**РЕЦЕНЗІЯ**

на дисертаційну роботу

Падалко Галини Анатоліївни

на тему “Моделі, методи та інформаційна технологія виявлення і аналізу текстової дезінформації та пропаганди в соціальних мережах”,

яка представлена на здобуття ступеня доктора філософії

в галузі знань 12 Інформаційні технології за спеціальністю 122 Комп’ютерні науки

**Актуальність теми дисертації**

У контексті розвитку соціальних мереж дезінформація та пропагандистські наративи стали ключовими інструментами негативного інформаційного впливу на суспільство. Особливо важливу роль у цьому процесі відіграють соціальні медіа, які сприяють стрімкому поширенню маніпулятивного контенту та ускладнюють ідентифікацію джерел інформаційних атак. Згадані явища досягли рівня повноцінних інформаційних кампаній для підриву довіри до державних інституцій, поглиблення внутрішніх конфліктів і дестабілізації суспільства. Ці практики регулярно фіксуються в аналітичних звітах міжнародних організацій і дослідницьких установ. У зв’язку з цим виникає гостра потреба у впровадженні автоматизованих рішень на основі штучного інтелекту, здатних ефективно аналізувати великі обсяги даних та виявляти мережі дезінформації. Актуальність обраної теми зумовлена необхідністю посилення інформаційної безпеки та підвищення стійкості суспільства до зовнішнього інформаційного впливу.

**Оцінка обґрунтованості наукових результатів дисертації, їх достовірності та новизни**

Матеріали, викладені у дисертації Падалко Галини Анатоліївни, дозволяють зробити висновок про обґрунтованість та достовірність її наукових результатів. Вони забезпечуються використанням сучасних методів дослідження, зокрема на теоріях обчислювального штучного інтелекту, а саме – машинного навчання та глибокого навчання, що дозволяють класифікувати та кластеризувати великі набори даних. З урахуванням проведеного аналізу в роботі поставлено та вирішено наукове завдання розробки технології дослідження текстів для аналізу маніпуляцій інформацією.

Розроблені моделі, методи та програмне забезпечення використані у Balsillie School of International Affairs (акт впровадження від 18 березня 2025), Center for International Governance Innovation (акт впровадження від 18 березня 2025), Державній установі “Харківський обласний центр контролю та профілактики хвороб при Міністерстві охорони здоров’я” (акт впровадження від 2 грудня 2024), а також у освітньому процесі (акт впровадження від 30 січня 2025) та науковій діяльності (акт впровадження від 4 березня 2025) Національного аерокосмічного університету “Харківський авіаційний інститут”.

Наукова новизна результатів дисертаційного дослідження полягає в наступному:

1. Вперше запропоновано фреймворк комплексного аналізу текстової дезінформації та пропаганди, заснований на комбінуванні технологій двоспрямованого кодувального представлення з трансформерів (BERT) та двоспрямованої довгої короткочасної пам’яті (Bi-LSTM) для створення інтерпретованих векторних представлень тексту, який на відміну від існуючих технік, об’єднує виявлення та тематичний аналіз дезінформації, що дозволяє забезпечити ефективне виявлення та аналіз текстової дезінформації та пропаганди в соціальних мережах.

2. Вперше розроблено модель класифікації текстової дезінформації та пропаганди, заснований на поєднанні трансформерів і архітектури Bi-LSTM для інтеграції лінійних та нелінійних залежностей текстових даних, яка на відміну від існуючих використовує глибокі контекстуальні ембединги і механізми уваги, що дозволяє підвищити точність моделі.

3. Удосконалено моделі глибокого навчання для класифікації текстової дезінформації та пропаганди, засновані на архітектурах XLNet, Bi-LSTM та Attention-Based Bi-LSTM, які на відміну від існуючих застосовують адаптивне оцінювання важливих фрагментів тексту, що дозволяє забезпечувати високий рівень продуктивності класифікації.

4. Удосконалено метод тематичного моделювання, заснований на застосуванні механізмів багатоголової уваги, що на відміну від існуючих використовує глибокі ембединги для контекстуалізації даних, що дозволяє забезпечити чітке розмежування між кластерами.

5. Дістали подальшого розвитку моделі для класифікації текстової дезінформації та пропаганди, засновані на ансамблевих методах машинного навчання, які на відміну від існуючих враховують балансування класів і адаптуються до змін вхідних даних, що дозволяє підвищити продуктивність моделей.

6. Дістали подальшого розвитку моделі для класифікації інформації, засновані на статистичних методах машинного навчання, які на відміну від існуючих адаптовані до класифікації текстової дезінформації та пропаганди, що дозволяє створення систем для виявлення місінформації, дезінформації та малінформації.

