РЕЦЕНЗІЯ

офіційного рецензента, докторки технічних наук, доцентки, завідувачки кафедри інформаційних технологій Київського національного університету будівництва та архітектури **Гончаренко Тетяни Андріївни**
на дисертаційну роботу Стецика Олексія Андрійовича «Інтелектуальна високонавантажена розподілена система обробки даних в соціальних мережах», що представлена до захисту на здобуття ступеня доктора філософії за спеціальністю 126 «Інформаційні системи та технології» галузі знань 12 «Інформаційні технології»

Актуальність теми дисертації

Соціальні мережі стали важливими для спілкування і обміну інформацією, але вони також створюють можливості для маніпуляції та утворення ехо «бульбашок», а також поширення спаму, пропаганди та дезінформації. Неправдива інформація та дезінформація може маніпулювати громадською думкою, підривати довіру до державних інституцій і підбурювати до насильства. Спам може зменшити задоволення користувача від користування платформою, що призводить до зниження рівня залученості та відмови від платформи.

Соціальні мережі стають все більш високонавантаженими і розподіленими, а велика кількість даних, яку необхідно аналізувати у режимі реального часу, додає технічної складності виявленню в них спаму та пропаганди. Тому потреба в надійних моделях і методах їх виявлення стає все більш критичною.

Саме цим актуальним у сучасному цифровому просторі аспектам розвитку соціальних мереж присвячене дисертаційне дослідження Стецика Олексія Андрійовича.

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

Викладені наукові положення, висновки і рекомендації є повністю обґрунтованими, а достовірність теоретичних положень підтверджується експериментальними даними і результатами верифікації запропонованих моделей та методів. Дисертація написана українською мовою. Матеріал і всі етапи роботи сформовані і описані послідовно, змістовно та доступно.

Основні висновки дисертації відображають проведене комплексне дослідження і пройшли ретельну апробацію, а саме: були висвітлені та обговоренні під час презентацій матеріалів дослідження на чотирьох міжнародних конференціях і опубліковані у трьох статтях, одна з яких індексується в наукометричній базі даних SCOPUS.

Достовірність і перспективність результатів роботи також підтверджується їх впровадженням у діяльність відділів управління інформаційної політики і організаційно-аналітичного забезпечення та контролю

Департаменту суспільних комунікацій Київської міської державної адміністрації.

Впевнена, що наукові положення та висновки дисертації є достатньо обґрунтованими та аргументованими.

Наукова новизна отриманих автором результатів.

Наукова новизна результатів дисертаційного дослідження полягає в розробці моделі для класифікації текстів. Ця модель є гібридом згорткової нейронної мережі та мережі довгої короткочасної пам'яті, яка використовує оптимізовані гіперпараметри і векторні представлення слів на основі нейромережі з трансформер архітектурою. На основі запропонованої моделі розроблено метод класифікації текстів.

Окрім того в дисертаційній роботі запропоновано метод диверсифікації стрічки новин, який дозволяє користувачам ознайомлюватися з думками і поглядами щодо дискусійних тем в широкому діапазоні, а також боротися з ефектом бульбашки в соціальних мережах. Метод класифікації текстів на спам і пропаганду оптимізований не тільки по точності, але й по швидкості.

Також у роботі запропоновано архітектуру системи високонавантаженої розподіленої системи соціальної мережі, в яку входять сервіси, що реалізують запропоновані в дисертаційному дослідженні методи класифікації текстів і диверсифікації стрічки новин. Архітектура системи орієнтована на високу масштабованість та низьку затримку відповіді на запит.

Проведене порівняння ефективності розроблених моделей і методів з іншими сучасними розробками показало їх конкурентну спроможність і підтвердило важливість подальшого дослідження в цьому напрямку.

Теоретичне та практичне значення одержаних результатів

Модель класифікації текстів, що запропонована в дисертаційному дослідженні розширює сферу обробки природньої мови, демонструючи те, як різні архітектури нейронних мереж можуть бути використанні для оптимізованої класифікації текстів.

Теоретичне значення запропонованого нового методу зменшення ефекту «бульбашки» шляхом диверсифікації стрічки новин по дискусійних темах полягає в розширенні розуміння того, як новини можна класифікувати та групувати для представлення різних точок зору, тим самим сприяючи більш збалансованому споживанню інформації.

Практична цінність одержаних в роботі наукових результатів підтверджується впровадженням запропонованих моделей і методів у діяльність Департаменту суспільних комунікацій Київської міської державної адміністрації та інтеграцією в навчальний процес Київського національного університету будівництва і архітектури.

Повнота викладення основних результатів дисертації в опублікованих працях.

Основні результати дослідження викладені в **двох** статтях, опублікованих у науковому фаховому виданні України категорії «Б», в **одній** статті у періодичному науковому виданні, проіндексованому у базі даних ***Scopus*** та в матеріалах трьох наукових конференцій. Публікації здобувача відповідають вимогам, встановленим «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

Дисертаційне дослідження Стецика Олексія Андрійовича «Інтелектуальна високонавантажена розподілена система обробки даних в соціальних мережах» присвячено вивченню і вдосконаленню високонавантаженим розподіленим системам соціальних мереж шляхом імплементації моделей і методів штучного інтелекту.

