УДК 004.93

А. Г. Кіслая, Л.Е. Чала, О.Є. Гриньова
ХНУРЕ, м. Харків, Україна, alla.kyslaia@nure.ua
ХНУРЕ, м. Харків, Україна , larysa.chala@nure.ua
ХНУРЕ, м. Харків, Україна , olena.hrinova@nure.ua

НЕЙРОМЕРЕЖЕВА МОДЕЛЬ ВИЯВЛЕННЯ БОТІВ В СОЦІАЛЬНИХ МЕРЕЖАХ

Виконано аналіз видів соціальних ботів, виявлено їх вплив на користувачів. Описано ознаки, за якими можна визначити спамерських пошукових роботів. Проаналізовано алгоритми розподілу інформації соціальними мережами. Запропонована архітектура нейронної мережі для виявлення ботів, а також надані результати її роботи для аналізу користувачів мережі Twitter і їх твітів.

НЕЙРОННА МЕРЕЖА, БОТИ, МЕТОД ВИЯВЛЕННЯ БОТІВ, РЕКУРЕНТНІ МЕРЕЖІ

А.Г. Кислая, Л.Э. Чалая, Е.Е. Гринева. Выявление ботов в соцсетях. Выполнен анализ видов социальных ботов, выявлено их влияние на пользователей. Описаны признаки, по которым можно определить ботов. Проанализированы алгоритмы распределения информации социальными сетями. Предложена архитектура нейронной сети для выявления ботов, а также предоставлены результаты ее работы для анализа пользователей сети Twitter и их твитов.

НЕЙРОННАЯ СЕТЬ, БОТЫ, МЕТОД ВЫЯВЛЕНИЯ БОТОВ, РЕКУРРЕНТНЫЕ СЕТИ

A.G. Kislaia, L.E. Chala, O.Y. Grynova. Bot detection in social networks. The analysis of types of social bots was performed, their influence on users was revealed. The signs by which it is possible to identify bots are described. The algorithms for the distribution of information by social networks are analyzed. A neural network architecture was proposed to identify bots, and the results of its work were provided for analyzing Twitter users and their tweets.

NEURAL NETWORK, BOTS, METHOD OF BOT DETECTION, RECURRENT NETWORKS

Вступ

З розвитком інформаційних технологій з'являється все більше способів маніпуляції людською свідомістю. Існує безліч інструментів і різних платформ, але саме соціальні мережі стали головним плацдармом для подібних дій. Йдеться про поширення так званих соціальних ботів – спеціальних програм, створених для імітації поведінки людей в соціальних мережах. Призначення цих програм може бути різним, але найчастіше до їх використання вдаються інтернет-маркетологи і кіберзлочинці.

Соціальні боти – це напівавтоматичні або автоматичні програми, які використовують звичайні функції соціальних мереж (додавання контенту, реагування на чужий контент тощо). За допомогою мереж ботів - програм, які керують фейковими аккаунтами в соціальних мережах, можна штучно впливати на актуальність будь-якої теми.

Запрограмовані армії ботів реагують в соціальних мережах на пости з певними ключовими словами. Виявивши задане слово або хештег, вони генерують лайки, репости, залишають під постом заготовлені позитивні або негативні коментарі. Згідно з високою штучною активністю згенерованих матеріалів, алгоритми соцмереж виводять ботему в тренд і передають її іншим користувачам. Так вона стає популярною і автоматично виходить в топ найбільш обговорюваних. При цьому за трендами в Twitter, Instagram, Facebook тощо стежать рекламодавці, політики, журналісти, які намагаються зрозуміти, що цікавить інтернет-користувачів.

Маніпулювати трендами можна по-різному. Наприклад, владі якоїсь авторитарної країни не подобається популярний хештег, який використовують опозиціонери. В цьому разі за допомогою ботів вони розкручують в мережі альтернативну кампанію і витісняють початковий хештег з топа. У підсумку до нього менше звертають увагу, а потім і зовсім забувають.

Коли дві третини учасників спільноти є ботами, то ефективність, наприклад, торгових кампаній суттєво зменшується, адже боти нічого не купують та не цікавляться новинами. Більш того, фальшиві користувачі створюють непотрібний інформаційний шум, а також загрожують звичайним користувачам.

