

Голові разової спеціалізованої
вченої ради PhD 8729
Хмельницького національного університету
доктору технічних наук, професору
Тетяні ГОВОРУЩЕНКО

ВІДГУК

офіційного опонента на дисертаційну роботу
Молчанової Марини Олексіївни
за темою «Методи виявлення та класифікації прийомів та об'єктів пропаганди у
текстовому контенті засобами штучного інтелекту»,
подану на здобуття наукового ступеня доктора філософії
за спеціальністю 122 – «Комп'ютерні науки»

Актуальність теми та зв'язок з науковими планами та програмами.
У сучасному інформаційному суспільстві текстовий контент відіграє ключову роль у формуванні суспільної думки, політичних переконань та світоглядних орієнтирів громадян. Зі зростанням обсягів цифрової інформації, особливо в соціальних мережах, онлайн-ЗМІ та блогах, спостерігається активне поширення пропагандистських матеріалів, які мають на меті маніпулювати свідомістю людей, викривлювати факти та формувати вигідні для певних сил наративи. У цьому контексті постає нагальна потреба в ефективних інструментах виявлення та аналізу таких впливів, особливо з урахуванням того, що традиційні методи, орієнтовані на ручну перевірку та експертну оцінку, не можуть впоратися з масштабами сучасного інформаційного потоку.

Використання методів штучного інтелекту надає можливості для автоматизованого аналізу великих обсягів текстових даних з метою виявлення прихованих маніпулятивних прийомів та пропагандистських об'єктів. Завдяки здатності сучасних моделей машинного навчання до обробки природної мови та семантичного аналізу, стає можливим виявлення як явних, так і завуальованих форм інформаційного впливу.

Особливої актуальності тема набуває в умовах гібридних воєн, інформаційного протистояння та зростання ризиків дестабілізації соціально-політичної ситуації внаслідок цілеспрямованого впливу через медіа. Україна, яка вже кілька років знаходиться в центрі такого протистояння, гостро потребує науково обґрунтованих та технологічно ефективних рішень для захисту інформаційного простору. Розробка систем виявлення і класифікації

пропагандистських прийомів на основі штучного інтелекту є не лише науково значущою, а й має вагоме практичне значення для забезпечення національної безпеки, медіаграмотності населення та збереження інформаційного суверенітету.

Дисертаційна робота Молчанової М. О. виконана в рамках виконання науково-дослідної роботи за держбюджетною темою Хмельницького національного університету №0121U112025 «Розроблення інформаційної технології прийняття контрольованих людиною критично-безпекових рішень за ментально-формальними моделями машинного навчання». Робота присвячена підвищенню точності та якості виявлення прийомів та об'єктів пропаганди за семантичними маркерами у текстовому контенті засобами штучного інтелекту з подальшим поясненням прийнятих рішень. Виходячи з наведеного, тема дисертаційної роботи є важливою та актуальною.

Аналіз змісту дисертації та основні результати роботи. Дисертаційна робота присвячена розробці методів для класифікації текстів за вмістом пропаганди й для виявлення прийомів і об'єктів пропаганди нейромережевими моделями з візуальною інтерпретацією прийнятих рішень. Структура роботи дозволяє послідовно обґрунтувати актуальність проблеми, висвітлити розроблені методи і їх експериментальні програмні реалізації, дослідити розроблені методи. Зокрема, у Вступі обґрунтовується актуальність теми, визначаються мета та завдання дослідження, формулюється наукова новизна та практичне значення роботи, описується зв'язок з науковими програмами.

У Розділі 1 «Аналіз методів, засобів та технологій для автоматизованого виявлення пропаганди у текстовому контенті» проведено огляд існуючих підходів до виявлення та класифікації прийомів і об'єктів пропаганди, наведено існуючі класифікації пропаганди, визначено термінологію предметної області, пропаганди та правове регулювання пропаганди в Україні.

У Розділі 2 «Метод виявлення пропаганди у текстовому контенті» наведено загальний підхід до виявлення та класифікації прийомів та об'єктів пропаганди у текстовому контенті й модель семантичної структури пропаганди, описано етичні аспекти та відповідальність впровадження концепції виявлення та класифікації прийомів та об'єктів пропаганди. Також в розділі наведено схему та кроки методу класифікації текстів за вмістом пропаганди нейромережевими моделями глибокого навчання. Визначено критерії оцінювання методів виявлення та класифікації прийомів та об'єктів пропаганди у текстовому контенті.

У Розділі 3 «Методи виявлення прийомів та об'єктів пропаганди у текстовому контенті» визначено обмеження методів виявлення прийомів та об'єктів пропаганди у текстовому контенті. Наведено схему та кроки методу виявлення прийомів пропаганди за маркерами із візуальною інтерпретацією прийнятих рішень та методу виявлення об'єктів пропаганди з візуальною інтерпретацією прийнятих рішень. Також в розділі наведено опис датасетів, що були використані для дослідження.

