

## ВІДГУК

офіційного опонента, кандидата технічних наук, доцента, доцента кафедри системного аналізу та теорії оптимізації ДВНЗ «Ужгородський національний університет» Андрашка Юрія Васильовича 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**

Обсяг даних, які генеруються в соціальних мережах, щодня зростає. Для ефективного аналізу і використання цих даних потрібні передові системи обробки, здатні обробляти великі обсяги даних в режимі реального часу. Високонавантажені системи забезпечують швидку обробку і аналіз даних, що є критичним для надання актуальної інформації користувачам соціальних мереж. Використання методів штучного інтелекту і машинного навчання дозволяє отримувати корисні інсайти з великих обсягів даних, що сприяє більш точному прогнозуванню поведінки користувачів та покращенню персоналізованих рекомендацій. Розподілені системи обробки даних забезпечують кращий контроль над збереженням та захистом даних користувачів, що є важливим аспектом кібербезпеки. Такі системи дозволяють ефективніше використовувати апаратні та програмні ресурси, забезпечуючи масштабованість і надійність обробки даних. Аналіз даних у соціальних мережах може мати значний вплив на суспільство, включаючи розуміння соціальних тенденцій, виявлення дезінформації та підтримку соціальних ініціатив. Таким чином, дослідження спрямовані на підвищення ефективності, швидкості та безпеки обробки даних у соціальних мережах, є надзвичайно важливим в умовах сучасного інформаційного суспільства.

Дисертаційне дослідження присвячено вирішенню у реальному часі проблем класифікації спаму, пропаганди та подоланню ефекту «бульбашки», які виникають у високонавантажених розподілених системах соціальних мереж при великій кількості користувачів.

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

Результати, представлені в даній дисертації, належним чином підтверджені та аргументовані, відображаючи всебічний процес дослідження. Ці результати ефективно ілюструються через авторські публікації та перевірені часом. Основні наукові висновки дисертації були представлені на міжнародних наукових форумах, включаючи конференції та семінари, де вони отримали визнання і апробацію вченої спільноти. Ці результати також підкріплені відповідними публікаціями дослідника і офіційною довідкою про їхнє впровадження, що свідчить про високий рівень наукової значущості та практичної цінності проведених досліджень. Загалом, рівень наукового дослідження в дисертації відповідає сучасним стандартам у вказаній галузі, а отримані результати, висновки і рекомендації заслуговують на визнання та подальше вивчення.

## **Наукова новизна отриманих автором результатів**

У дослідженні "Інтелектуальна високонавантажена розподілена система обробки даних в соціальних мережах" розроблено модель класифікації тексту у стрічці новин на основі модифікації гібриду згорткової нейронної мережі та двосторонньої мережі довгої короткочасної пам'яті. На основі розробленої моделі запропоновано метод класифікації розробленою нейромережею новин за конкретними дискусійними категоріями для зменшення ефекту «бульбашки». Удосконалені методи виявлення спаму і пропаганди є актуальними в сучасних умовах. Набула подальшого розвитку концепція створення гібридних систем, що заснована на моделях штучного інтелекту і машинного навчання в напрямку обробки даних для отримання чистішого і різноманітнішого середовища в соціальних мережах.

Ці результати сприяють поглибленню теоретичних та практичних знань у сфері обробки великих даних і машинного навчання, зокрема у контексті розробки ефективних моделей для класифікації тексту та боротьби з негативними явищами в соціальних мережах. Вони також підкреслюють важливість інтеграції інтелектуальних технологій для підвищення якості інформаційного простору та

забезпечення більшої безпеки і достовірності інформації, що поширюється серед великої кількості користувачів.

### **Теоретичне та практичне значення одержаних результатів**

Теоретичне та практичне значення одержаних результатів полягає у кількох важливих аспектах. Теоретично, розробка моделі нейронної мережі для класифікації тексту у стрічці новин сприяє поглибленню знань у галузі машинного навчання та нейронних мереж, зокрема щодо їх застосування для класифікації тексту. Вперше запропоновані методи зменшення ефекту «бульбашки» та диверсифікації стрічки новин розширюють концептуальне розуміння інтеграції нейромереж для досягнення більш тонкої і різноманітної інформаційної екосистеми. Практично, запропоновані методи дозволяють ефективніше виявляти спам і пропаганду в соціальних мережах, що є критичним для підвищення якості інформації та захисту користувачів від дезінформації. Крім того, удосконалена архітектура систем соціальних мереж, спрямована на підвищення швидкодії та ефективності обробки даних, забезпечує можливість роботи в режимі реального часу при великій кількості користувачів, що є важливим для масштабованих платформ.

Впровадження результатів дослідження в діяльність Департаменту суспільних комунікацій Київської міської державної адміністрації та в навчальний процес в Київському національному університеті будівництва і архітектури підтверджує практичне значення одержаних результатів.