Блокувати ботів доцільно зі старту, при додаванні в друзі або після підписки на сторінку. Це визначає актуальність розпізнавання фейкових аккаунтів, для чого можуть бути ефективно використовувані нейронні мережі та методи машинного навчання.

Метою даної статті є аналіз можливості автоматичного виявлення ботів в популярних соціальних мережах за допомогою запропонованих нейромережевих моделей.

1. Негативний вплив ботів на ефективність соціальних мереж

Мертві аккаунти в соцмережах дуже часто створюються для поширення спаму або збільшення рейтингу певним групам. Такі боти не загрожують групі і читачам.

Більш серйозні боти можуть публікувати тематичний контент і навіть починати дискусії в коментарях. Вони насправді загрожують інформаційній безпеці.

По-перше, боти можуть завалити стрічку новин рекламою. По-друге, вони часто поширюють нелегальний контент. По-третє, займаються шахрайством, наприклад, фішингом або зараженням комп'ютерів користувачів шкідливими програмами. По-четверте, боти часто використовуються для політичної пропаганди, збираючи інформацію про дії та вподобання користувачів.

Як не дивно, незважаючи на всю небезпеку ботів, SMM-менеджери часто самі звертаються до залучення мертвих аккаунтів. Втім штучно створена група з 10 тисячами читачів і сотнями записів не завжди сприяє просуванню бізнесу. Відсутність взаємодій з аудиторією не може допомогти такому просуванню та збільшити показники продажу товару.

У 2013 році боти атакували американське інформаційне агентство Associated Press. Зловмисникам вдалося поширити на сайті неправдиві повідомлення про атаку на Білий дім і пораненні Президента США Барака Обами. Ця новина швидко розійшлася і вплинула на американський фондовий ринок, при цьому деякі акції впали в ціні до критичних показників.

Дуже великою популярністю бот-атаки користуються в Росії. Тільки за 2016 рік там було зафіксовано більше тисячі атак на фінансові та політичні ресурси. Наприклад, online-обговорення «податку на інтернет» в Росії виявилося під загрозою зриву через бот-атаки. Було зафіксовано, що, починаючи з 20 квітня і до кінця обговорення, на сайт надходили виключно позитивні відгуки. Модератори сайту закрили коментарі на два дні і перевірили аккаунти. В ході розслідування були вилучені 3000 повідомлень від ботів.

Програмісти University of British Columbia провели цікавий експеримент, поширивши в мережі Facebook програму, що здатна самостійно зареєструвати аккаунт. При цьому боти вели себе в соціальній мережі, як звичайні користувачі, тобто обмінювалися фотографіями, статусами, коментували записи, шукали друзів.

Експериментатори помітили, що в друзі користувачі частіше додають невідомого, якщо бачать, що раніше його вже додали знайомі. Згоду на запити до друзів боти отримували в 60% випадків. Таким чином, боти завантажили 250 гігабайтів особистої інформації зі сторінок своїх нових друзів. Ініціатори експерименту запевняють, що вся зібрана ними інформація не отримала подальшого поширення і застосування. Дані були зашифровані, а згодом знищені.

Вивчає маніпуляції в соцмережах німецький політолог Симон Хегель (Simon Hegelich), який в 2014 році виявив мережу ботів, що впливали на тренди в розпал конфлікту на Україні. 15 тисяч фейкових аккаунтів залишали по 60 тисяч постів в день: вони публікували жарти, обговорювали спортивні події і між справою підтримували «Правий сектор». С. Хегель підтвердив в інтерв'ю DW і те, що армії ботів були активні під час виборчої кампанії в США в 2016 році. Агітаційні повідомлення нібито розсилали привабливі дівчата, за якими насправді ховалися спеціальні програми [1].

2. Види і задачі ботів

Розглянемо 6 основних груп соціальних та пошукових роботів.

Технічні боти. В їх структурі практично завжди є спеціально прописані програми. Головними завданнями таких програм є накопичення лайків і односкладових коментарів під потрібним записом, створення великої маси друзів для збільшення довіри до боту і поширення публікацій через репости. Це найбільш популярний вид ботів, який використовується в будь-яких соціальних мережах (найчастіше це Facebook, ВКонтакте або Instagram).

