явлення достатнє близьких вхідних векторів давало однакові виходи. Процес навчання виділяє статистичні властивості навчальної множини та групує подібні вектори в класи. Пред'явлення на вхід вектора з даного класу дасть певний вихідний вектор, але до навчання неможливо передбачити, який вихід буде проводитися даним класом вхідних векторів. Отже, виходи подібної мережі повинні трансформуватися в деяку зрозумілу форму, обумовлену процесом навчання.

При змішанім навчанні частина ваг визначається за допомогою навчання із учителем, а інша частина отримується за допомогою самонавчання.

3. Опис методу адаптивного тестування із застосуванням логічних мереж

3.1. Розробка контрольньо-вимірювальних матеріалів

Для якісного та вірогідного виміру знань студентів необхідно діяти згідно такої процедури: відібрати зміст навчання, для цього розглядається програма обраного (необхідного) курсу; розробити систему тестів по обраній дисципліні (створити базу завдань); розробити необхідне програмне забезпечення для здійснення тестування; побудувати ШНМ та навчити її на результатах попередніх тестувань; провести тестування; виконати обробку та інтерпретацію отриманих результатів.

Згідно цієї процедури, на першому етапі для проведення дослідження було обрано курс «Інформа-

цінні забезпечення професійної діяльності» (ІТПД), який курс викладається студентам спеціальностей «Правознавство» та «Правоохоронна діяльність» Харківського національного університету внутрішніх справ (ХНУВС) та містить одинадцять тем. Загальний обсяг бази результатів тестування складає біля 300 студентів, оскільки в умовах дистанційного навчання тестування проходить кожен студент. Це достатньо велика кількість для проведення нашого дослідження. Для навчання та тестування студентів використовується система Moodle.

Загальний обсяг бази – 100 тестових завдань. Для реалізації тестування питання проходили попередню обробку – розподіл по темах та оцінку складності.

Для попередньої обробки тестових завдань було розраховано кілька величин: ємність відповіді на завдання, імовірність правильної відповіді, імовірність «сліпого» угадування правильної відповіді, складність завдання.

Ємність відповіді на завдання I . Дана характеристика показує кількість інформації, що міститься у відповіді на завдання. Будемо вважати, що для студента всі відповіді рівнозначні. Тоді для того, щоб оцінити величину I , необхідно буде скористатися формулою для рівномірних подій:

$$I = \text{Log}_2 M, \quad (1)$$

де M – кількість можливих відповідей на завдання.

Імовірність правильної відповіді на завдання P . Ця характеристика відображає знання групи студентів з розділу, якому відповідає завдання визначається як відношення кількості вірних відповідей до загальної кількості відповідей на це завдання.

Імовірність «сліпого» вгадування правильної відповіді P_y . Цей показник характеризує подію тільки в тому випадку, коли студент не читає текст завдання або не розуміє його. У цьому випадку вибір правильної відповіді носить випадковий характер (вгадування). У даної характеристики немає єдиної формули, тому для кожного виду завдання вона буде розраховуватися окремо.

Складність завдання Q враховує статистику відповідей студентів та їх знання завдання. Для визначення рівня складності завдання будемо застосовувати формулу, що відбиває невизначеність стану об'єкту:

$$Q = -\sum_{i=1}^M P(X_j) \text{Log}_2 P(X_j), \quad (2)$$

де M – кількість можливих відповідей на завдання; $P(X_j)$ – ймовірність j -ї відповіді.

Початкові умови тестування зазвичай завдаються таким чином, що спочатку кожен з варіантів відповідей на завдання рівномірнісний. Для того, щоб більш точно визначити рівень складності завдань, необхідно щоб усі вони пройшли експериментальну перевірку. По завершенню цієї перевірки буде

отримана інформація про необхідні статистичні характеристики завдань.

Складність Q також може визначатися на основі експертної оцінки. Зазвичай у ролі експерта виступає викладач, що часто є виправданим у випадку невеликої кількості завдань, хоча може привнести певний суб'єктивізм.

3.2. Процедура адаптивного тестування знань

Для реалізації адаптивного тестування було запропоновано процедуру, яка реалізує багатокрокову варіюючу стратегію. Ця стратегія під час тестування здійснює вибір наступного питання з банку питань на базі результату відповіді на два попередні з урахуванням його рівня складності.

Тестові завдання адаптивного тесту передбачаються закритої форми з вибором одного з 4-х запропонованих варіантів. На кожному кроці тестування по кожному рівню складності студенту дається два завдання, і за результатами відповідей на них визначається рівень складності для наступних завдань. Така кількість завдань (два) дозволяють більш адекватно оцінювати рівень знань, ніж одне завдання, і в той же час не дає великої кількості комбінацій варіантів відповідей, як у випадку трьох та більшої кількості завдань.

