

РЕЦЕНЗІЯ

офіційного рецензента, доктора технічних наук, професора, завідувача кафедри інформаційних технологій проектування та прикладної математики Київського національного університету будівництва та архітектури **Терентьєва Олександра Олександровича** на дисертаційну роботу Стецика Олексія Андрійовича «Інтелектуальна високонавантажена розподілена система обробки даних в соціальних мережах» що представлена на здобуття ступеня доктора філософії за спеціальністю 126 «Інформаційні системи та технології» галузі знань 12 «Інформаційні технології»

Актуальність теми дисертації

Запропонована дисертація щодо виявлення спаму та пропаганди в соціальних мережах базується на Законах України: «Про схвалення Концепції розвитку штучного інтелекту в Україні» (2020 р.) «Про заборону пропаганди російського нацистського тоталітарного режиму, збройної агресії Російської Федерації як держави-терориста проти України, символіки воєнного вторгнення російського нацистського тоталітарного режиму в Україну» (2022 р.), «Про засудження та заборону пропаганди російської імперської політики в Україні і деколонізацію топонімії» (2023 р.). Що підтверджує її надзвичайну актуальність і своєчасність для України. Але і в мирних умовах у сучасну цифрову епоху швидке поширення неправдивої інформації та дезінформації на платформах соціальних мереж створює низку викликів, які впливають на демократичні суспільства, підривають довіру до державних інституцій, спотворюють суспільний дискурс і вводять в оману користувачів.

Здатність ефективно виявляти та зменшувати шкідливий контент має вирішальне значення для збереження цілісності та довіри до цих платформ, а розробка надійних методів виявлення пропаганди та спаму, здатних ефективно працювати в розподілених соціальних мережах з високим трафіком, є не лише технічною необхідністю, а й нагальним практичним завданням.

Таким чином, актуальність дослідження Стецика Олексія Андрійовича підкреслюється нагальною потребою захисту публічного дискурсу, покращення користувацького досвіду та підтримки демократичних цінностей, яким загрожує поширення неправдивої інформації у високонавантажених соціальних мережах.

Ступінь обґрунтованості наукових положень, висновків і рекомендацій, сформульованих у дисертації

Обґрунтованість наукових положень і висновків, сформульованих у дисертації, підтверджується публікаціями 2 статей у науковому фаховому виданні України категорії «Б» і статті у періодичному науковому виданні іншої держави, яка входить до ЄС. Окрім того матеріали дослідження доповідались на чотирьох міжнародних конференціях, пройшли ретельну апробацію та отримали підтвердження у вигляді рецензій і висновків фахівців галузі інформаційних технологій.

Обґрунтованість рекомендацій щодо побудови інтелектуальних високонавантажених соціальних мереж та їх практичного застосування доведена довідками про впровадження в роботу Департаменту суспільних комунікацій Київської міської державної адміністрації. Таким чином, ступінь обґрунтованості наукових положень, висновків і рекомендацій, сформульованих у дисертації, відповідає сучасним стандартам галузі.

Наукова новизна отриманих автором результатів.

У дисертації представлено кілька важливих інновацій у сфері класифікації тексту в стрічках новин.

Застосування сучасних методів машинного навчання, рекурентних нейронних мереж, згорткових нейронних мереж, нейронних мереж з трансформер архітектурою та сучасних технологій для побудови архітектур високонавантажених соціальних мереж *по-перше* надає можливість удосконалити архітектуру систем соціальних мереж в напрямку наближення швидкості класифікацій текстів і генерації диверсифікованої стрічки новин до режиму реального часу шляхом комбінації масштабованих технологій і натренованої моделі класифікації; *по-друге* пропонує цінну інформацію та потенційні напрямки подальшого розвитку і вдосконалення інтелектуальних високонавантажених розподілених систем та *по-третьє* забезпечує конкурентну спроможність системи.