Бойові боти. Потрібні в першу чергу для зниження репутації або блокування певної сторінки в соціальній мережі шляхом відправлення великої кількості скарг і негативних коментарів.

Зливні боти. Іноді для поширення тієї чи іншої інформації використовуються спеціальні боти, які спочатку поводяться як реальні користувачі, але в певний момент починають поширювати інсайдерську інформацію. Згодом багато інтернет-видань та ЗМІ будуть посилатися на ці фальшиві джерела.

Гіперболізуючі боти. Це, мабуть, найбільш тонкий і досвідчений вид соціальних ботів, які призначені для входження в довіру до клієнтів конкурента замовника з подальшим створенням антиреклами. З першого погляду фейковий користувач повністю підтримує ідеологію і погляди опонента. Але в певний момент починає поширювати гіперболізовану (перебільшену) інформацію. Наприклад, бот проникає в групу любителів певної марки автомобіля і повідомляє, що всі власники автомобілів інших підприємств є не мають смаку, а пізніше його підтримують інші боти. В результаті багато потенційних покупців почнуть вельми скептично ставитися до всіх власників даних автомобілів і відповідно до всього бренду в цілому.

Інтелектуальні бойові боти. Ці боти цікаві тим, що використовують власний вкладений інтелектуальний ресурс. Тобто програма з необхідними закладеними даними відправляється на інформаційну війну для пропаганди сторонньої думки. Головне завдання такого бота -– спілкуватися на вузькоспеціалізовані теми або обговорювати певні гілки повідомлень. Крім цього, такий бот часто може переходити на образи чи провокаційні висловлювання щодо інших користувачів, тим самим відволікаючи уваги від основної теми бесіди. Цей

вид бота найбільш популярний при обговоренні політичних і соціальних явищ.

Боти від сайтів. До їх допомоги вдаються маловідомі компанії, створюючи аккаунт на віртуальну особистість. Зазвичай такі боти малоактивні і їх легко виявити [2].

3. Загальні ознаки ботів

Існує безліч різних ознак, за якими можна розпізнати бота. Серед них є як очевидні, так і виявлені за допомогою тривалого аналізу даних.

На підробленій сторінці часто зображені дуже приваблива дівчина або хлопець. Люди схильні звертати увагу на привабливі фотографії, такі фотографії збільшують ймовірність того, що запит в друзі буде прийнятий. Щоб переконатися, що зображення реальне, потрібно виконати пошук в Google Image – якщо знайдеться безліч сайтів з цим зображенням, можна припустити, що обліковий запис є фальшивим.

Найчастіше боти не викладають велику кількість фотографій або новин на своїх сторінках (якщо це мережа на зразок Facebook). Їх головна мета – докласти найменшу кількість зусиль, щоб створити видимість того, що за сторінкою в соцмережі є дійсно реальна людина. Тому, якщо сторінка контролюється ботом, швидше за все на ній не викладають фотографій.

Якщо біографічна інформація надто дивна або явно вигадана, то є велика ймовірність того, що це підроблений аккаунт.

Боти часто можуть приймати запит в друзі, але не можуть відповідати на повідомлення (або їхні відповіді обмежені певним набором фраз). Якщо є сумніви, чи аккаунт є реальним, то йому можна надіслати повідомлення і зачекати на реакцію.

Деякі акаунти, контрольовані ботами, налаштовані на те, щоб «лайкати» певну кількість сторінок в день (воно може бути обмежене, щоб не потрапити під простий алгоритм пошуку ботів, тому боти можуть «лайкати» велику кількість сторінок в день, і на цьому будуть ідентифіковані) .

Кількість друзів і читачів у фейкового аккаунта може бути більше тисячі, але більшість серед сумнівні або схожі на спамерських пошукових роботів.

Власний контент у ботів майже відсутній: бот не діляться спогадами, думками або ідеями. Для того, щоб сторінка не була порожньою і не виглядала підозрілою, дуже часто фейковий акаунти наповнюють її переписами.

4. Аналіз роботи алгоритмів соцмереж

В умовах постійно зростаючої кількості інформації соціальні мережі фільтрують її за допомогою різних методів.