У Розділі 4 «Експериментальне дослідження методів виявлення та класифікації прийомів та об'єктів пропаганди у текстовому контенті» наведено опис розроблених застосувань для класифікації текстів за вмістом пропаганди й для виявлення прийомів і об'єктів пропаганди нейромережевими моделями з візуальною інтерпретацією прийнятих рішень. Також в розділі наведено результати досліджень розроблених методів: методу класифікації текстів за вмістом пропаганди нейромережевими моделями глибокого навчання, методу виявлення прийомів пропаганди за маркерами із візуальною інтерпретацією прийнятих рішень та методу виявлення об'єктів пропаганди нейромережевими моделями глибокого навчання з візуальною інтерпретацією прийнятих рішень.

Основні результати роботи:

1. Проведено аналіз методів, засобів та технологій для автоматизованого виявлення пропаганди у текстовому контенті.
2. Удосконалено метод класифікації текстів за вмістом пропаганди нейромережевими моделями глибокого навчання.
3. Розроблено метод виявлення прийомів пропаганди за маркерами із візуальною інтерпретацією прийнятих рішень.
4. Розроблено метод виявлення об'єктів пропаганди нейромережевими моделями глибокого навчання з візуальною інтерпретацією прийнятих рішень.
5. Розроблено інтелектуальну інформаційну систему для валідації запропонованих методів і проведено експериментальні дослідження.

Наукова новизна, оцінка обґрунтованості наукових положень дисертації та їх достовірності. Авторка претендує на наукову новизну наступних отриманих результатів:

- удосконалено метод класифікації текстів за вмістом пропаганди нейромережевими моделями глибокого навчання, який відрізняється від існуючих модифікованою архітектурою нейромережі та обсягом вхідних даних, що дало змогу підвищити точність класифікації;
- розроблено новий метод виявлення прийомів пропаганди за маркерами із візуальною інтерпретацією прийнятих рішень, який відрізняється від існуючих

використанням доповненої множини маркерів для виявлення прийомів пропаганди, що дало змогу пояснювати отримані результати та підвищити точність і якість виявлення пропаганди;

– розроблено новий метод виявлення об'єктів пропаганди нейромережевими моделями глибокого навчання з візуальною інтерпретацією прийнятих рішень, який відрізняється від існуючих групуванням об'єктів пропаганди, що дало змогу покращити результати виявлення об'єктів пропаганди та візуально їх інтерпретувати.

Обґрунтованість і достовірність наукових результатів, висновків та рекомендацій, наведених у дисертації, досягається ретельним аналізом існуючих результатів у сфері наукових досліджень. Коректне використання методів дослідження, математичного апарату й програмних результатів підтверджується результатами числових експериментів.

Теоретичне та практичне значення одержаних результатів. Теоретичне значення роботи полягає у розробці та обґрунтуванні нового підходу до виявлення та класифікації прийомів та об'єктів пропаганди у текстовому контенті, що полягає в трьохетапному аналізі текстового контенту на виявлення прихованих маніпулятивних прийомів та пропагандистських об'єктів. За такого підходу, спершу виконується класифікація текстів за вмістом пропаганди нейромережевими моделями глибокого навчання, й лише у випадку перевищення оцінки наявності пропаганди у тексті вище порогового значення, виконується подальший аналіз, який полягає у виявленні прийомів та об'єктів пропаганди. При цьому, спершу виконується виявлення маркерів пропаганди, за якими визначаються застосовані в тексті прийоми пропаганди, з подальшою візуальною інтерпретацією прийнятих рішень. Використання доповненої множини маркерів для виявлення прийомів пропаганди дало змогу деталізовано пояснювати отримані результати та підвищити точність і якість виявлення пропаганди. Після визначення застосованих у тексті прийомів пропаганди виконується виявлення та групування відповідних об'єктів пропаганди нейромережевими моделями глибокого навчання з візуальною інтерпретацією прийнятих рішень.

Практичне значення отриманих результатів полягає у доведенні теоретичних результатів дисертаційної роботи до реалізації та у безпосередньому використанні їх у виробничій діяльності підприємств. Реалізована система комплексного виявлення та класифікації прийомів та об'єктів пропаганди у текстовому контенті засобами штучного інтелекту. Розроблена інтелектуальна система надає можливість автоматизованого аналізу

текстів, забезпечуючи комплексний результат у вигляді: загальної оцінки прояву пропаганди у тексті; виявлених пропагандистських прийомів із числовими характеристиками сили їх прояву; оцінки приналежності виявлених об'єктів пропаганди до використаних прийомів. Також інтелектуальна система дозволяє отримувати візуальну інтерпретацію прийнятих рішень, що сприяє підвищенню прозорості та довіри до отриманих результатів.

Повнота викладу результатів дисертації в опублікованих працях.

Основні результати дисертації достатньо повністю наведено у 4 наукових статтях у виданнях, включених до Переліку наукових фахових видань України, які, на дату опублікування, віднесені до категорії «Б». Три статті одноосібні, тому згідно Підпункту 1 пункту 8 в редакції Постанови КМ № 507 від 03.05.2024, зараховуються повністю. Одна стаття має 2 (два) співавтори, тому згідно наведеної вище постанови, також зараховується повністю.

Виходячи з наведеного вище, здобувач має 4 публікації, у яких викладені основні результати дисертації, чого достатньо, згідно чинних вимог. Зазначу, що для кожної публікації, де автор мав співавторів, чітко вказаний його особистий внесок.