У дисертації запропоновано нейронну мережу, розроблену саме для класифікації тексту в стрічках новин. Ця модель, поєднуючи гібрид згорткових нейронних мереж і двосторонніх мереж з довгою короткочасною пам'яттю, досягає ефективною та швидкою класифікації тексту. На основі зазначеної моделі

удосконалено метод виявлення спаму і пропаганди, в напрямку отримання балансу між точністю і швидкістю класифікації. Також у дисертації вперше розроблено метод зменшення ефекту «бульбашки» через диверсифікацію стрічки новин, який ґрунтується на класифікації новин за конкретними дискусійними категоріями і подальшій кластеризації текстів в цих дискусійних категоріях, що виконуються запропонованою нейромережею.

Теоретичне та практичне значення одержаних результатів

Запропонований метод зменшення ефекту «бульбашки», що містить елементи класифікації тексту, кластеризації і користувацького досвіду, концептуально розширює межі застосування гібридних моделей до вирішення задачі обробки даних у високонавантажених системах.

Запропонований метод класифікації стрічки новин має практичне значення в боротьбі з спамом і пропагандою в соціальних мережах, а удосконалена архітектура соціальної мережі підтримує інтеграцію текстового класифікатора та методу диверсифікації стрічки новин, оптимізованого для високонавантажених середовищ, що працюють у режимі, близькому до реального часу. Практична корисність цих методів і моделей підтверджена їх впровадженням в роботу Департаменту суспільних комунікацій Київської міської державної адміністрації для вирішення завдань виявлення пропаганди в текстах новин і зменшення ефекту «бульбашки», що використовуються в сучасних соціальних мережах з метою розшарування суспільства.

Повнота викладення основних результатів дисертації в опублікованих працях.

Основні результати дослідження викладено у трьох публікаціях, одна з яких індексується у наукометричній базі даних SCOPUS, а також представлено і опубліковано в збірниках матеріалів трьох наукових конференцій, одна з яких індексується у наукометричній базі даних SCOPUS,.

Публікації здобувача відповідають вимогам, що встановлені «Порядком присудження ступеня доктора філософії та скасування рішення разової спеціалізованої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії» затвердженим постановою Кабінету Міністрів України № 44 від 12 січня 2022 року.

Зміст дисертації

Дисертацію присвячено вивченню проблем, що характерні сучасним високонавантаженим розподіленим системам соціальних мереж, і їх вдосконаленню шляхом впровадження моделей і методів штучного інтелекту.

Дисертація складається із анотації, вступу, чотирьох розділів та висновків, списку використаних джерел та додатків.

Анотація повністю розкриває зміст дисертаційної роботи, а її основні положення ідентичні основним положенням дисертаційного дослідження.

У вступі сформульовано мету та задачі дослідження, визначено об'єкт, предмет і методи дослідження, наукову новизну і апробацію результатів дисертації.

У першому розділі розглянуто виклики, з якими стикаються сучасні високонавантажені розподілені соціальні мережі. В ньому проведено аналіз існуючих підходів до боротьби зі спамом, пропагандою та ефектом «бульбашки», а також розглянуто причини виникнення вразливостей і методи виявлення та усунення несправностей у розподілених системах.

У другому розділі досліджено переваги і недоліки використання машинного навчання та нейронних мереж для класифікації спаму і пропаганди. Запропоновано модель класифікації текстів на основі гібриду мережі довгої короткочасної пам'яті та згорткової нейронної мережі. На основі даної моделі запропонований метод класифікації текстів, оптимізовано на точність і швидкодію. Також у розділі запропоновано метод диверсифікації стрічки новин, по дискусійним темам.

У третьому розділі описано основні операції користувачів у соціальних мережах і створено масштабовану модель для них. У ньому детально розглянуто потенційні можливості розширення цієї моделі за рахунок збільшення функціональності та визначено її вразливі місця. Основну увагу приділено продуктивності та пропускній здатності системи.

У четвертому розділі розроблено інструменти для кластеризації текстів та їх класифікації як спаму або пропаганди. Результати класифікації текстів за допомогою різних моделей і методів машинного навчання та нейронних мереж порівняно за показниками точності, повноти, влучності та f-міри. Також проведено експерименти з кластеризації текстів на дискусійні теми.