Дисертація складається із анотації, вступу, чотирьох розділів, висновків до розділів, основних висновків, списку використаної літератури та додатків.

У першому розділі досліджуються виклики, з якими стикаються користувачі і розробники сучасних розподілених соціальних мереж. Проведено аналіз таких проблеми, як спам, пропаганда та ефект «бульбашки», аналізуються першопричини цих проблем та сучасні методи боротьби з ними. У розділі також розглядаються такі архітектурні вразливості, як збої в роботі серверів, неочікувані сплески кількості користувачів і аналізуються методи виявлення та пом'якшення цих недоліків.

У другому розділі запропоновано модель на основі нейронних мереж для виявлення спаму і пропаганди. Це гібридна модель поєднує згорткову нейронну мережу і мережу з довгою короткочасною пам'яттю. У розділі також представлено новий підхід до зменшення ефекту «бульбашки» шляхом кластеризації текстів за дискусійними темами.

У третьому розділі запропоновано розширену масштабовану модель для платформ соціальних мереж на основі дій користувачів. У ньому висвітлюються потенційні можливості розширення та визначаються вразливі місця в моделі. Особлива увага приділяється продуктивності системи та її здатності витримувати високі навантаження. Обговорюється інтеграція моделі соціальних мереж з методами класифікації При цьому робиться акцент на масштабованість і точність.

У четвертому розділі розроблено програмні інструменти і код для моделі класифікації текстів на спам і пропаганду та методу кластеризації текстів, що описаний у другому розділі. У розділі порівнюється продуктивність різних моделей машинного навчання та нейронних мереж за такими показниками, як точність, влучність, повнота та f-міра. У ньому також описано експерименти з кластеризації текстів на дискусійні теми, які демонструють ефективність запропонованих методів у реальних ситуаціях.

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

Зауваження

У цілому позитивно оцінюючи дисертацію, яку можна використовувати в навчальному процесі в якості новітнього посібника при розробці сучасних інтелектуальних високонавантажених розподілених систем, і висловлюючи побажання побачити роботу цієї системи з новинами, шо написані українською мовою, варто звернути увагу на окремі непорозуміння, помилки та недоліки:

1. У вступі недостатньо чітко визначена актуальність роботи, а висновки до розділів потребують розширення і уточнень.
2. У роботі зроблено акцент на точності та швидкодії класифікації текстів. Було б добре застосувати техніку квантизації для нейронних моделей та оцінити наскільки це б в вплинуло на швидкодію та точність.
3. У архітектурі системи соціальної мережі використано нереляційну базу даних Кассандра. Існують і інші бази даних (наприклад Монго), які теж мають властивість масштабування. На мою думку, потрібно було навести більш змістовну аргументацію, чому обрано зазначений тип бази даних.
4. У роботі описано сервіс користувацьких операцій, який між користувачем і даним сервісом знаходиться балансувальник навантаження, який має балансувати навантаження між сервісами користувацьких операцій. Не зрозуміло, як масштабується сам сервіс користувацьких операцій при великому напливі користувачів.
5. Потребує уточнення питання щодо апробації матеріалів дисертації в частині «Наукові праці, які засвідчують апробацію матеріалів дисертації» абстракту і «Апробація результатів дисертації» вступу роботи, які містить різний перелік конференцій.
6. У роботі присутні незначні друкарські, граматичні та стилістичні помилки.

Висновок

Незважаючи на окреслені зауваження, вважаю, що дисертаційна робота Стецика Олексія Андрійовича «Інтелектуальна високонавантажена розподілена система обробки даних в соціальних мережах» є завершеною науковою працею, яка має теоретичне та практичне значення. Сформульована в дисертації мета дослідження досягнута, а зроблені зауваження мають переважно рекомендаційний характер і не зменшують цінність результатів роботи в цілому.

Робота відповідає вимогам, які висуваються до дисертаційних робіт, зокрема зміст дисертації загалом відповідає галузі знань 12 «Інформаційні технології», спеціальності 126 «Інформаційні системи та технології», та «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 на присудження наукового ступеня доктора філософії в галузі знань 12 «Інформаційні технології» за спеціальністю 126 «Інформаційні системи та технології».

Офіційний рецензент,
завідувачка кафедри інформаційних технологій
Київського національного університету
будівництва та архітектури
докторка технічних наук, доцентка Тетяна ГОНЧАРЕНКО

Особистий підпис завідувачки кафедри інформаційних технологій, докторки технічних наук, доцентки Тетяни ГОНЧАРЕНКО **засвідчую.**

Вчений секретар Вченої ради
Київського національного університету
будівництва та архітектури
кандидат технічних наук, доцент Микола КЛИМЕНКО