Найпростіший спосіб вирішити, яку інформацію показувати користувачеві, а яку ні (або ж яку інформацію показувати в першу чергу) – це лінійні моделі з ваговими коефіцієнтами, що у загальному випадку мають таку форму:

$$y = \beta_0 + \sum \beta_i * x_i + \varepsilon_i.$$

Алгоритм випадковим чином призначає ваги для різних властивостей інформації, а потім уточнює цю інформацію на кожному прикладі, для якого відомою є правильна відповідь – це навчання з учителем.

У випадку з записами в Facebook, про які алгоритму треба відповісти «показувати / НЕ показувати», є ще одне корисне джерело інформації – це слова, з яких складено запис. Якщо модель знає, що людина любить читати про ракети, Марс і Ілона Маска, то за кожне з цих слів видасть багато балів, і випадкові знайомі, справами яких він не цікавився останні п'ять років, мають шанс пробитися до нього у френд-стрічку зі своїми міркуваннями про запуск автомобіля в космос.

Відзначимо, що прості і швидкі алгоритми не реагують на інтонації і стилі записів. Так, вони не відрізняють віршів від прози, приховані цитати і сарказм є за межами їх розуміння,тому що для простої моделі запис є мішком слів (у фахівців з машинного навчання це усталений термін): слова або є, або немає, а в якому порядку йдуть і що означають разом, вже неважливо. Також прості лінійні моделі не вміють шукати сенс в комбінаціях ознак.

Ще один спосіб ранжування інформації – моделі вирішальних дерев:

$$\beta_v(x, j, t) = \left[x_j < t \right].$$

Приклад використання таких моделей – постановка діагнозу лікарем. Сценарії діалогу лікаря з хворим можна представити в формі гіллястого дерева: кожне наступне питання залежить від відповіді на попереднє. Такий алгоритм добре справляється з десятками і сотнями ознак, але далі починає зазнавати труднощів. Тексти, в яких присутні десятки тисяч різних слів, або багатопіксельні малюнки, таким алгоритмом обробити не можна. Така інформація піддається аналізу тільки в напівобробленому вигляді, коли текст або малюнок уже перетворені в обмежений набір ознак. В такому разі набагато ефективнішим є використання нейронних мереж.

З точки зору математики кожний з окремих нейронів таких мереж є лінійною моделлю: він отримує електричні сигнали від інших нейронів, оцінює їх (в умовних балах) і видає сумарну оцінку у вигляді власного електричного сигналу.

Але один нейрон рідко приймає остаточне рішення: в мозку вони часто організовані в шари, і відповідь буде готовий, коли нервовий імпульс пройде їх все наскрізь, від верхніх до найглибших. Нейрони першого шару отримують «сирі» сигнали

- наприклад, вони можуть бути паличками або колбочками сітківки, які реагують на світло. Кожен нейрон другого шару буде обробляти імпульси від багатьох клітин першого і формувати свій електричний імпульс, щоб передати його далі.

Нейромережі можна довірити все відразу, тому що великим питанням потрібні великі обчислювальні потужності: «багатоповерхові» алгоритми вимагають більш громіздких розрахунків, ніж «одноповерхові». У листопаді 2016 року Google перемкнув на нейромережі свою систему машинного перекладу, але користувачі потребують перекладу текстів все-таки рідше, ніж оновлюють стрічку Facebook. У травні 2017 року на блозі компанії Twitter з'явилася новина, що тепер нейромережі беруть участь і в ранжуванні твітів (тобто вирішують, які показувати вище, а які нижче). Ключове слово тут «беруть участь»: вони як і раніше беруть на себе тільки частину роботи. Тому долю френдстрічки як і раніше вирішують прості математичні моделі.

Використовуючи знання алгоритмів ранжирування інформації, генератор соціальних ботів може вивести в тренд свою тему за допомогою простого програмування бота на відповідний чин поведінки.

5. Опис використовуваних даних

Розглянемо можливість використання нейромереж для виявлення ботів на прикладі даних соціальної мережі Twitter.

Аналіз проводився на двох рівнях: на рівні аккаунтів користувачів і на рівні їх твітів (рис.1, рис.2).