У висновках перелічено отримані наукові результати, які повністю відповідають задачам дисертаційного дослідження.

Список використаних джерел налічує 171 посилання.

У додатках розміщено акти впровадження результатів дисертаційного дослідження Департаментом суспільних комунікацій і довідку про впровадження результатів дисертаційного дослідження в навчальний процес Київського національного університету будівництва та архітектури .

Зауваження

Подана до захисту дисертаційна робота Стецика Олексія Андрійовича не зважаючи на стилістичні, граматичні помилки і недоліки оформлення, легко читається і сприймається. Але, позитивно оцінюючи цю роботу в цілому, слід звернути увагу на окремі дискусійні положення, які потребують роз'яснення і уточнення:

1. У розділі 2 в моделі класифікації, що запропонована у дисертації, на рис.2.10 використано функцію активації, яка використовується при бінарній класифікації, тому незрозуміло як ця модель застосовується при вирішенні задачі багатокласової класифікації текстів у 4-му розділі.

2. У розділі 3 в 3.3 запропоновано архітектуру системи, практичним застосуванням якої має бути соціальна мережа. На мою думку, варто було б оцінити у якому хмарному провайдері (Amazon, Azure чи GCP) найдешевше запустити цю систему.

3. У розділі 3 в 3.4 сервіс для післяобробки постів отримує повідомлення від брокера повідомлень Кафка, але не зрозуміло, що відбуватиметься, якщо сервіс для післяобробки постів не встигатиме в реальному часі обробляти повідомлення, які потрапили в брокер повідомлень Кафка.

4. Дисертаційне дослідження сконцентроване на роботі з англійсько мовними текстами, хоча було б цікаво побачити як запропонована система виконує цю обробку текстів, написаних російською та українською мовами

Висновок

Вважаю, що дисертаційна робота Стецика Олексія Андрійовича «Інтелектуальна високонавантажена розподілена система обробки даних в соціальних мережах» є повністю завершеним науковим дослідженням, що виконане на високому професійному рівні та не порушує принципів академічної

добросовісності. Результати, що одержані в роботі є новими, відповідають поставленим цілям і задачам та мають теоретичне та практичне значення. Усі наведені зауваження орієнтовані на перспективу подальших науково-практичних розробок і не зменшують цінність результатів роботи.

Дисертаційна робота відповідає вимогам, які висуваються до дисертаційних робіт, зокрема зміст дисертації загалом відповідає галузі знань 12 «Інформаційні технології», спеціальності 126 «Інформаційні системи та технології», «Вимогам до оформлення дисертації», затвердженим Наказом Міністерства освіти і науки України від 12.01.2017 р. № 40 (із змінами, внесеними згідно з Наказом Міністерства освіти і науки України від 31.05.2019 № 759) та «Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії», затвердженому Постановою Кабінету Міністрів України від 12 січня 2022 р. № 44 (в редакції постанови Кабінету Міністрів України від 19 травня 2023 р. № 502), а її автор – Стецик Олексій Андрійович заслуговує присудження наукового ступеня доктора філософії в галузі знань 12 – «Інформаційні технології» спеціальності 126 – «Інформаційні системи та технології».

Офіційний рецензент,

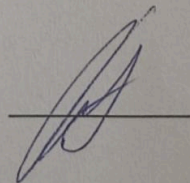
завідувач кафедри інформаційних технологій проектування

та прикладної математики

Київського національного університету

будівництва та архітектури

доктор технічних наук, професор



Олександр ТЕРЕНТЬЄВ

Особистий підпис завідувача кафедри інформаційних технологій проектування та прикладної математики, доктора технічних наук, професора Олександра Терентьєва засвідчую.

Вчений секретар Вченої ради

Київського національного університету

будівництва та архітектури

кандидат технічних наук, доцент



Микола КЛИМЕНКО