Алгоритм можна описати наступним чином. У тесті є m рівнів складності завдань (у тестовому випадку їх 3). Вводиться коефіцієнт $K_i = 100/m$. На другому кроці позначимо поточний рівень знань студента як t , t_n – нижній рівень знань, t_v – верхній рівень знань. Усі рівні знань будуть вимірятися від 0 до 100 (0 – немає знань, 100 – абсолютні знання).

Спочатку будемо вважати, що студент має середній рівень підготовки. Тому встановимо $t = 50$, $t_n = 0$, $t_v = 100$. Обчислюємо поточний рівень складності: $tt = t/K_i$.

На наступному кроці пред'явимо студенту два завдання складності tt , при цьому стежимо за кількістю правильних відповідей k_{pr} .

Перераховуємо рівень знань із урахуванням відповідей на два завдання:

Якщо $k_{pr} = 2$, то $t_n = t$; $t_v = t_v + 0,5t$. Якщо $t_v > 100$, то $t_v = 100$.

Якщо $k_{pr} = 1$, то $t_n = t_n/4$; $t_v = t_v + 0,1t$. Якщо $t_v > 100$, то $t_v = 100$.

Якщо $k_{pr} = 0$, то $t_n = t_n/2$; $t_v = t$.

Обчислюємо $t_v = (t_n + t_v)/2$.

Якщо $|t - t_v| > 0$, то $t = t_1$.

Якщо досягнуто критичний рівень кількості завдань або балів за завдання, то рівень знань рівний t_1 . Вихід.

У протилежному випадку перейти до другого кроку.

Отриманий результат перераховується згідно обраної шкали оцінок.

Результати тестування у бінарному вигляді представляються у вигляді таблиці. Дані таблиці є вхідною інформацією для другого кроку адаптації – оцінки складності тестових запитань.

Систем вважає необхідним підвищити рівень складності завдань з номерами 2, 4, 14 та 20, а також знизити рівень складності для завдання № 25. Для решти завдань рівень змінювати не потрібно.

Ця інформація повинна бути розглянута експертом (викладачем) для аналізу та прийняття рішення.

Наступний етап обробки результатів тестування – надання рекомендації щодо зміни часу на окремі теми дисципліни.

У результаті роботи ШНМ педагогам-розроблювачам комп'ютерних навчальних матеріалів надається інформація про номер завдання, його початковий рівень складності та рекомендація щодо можливої зміни рівня складності (підвищити, залишити без змін, знизити). Результати обробки вхідних даних представлено на рис. 2.

Дійсно, якщо в ході адаптивного тестування виявлене загальне нерозуміння студентами однієї з тем навчальної дисципліни, то при наступному навчанні педагог повинен затратити більше навчального часу й приділити більшій увазі вивченню цієї теми. Практичне впровадження системи коректувань змісту й методики освіти залежно від рівня підготовленості студентів, виявленого в ході адаптивного тестування, стикається із проблемами обробки й обліку більших інформаційних потоків, рішення яких також можливо на основі використання ШНМ.

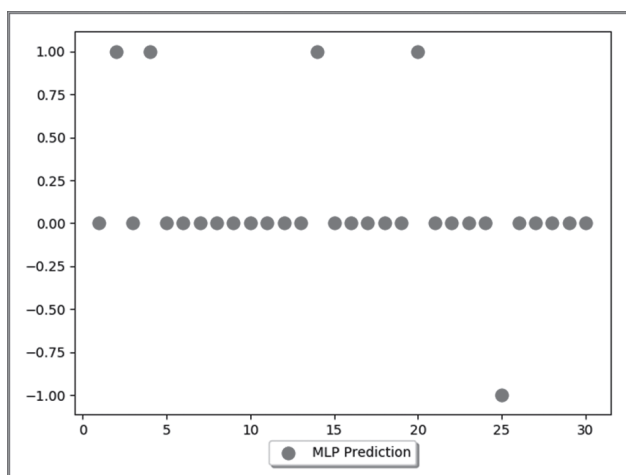


Рис. 2. Результати роботи ШНМ

На підставі проведення тестів можна стверджувати, що створений прототип виконує всі функції, що зазначені у вихідних даних на розробку й відповідає всім вимогам, що пред'явлені до нього.