Рис. 1. Аккаунт в мережі Twiter

З акаунтів користувачів використовувалися наступні дані:

а) статуси;
б) читачі;
в) друзі;
г) відношення до сторінок інших користувачів;
д) опис профілю і його фото;
е) геолокація;
ж) використання фону;
з) верифікованість;
и) захищеність.

Рис. 2. Приклад твіту

На рис. 2 представлений приклад твіта. Тіло твіта може складатися з тексту повідомлення (що включає в себе фото, відео, посилання), тегів, по яких потім можна шукати твіти, згадок інших користувачів мережі. Також твіт характеризується кількістю коментарів, ретвітів і лайками. Таким чином, в експерименті були використані наступні складові:

а) кількість ретвітів;
б) кількість коментарів;
в) кількість лайків;
г) кількість хештегів;
д) кількість посилань;
е) кількість згадок.

6. Опис архітектури нейромережі

Метадані користувачів є найкращим предиктором для виявлення ботів. Це дозволяє використовувати чимало існуючих стандартних методів машинного навчання в нейромережах для ідентифікації ботів.

Було виявлено, що більшість з цих методів є достатньо ефективною, причому більшість підходів перевершувало точність передбачення більш ніж 90%. Найуспішніший підхід полягав у використанні класифікатора Random Forest, що дає точність 95%. Однак спостерігався значний приріст продуктивності при балансуванні набору даних з використанням методів передискретизації, зокрема, методу передискретизації синтетичної меншини (SMOTE) [3].

Алгоритм SMOTE генерує семпли, засновані на просторі прикладів меншини (класу, який має меншу кількість помічених точок даних), і є потужним методом, який успішно використовується в багатьох доменах. Зокрема, була використана комбінація SMOTE і двох методів передискретизації. Такі методи поліпшення даних використовуються для усунення будь-якого зсуву, введеного за допомогою передискретизації. SMOTE був об'єднаний з поліпшенням даних через Edited Nearest Neighbors (ENN) [4] і Tomek Links [5]. Було виявлено, що ці дві комбінації дають цілком задовільніі результати по незбалансованим даним. Хоча об'єднання SMOTE і субдіскретізація через Tomek Links (SMOTOMEK) не покращує результати в значній мірі, значне поліпшення спостерігається завдяки об'єднанню SMOTE і субдіскретізація через ENN (SMOTENN) для всіх моделей.

У випадку з аналізом твітів, щоб подолати обмеження традиційних методів нейронних мереж, була використана модель Long Short Term Memory

(LSTM) [6], що перевершує варіант рекурентних нейронних мереж (RNN) (рис. 3) [7]. RNN і їх варіації були визнані ефективними для задач НЛП, враховуючи їх здатність вивчати зв'язку в послідовних даних.

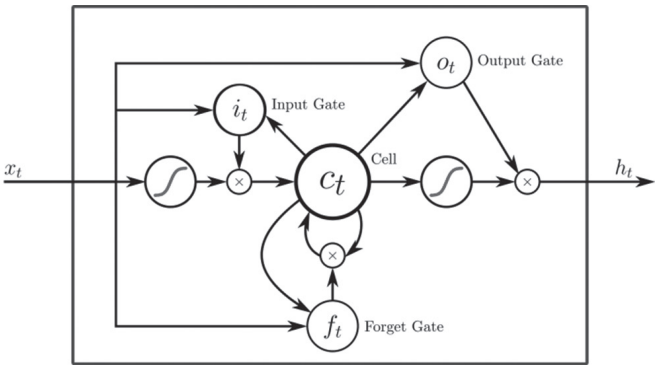

Рис. 3. Алгоритм LSTM

Для перетворення твітів у форму, придатну для LSTM, був використаний попередньо підготовлений набір глобальних векторів для подання слів (GloVE), призначений для даних Twitter [8]. GloVE – це глобальна лог-білінейна регресійна модель, яка використовує глобальну матричну факторизацію і методи локального контекстного вікна для ефективного вивчення субструктури природної мови шляхом навчання за випадковим збігом слів.

До навчання LSTM на твітах, твіти були оброблені шляхом формування рядка токенів з кожного твіту.

Хештеги, посилання і числа, згадані користувачем в повідомленні, були замінені на відповідні теги. Аналогічним чином, більшість поширених смайлів було замінено на теги, що відповідають цим емоціям.