### **Повнота викладення основних результатів дисертації в опублікованих працях**

Ознайомлення з дисертацією та науковими публікаціями Стецика Олексія Андрійовича дозволяє зробити висновок, що основні положення дисертаційної роботи в основному висвітлені в 6 наукових публікаціях, у тому числі 3 наукових статтях у фахових виданнях, одна із яких індексується в наукометричній базі SCOPUS та 3 працях апробаційного характеру у збірках матеріалів конференцій.

### **Зміст дисертації**

Дисертація складається із анотації, вступу, чотирьох розділів і висновків.

Анотація в повній мірі розкриває зміст дисертаційної роботи, а основні положення, наведені в анотації, ідентичні основним положенням дисертаційної роботи.

Дисертація присвячена аналізу і моделюванню сучасних інтелектуальних високонавантажених розподілених систем соціальних мереж.

У першому розділі досліджуються основні проблеми таких систем, включаючи спам, пропаганду, ефект "бульбашки", архітектурні вразливості, що впливають на продуктивність і надійність системи.

Другий розділ описує моделі та методи виявлення спаму, пропаганди і зменшення ефекту "бульбашки" за допомогою машинного навчання і нейромереж. Зокрема, розроблено нову модель нейронної мережі для класифікації тексту у стрічці новин, яка поєднує згорткову нейронну мережу і мережу довгої-короткочасної пам'яті.

Третій розділ зосереджений на створенні моделі високонавантаженої розподіленої системи соціальної мережі, включаючи опис користувацьких операцій, масштабовану модель, методи покращення продуктивності, інтеграцію класифікації спаму і пропаганди, та диверсифікацію стрічки новин для зменшення ефекту "бульбашки".

У четвертому розділі проводяться експериментальні дослідження розроблених моделей і методів, включаючи порівняння різних моделей і методів класифікації текстів за точністю, повнотою та іншими метриками, а також експерименти з кластеризації текстів по дискусійних темах.

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

### **Зауваження**

У цілому позитивно оцінюючи подану до захисту дисертацію, водночас, варто звернути увагу на окремі дискусійні положення дисертаційної роботи:

1. Висновки дослідження сформовано як коротку анотацію отриманих результатів відповідно до поставлених задач для досягнення мети. Проте у висновках відсутня стисла інтерпретація результатів а також кількісні та якісні порівняльні оцінки результату. Також слід зауважити, що мета не є сумою задач.

Задачі – це шлях до досягнення мети, а мета – це те, до чого треба дійти, вирішуючи задачі. Тому у висновках слід обґрунтувати досягнення мети дослідження.

2. В розділі 2.3 описана модель нейронної мережі для класифікації тексту у стрічці новин, яка містить понад 6 мільйонів параметрів, проте в дослідженні не вказано чому обрано саме таку кількість шарів. Доцільно було б дослідити як впливає зміна архітектури мережі, наприклад видалення шарів Conv-Maxpooling чи зміна довжини вектору в мережі довготривалої короткочасної пам'яті, на точність та швидкість моделі.

3. В розділі 3.4, описано архітектуру системи соціальної мережі, реалізація якої передбачає використання мікросервісів, Kafka, Amazon S3, СУБД Postgresql та СУБД Cassandra, проте в роботі не вказані результати тестування системи (чи її прототипу) яке б підтвердило «генерацію диверсифікованої стрічки новин до режиму реального часу», що заявлено в науковій новизні одержаних результатів.

4. В розділі 4 описано результати тестування розробленої та навченої моделі нейронної мережі для класифікації тексту у стрічці новин для задач класифікації спаму та пропаганди, а також для кластеризації дискусійних текстів. Із наведених результатів не зрозуміло які саме тексти використано при визначенні точності моделі, проте певний опис тестових наборів даних вказано в описі моделі в розділі 2. Вважаю, що опис тестових даних було б доцільно перенести в розділ 4. Також варто було б вказати статистичні дані тестових наборів даних, зокрема середню мінімальну та максимальну довжину текстів в наборі.

В роботі наявні певні синтаксичні та пунктуаційні помилки, описки.

Водночас, висловлені зауваження та пропозиції мають, в основному, рекомендаційний і дискусійний характер, впливають із складності й багатогранності досліджуваної теми та окреслюють перспективні напрями подальших наукових розробок досліджуваної проблематики. Саме тому вони не знижують загальної позитивної оцінки дисертаційного дослідження, його наукової та практичної цінності.

## Висновок

Зазначені зауваження не впливають на високий науковий та прикладний рівень поданої дисертації. Вважаю, що дисертація Стецика Олексія Андрійовича «Інтелектуальна високонавантажена розподілена система обробки даних в соціальних мережах» є завершеною науковою працею, в якій отримані обґрунтовані наукові результати.


Дисертація 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" зі спеціальності 126 «Інформаційні системи та технології».

Офіційний опонент,

доцент кафедри системного аналізу та теорії оптимізації

ДВНЗ «Ужгородський національний університет»

кандидат технічних наук, доцент

 Юрій АНДРАШКО

02.08.2024