Додатково положення роботи апробовано на 5 Міжнародних конференціях, за їх результатами виконано 5 (п'ять) публікацій, що індексуються в міжнародній наукометричній базі "Scopus". У кожній спільній публікації чітко зазначено особистий внесок дисертанта, а загальна кількість та рівень джерел повністю відповідають вимогам до дисертаційних робіт.

Зауваження та побажання.

1. Інтерфейс експериментальних застосунків має елементи як на українській, так і на англійській, що видно на рис. 4.1, 4.2, 4.3. Було б коректним забезпечити елементи інтерфейсу на українській.

2. В роботі мета та завдання дослідження (п.1.7) та висновки до розділу 1 (п.1.8) наведено як окремі пункти, хоча за логікою висвітлення роботи їх доцільно об'єднати й розмістити в зворотному порядку, визначивши мету та завдання дослідження за результатом обговорення першого розділу роботи.

3. Хоча мовою дисертації є українська, проте в Переліку скорочень не для всіх англійських термінів і скорочень наведено українськомовний аналог. Наприклад, не наведено для скорочень: RNN, NLP, LLM.

4. У п.1.6 відомості щодо маркерів пропаганди та їх взаємозв'язку з прийомами пропаганди викладені в вигляді тексту, що ускладнює розуміння. Кращим варіантом була б систематизація цих відомостей у табличному вигляді.

5. У п.2.6 частина матеріалу відведена опису критеріїв оцінювання методів виявлення та класифікації прийомів та об'єктів пропаганди у текстовому контенті, які базуються на відомих формулах. Наприклад, формули обрахунку точності (Accuracy) та влучності (Precision). Більш коректно було б ці формули винести у Розділ 1, або навести посилання на джерела, звідки вони взяті.

6. З метою підвищення точності виявлення пропаганди запропоновано нейромережеву модель BiLSTM гібридної архітектури, наведену в п 2.5.4. Проте в п.2.5.3 розглянуто особливості використання нейромережевих моделей BiLSTM та GRU для класифікації текстів за вмістом пропаганди. Це пояснюється необхідністю порівняння запропонованих рішень з існуючими аналогами, очевидно, що наведені аспекти виходять за межі поставленої мети і є надлишковими в роботі.

7. У Таблиці 3.1 наведено приклади типових показників прояву маркерів для 10 прийомів пропаганди. Однак у роботі загалом розглядається 17 прийомів пропаганди. Для повноти відображення, було б корисним навести типові показники прояву маркерів для всіх 17 прийомів пропаганди.

8. В роботі виявлено ряд технічних помилок граматичного та синтаксичного характеру. Наприклад, на стор. 143 в рядку 12 речення завершено комою замість крапки.

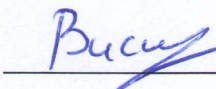
Однак вказані зауваження не є суттєвими, переважно відносяться до оформлення результатів та істотно не впливають на загальну позитивну оцінку дисертаційної роботи, не знижують її наукову й практичну цінність.

Загальний висновок. Подана дисертаційна робота є завершеною науково-дослідною роботою, в якій розроблено системний підхід до автоматизованого виявлення прийомів та об'єктів пропаганди, що дозволяє комплексно аналізувати взаємозв'язки виявлених прийомів і об'єктів. Вирішується суперечність між необхідністю ефективного виявлення пропаганди та недосконалістю існуючих методів і засобів, які не враховують комплексний аналіз взаємозв'язків прийомів і об'єктів пропаганди. Розроблені методи і засоби сприятимуть підвищенню точності автоматизованого виявлення пропаганди. Тема і зміст дисертації відповідає Стандарту спеціальності 122 – «Комп'ютерні науки» для третього рівня вищої освіти.

Оцінюючи дисертаційну роботу в цілому, є всі підстави стверджувати, що за актуальністю теми, обсягом виконаних досліджень, науковою новизною, цінністю одержаних результатів і науково-теоретичним рівнем їх обґрунтованості робота цілком відповідає вимогам пунктів 6, 7, 8, 9 «Порядку

присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії», затвердженого Постановою Кабінету Міністрів України від 12 січня 2022 р. №44 (зі змінами, внесеними згідно з Постановами Кабінету Міністрів України № 341 від 21.03.2022, № 502 від 19.05.2023, № 507 від 03.05.2024), а її авторка Молчанова Марина Олексіївна заслуговує на присудження їй ступеня доктора філософії за спеціальністю 122 – «Комп'ютерні науки» в галузі знань 12 – «Інформаційні технології».

Офіційний опонент,
доктор технічних наук, доцент,
професор кафедри інформаційних систем та мереж
Інституту комп'ютерних наук
та інформаційних технологій
Національного університету
«Львівська політехніка»

 Вікторія ВИСОЦЬКА

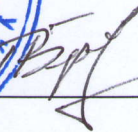
«Підпис Висоцької В.А. засвідчую»:

Вчений секретар

Національного університету

«Львівська політехніка», к.т.н., доцент





Роман БРИЛИНСЬКИЙ