Для слів, написаних великими літерами, або для слів, що містять більше 2 повторюваних букв, використовувався тег, що проставляється після появи слова. Всі маркери перетворені в нижній регістр.

В подальшому ці токенізовані твіти перетворюються у вхідну інформацію, використовуючи вищезгадану попередньо підготовлену модель GloVE. Результуюча формується виходить вектор з 32 ознаками, який надходить на вхід 2 активуючих ReLU шарів розмірностями 128 і 64 для отримання кінцевого результату. Модель скидає стан після кожної нової порції даних, тому навчальна послідовність виходить в рамках одного твіту, а не послідовності твітів.

Відзначимо, що такий підхід використовує тільки текстовий вміст твітів і не розглядає пов'язані з ним метадані. Використання метаданих не дозволяє точно визначити природу користувача, але метадані є слабкими предикторами характеру облікового запису (бот чи ні), тому їх доцільно також включити в мережу для більш точного аналізу.

В варіанті нейромережі з метаданими вводиться допоміжне введення даних в вихідний шар. Аналогічно раніше описаної моделі аналізу тільки твітів, головний вихід -– це текст твіта, який є токенізірованим і перетвореним в набір GloVE векторів перед обробкою їх LSTM. В результаті формується вектор, конкатенований з висновком, отриманим при обробці метаданих з двошаровою нейронною мережею, що використовує функцією активації ReLU. Отримані розмірності такі ж, як і в попередній моделі.

Як механізм регуляризації пропонується вихід метаданих, який також може надати можливість класифікації твітів. Втрата складається з зваженого середнього виводу метаданих та головного виведення в співвідношенні 2 до 8.

7. Опис експерименту

Для збору даних використовувалося Twitter API. Було зібрано близько 9 мільйонів твітів у восьми тисяч користувачів.

Незважаючи на те, що багато просунутих технік використовують велику кількість властивостей (1500 і більше), недавні дослідження показали, що досить використовувати обмежену, мінімальну кількість властивостей [9].

Лімітування кількості використовуваних властивостей пояснюється наступними причинами:

а) можливість інтерпретації. Обмежений набір властивостей зі зрозумілим значенням, на кшталт запропонованого в роботі, дозволяє продукувати інтерпретовані моделі. Це дуже важливий момент, особливо в комбінації зі стратегіями глибокого навчання, які, зазвичай, важко піддаються трактуванню;

б) ефективність моделі. Зменшений набір властивостей допомагає в отриманні ефективних моделей нейронних мереж, які потім можуть бути швидше натреновані і більш стійкі до перенавчання, що є досить поширеною проблемою в соціальних даних завдяки наявності великої кількості викидів.

Такий вибір кількості властивостей протиставлений системам, що використовують великий набір, які спроектовані для підтримки таких великих наборів, але обчислювальна ефективність яких неоптимальна, і можливість інтерпретації яких можна поставити під сумнів.

Для взаємодії з Twitter API була обрана бібліотека hbc, розроблена командою Твітетра, що працює з протоколом HTTP і взаємодіє з API безпосередньо за допомогою запитів.

Як бібліотека з реалізованою і кастомізованою нейромережею була обрана бібліотека Deeplearing4j - бібліотека програм на мові Java, щоа використовується як фреймворк для глибокого навчання.

Як БД обрана база MySQL(вільна реляційна система управління базами даних).

8. Аналіз отриманих результатів

Тестування нейромережі проводилося на виявлених заздалегідь 300 акаунтах ботів зі згаданих 8000 користувачів, а також на 500 аккаунтах звичайних користувачів. Був проаналізований 1 мільйон твітів.

Класифікатор Random Forest показав точність 90%.; класифікатори SMOTE та з ENN з субдіскретізацю (SMOTENN) показали точність близько 92.1%. Це говорить про те, що стратегія передискретизації синтетичного меншини дійсно ефективна при вирішенні таких завдань класифікації на рівні облікового запису. При підході, пов'язаному з передискретизацією за допомогою SMOTE, та субдискретизацією з TOMEK (SMOTOMEK), е менш ефективні, ніж при використанні ENN. В цілому, була продемонстрована можливість високоточного виявлення ботів на основі простих алгоритмів в поєднанні з синтетичними методами передискретизації меншини для поліпшення системуи адаптації.