Визначено ряд обмежень логічного характеру, що визначають як перехід зі стану в стан, видалення станів, так і обмеження кількості переходів (тестування не може проводитися нескінченно й питання

не повинні повторюватися). Однак, при невеликій різниці між зазначеними передбачуваними й дійсними властивостями групи випробуваних застосування зазначеної методики дає позитивний ефект, із чого можна зробити висновок, що застосування запропонованого методу є ефективним для будь-яких обсягів і інших властивостей, що визначають тестування як процес.

Висновки

Ідея адаптивного тестування, заснована на блоках питань, має безпосереднє відношення до одного із самих розповсюджених форматів багатоступінчастого тестування, при якому суб'єкт навчання проходить через послідовність тестів, рухаючись у бік більш складних питань при успішних відповідях або до більш простих, якщо його відповіді невірні. При цьому перехід від питання до питання відбувається за певними правилами. Автоматизація у практиці тестування надає можливість статистичної оцінки знань на кожному кроці тестування.

Очікуваний результат по використанню розроблених моделей стосовно до 3-х параметричної моделі тестування має бути досягнутий при апроксимації функції успіху, що припускає визначення уточнених значень дискримінаторів завдань шляхом рішення завдання оптимізації при використанні відомих алгоритмів. Це пов'язане насамперед з тим, що застосовані алгоритми оптимізації є універсальними, а рішення виниклої проблеми вимагає додаткового дослідження метою якого є часткове заміщення цих алгоритмів операційними методами, що використовують як властивість поставленого завдання тестування так і іншої апріорної інформації.

Різниця в рівнях складності основних і додаткових питань і запропонований зв'язок між основними питаннями і гілками додаткових питань дозволяє в процесі тестування як мінімізувати кількість необхідних відповідей суб'єкта навчання для визначення рівня його знань, так і суттєво поліпшити адаптаційні властивості тестування.

Для розв'язання цієї проблеми було запропоновано використовувати ШНМ з тією ж структурою. Входом мережі є вектор оцінок студентів з певної теми. Вихідна реакція – рекомендація щодо можливих змін у обсягу часу на вивчення цієї теми (збільшити, не змінювати, знизити).

Після попереднього навчання на вхід системи було подано вхідні сигнали та отримана рекомендація «не змінювати» обсяг часу на вивчення конкретних тем за досліджуваними наборами тестів з дисциплін.

Для обробки результатів тесту та врахування додаткових чинників, які можуть впливати на результати тестування, можуть використовуватися також

статистичні методи. Ці методи дозволяють отримати оцінки додаткових показників тестів:

- відсоток вірних або невірних відповідей;
- відносну успішність групи студентів по одній темі стосовно всіх інших тем курсу;
- розподіл знань студентів по всім темам;
- середня навченість групи та ін.

Список літератури:

- [1] Baker R. “Educational data mining and learning analytics.” The Cambridge handbook of the learning sciences, 2019, p. 274.
- [2] Rzhetska S. “Experience of using clustering methods for analyzing the results of distance learning.” Informatization of engineering education: materials of international scientific-practical conf., No. 56, 2016, pp. 617 – 620.
- [3] I.Shubin, V.Skovorodnikova, A.Kozyriev, Mining Methods for Adaptation Metrics in E-Learning, in: Proceedings of the 3rd International Conference Computational Linguistics and Intelligent Systems CoLInS 2019, NTU “KhPI”, Kharkiv, (2019) 288–300.
- [4] Chetverikov G.G., Vechirska I.D., Tanyanskiy S.S. The methods of algebra finite predicates in the intellectual system of complex calculations of telecommunication companies // International Conference Proceedings Crimean Microwave and Telecommunication Technology (CriMiCo).-2014, 6959425. - pp. 346-347.
- [5] Sharonova, N., Kyrychenko, I., Tereshchenko, G. Application of big data methods in E-learning systems CEUR Workshop Proceedingsthis link is disabled, 2021, 2870, pp. 1302–1311
- [6] Gruzdo, I., Kyrychenko, I., Tereshchenko, G., Cherednichenko, O. Application of paragraphs vectors model for semantic text analysis CEUR Workshop Proceedingsthis 2020, 2604, pp. 283–293
- [7] G. Bezhanishvili “Locally finite varieties” Algebra Universalis, Vol. 46, no. 4, 2001, pp. 531–548.
- [8] Chetverikov G., Puzik O., Vechirska I. Multiple-valued structures of intellectual systems //Proceedings of the with Internations Computer Sciences and Information Technologies (CSIT). 2016, 7589907. -pp. 204-207.
- [9] A Guide to the SCRUM BODY OF KNOWLEDGE (SBOK™ GUIDE) URL: <https://www.scrumstudy.com/SBOK/SCRUMstudy-SBOK-Guide-3rd.pdf>.

Надійшла до редколегії 19.05.2021