Використовуючи тільки твіти, система LSTM забезпечила точність класифікації 85.13%. Цей результат можна вважати перспективним: використання такої архітектури дає максимальну продуктивність в розмірі близько 5%, навіть використовуючи тексти твітів.

Контекстний LSTM показав підвищення продуктивності, що приводить до точності класифікації близько 89.42%. Варто відзначити, що різні конфігурації GloVE, а саме зміна розміру простору вкладення слова, не суттєво впливають на продуктивність: загальна тенденція вказує на те, що більш висока розмірність дає ненабагато кращу продуктивність. З проведеного аналізу можна зробити висновок, що, хоча метадані твітів є слабким предиктором, пропонована архітектура, що використовує додаткову інформацію, надану метаданими, дозволяє отримати більш точні результати прогнозування. Слід зазначити, що метадані не були передіскретізіровані для систем контекстного LSTM.

Глибокі моделі нейронних мереж часто описуються як неінтерпретовані «чорні ящики». Хоча вони дають вражаючі результати в багатьох практичних додатках, але вимагають уваги до питання інтерпретування. Перспектиним є проведення досліджень внутрішнього стану рекурентних моделей з візуалізацією відповідної прихованої динаміки.

Висновки

У статті були розглянуті види ботів, їх вплив на користувачів, а також способи виявлення ботів в соціальних мережах за допомогою нейронних моделей.

Сьогодні для маніпулювання громадською думкою в мережу постійно випускаються тисячі ботів, які вже здатні спілкуватися майже без участі людини. Комп'ютерні програми-боти здатні маніпулювати соціальними мережами і в більш широкому масштабі - наприклад, загострювати увагу на якихось громадських проблемах, підтримувати гуманітарні акції або, навпаки, пригнічувати соціальну активність.

Запропонований нейромережевий підхід дозволяє виявляти ботів в мережах типу Twitter на двох рівнях: на рівні профілів користувачів, а також на рівні твітів. Результати експериментів підтвердили, що обидва підходи забезпечують достатньо високу точність передбачення.

Список літератури:

1. Dark side of SMM: чем опасны бот-аккаунты [Электронный ресурс] / M|C Today – Режим доступа: https://mc.today/dark-side-of-smm-chem-opasny-bot-akkaunty/ – 2017. – Загл. с экрана. **2.** *Марголін О.Г.* Система виявлення інформації у текстових повідомленнях користувачів / О.Г. Марголін // Штучний інтелект. — 2016. — № 4. — С. 85-91. **3.** *Chawla N. V.* SMOTE: synthetic minority oversampling technique [Text] / N. V. Chawla, K. W. Bowyer, L. O. Hall // Journal of artificial intelligence research. – 2002. – Vol.16. – P. 321–357. **4.** *Dennis L. W.* Asymptotic properties of nearest neighbor rules using edited data [Text] / L. W. Denis // IEEE Transactions on Systems, Man, and Cybernetics. – 1972. – Vol 2, №3. – P. 408-421. **5.** *Tomek I.* Two modifications of CNN [Text] / I. Tomek // IEEE Trans. Systems, Man and Cybernetics. – 1976. – Vol. 6 – P. 769–772. **6.** *Hochreiter S.* Long short-term memory [Text] / S. Hochreiter, J. Schmidhuber // Neural Computation. – 1997. – Vol. 9, №8. – P. 1735-1780. **7.** *Jozefowicz R.* An empirical exploration of recurrent network architectures [Text] / R. Jozefowicz, W. Zaremba, I. Sutskever // In Proceedings of the 32nd International Conference on Machine Learning (ICML-15). – 2015. – P. 2342–2350. **8**. *Samek W.* Understanding, Visualizing and Interpreting Deep Learning Models [Text] / W. Samek, T. Wiegand, K-R. Müller // Explainable Artificial Intelligence. – 2017. – №. 1708.08296. **9.** *Ferrara E.* The rise of social bots [Text] / E. Ferrara, O. Varol, C. Davis // Commun. ACM. – 2016. – Vol. 59, №7. – P. 96-104.

Надійшла до редколегії 11.04.2018