**Оцінка змісту дисертації, її завершеність та дотримання принципів академічної доброчесності**

За своїм змістом дисертаційна робота здобувачки Падалко Галини Анатоліївни повністю відповідає Стандарту вищої освіти за спеціальністю 122 Комп’ютерні науки. Дисертаційна робота є завершеною науковою працею, а її результати свідчать про вагомий особистий внесок здобувачки у науковий напрям.

Можна зробити висновок, що дисертаційна робота Падалко Галини Анатоліївни є результатом самостійних досліджень здобувачки і не містить елементів фальсифікації, компіляції, фабрикації, плагіату та запозичень. Використані ідеї, результати і тексти інших авторів мають належні посилання на відповідне джерело.

**Мова та стиль викладення результатів**

Дисертаційна робота написана українською мовою, котра відповідає особливостям стилю наукових досліджень з використанням правильних термінів і понять, характерних для стандартної фахової та наукової термінології. Дисертаційна робота є цілісною і в ній досягнуті тематична повнота та розкриття головної наукової ідеї автора. Подані в роботі наукові та практичні положення логічно викладені та достатньо обґрунтовані.

Дисертація складається з анотації, змісту, переліку умовних скорочень, вступу, чотирьох розділів, висновку, списку використаних джерел та додатків. Повний обсяг роботи становить 207 сторінок друкованого тексту, з них анотація – на 7 стор., зміст – на 4 стор., перелік умовних скорочень – на 2 стор., основний текст – на 124 стор., список із 151 використаних джерел – на 17 стор., додатки – на 22 стор. Дисертація містить 54 рисунки, та 16 таблиць.

**Мова та структура дисертації**

Робота виконана українською мовою, витримана в науковому стилі, з логічною послідовністю викладу та коректним використанням термінології. Оформлення відповідає чинним вимогам. Дисертація складається з анотації, змісту, списку скорочень, вступу, чотирьох розділів, висновків, списку джерел та додатків. Загальний обсяг — 207 сторінок, зокрема 54 рисунки та 16 таблиць.

**Зміст дисертаці**ї

У вступі обґрунтовано актуальність теми, сформульовано мету, завдання, об’єкт і предмет дослідження, окреслено наукову новизну, методологічні підходи, ступінь апробації та зв’язок із науковими програмами.

**Розділ 1** присвячений теоретичному аналізу дезінформації, місінформації та малінформації в умовах гібридної війни, з акцентом на їх використання державними акторами (зокрема Росією) через соціальні мережі й ШІ. Описано сучасні методи виявлення інформаційних загроз.

**Розділ 2** містить порівняльний аналіз алгоритмів машинного навчання (SVM, LightGBM тощо) для виявлення дезінформації. Найкращі результати продемонстровані на корпусах WELFake і Weibo, із точністю до 89% та F1-мірою 0.977.

**Розділ 3** присвячено глибоким моделям (BERT, XLNet, BiLSTM тощо) та запропонованій гібридній архітектурі з attention-механізмом, яка досягла точності 99.47% на різних датасетах.

**Розділ 4** описує інтегрований фреймворк, що поєднує класифікацію та тематичне моделювання (BERTopic) для виявлення ключових наративів. Модель виявляє ідеологічні патерни з високою точністю (до 99.47%).

**У висновках** підсумовано наукові й практичні результати, окреслено напрями подальших досліджень.

**Додатки** включають список публікацій, дані про апробацію, акти впровадження та опис використаних датасетів.

Дисертаційна робота оформлена відповідно до вимог, сформульованих у наказі МОН України від 12 січня 2017р. № 40 «Про затвердження вимог до оформлення дисертації».

**Повнота викладення результатів досліджень в опублікованих працях**

За матеріалами дисертації опубліковано 16 наукових праць, розроблені моделі, методи та програмне забезпечення використані у Balsillie School of International Affairs (акт впровадження від 18 березня 2025), Center for International Governance Innovation (акт впровадження від 18 березня 2025), Державній установі “Харківський обласний центр контролю та профілактики хвороб при Міністерстві охорони здоров’я” (акт впровадження від 2 грудня 2024), а також у освітньому процесі (акт впровадження від 30 січня 2025) та науковій діяльності (акт впровадження від 4 березня 2025) Національного аерокосмічного університету “Харківський авіаційний інститут”.

Наукові публікації здобувачки містять опис наукових досліджень, проведених в рамках дисертаційного пошуку, аналіз сутності проблеми, методів і результатів проведених досліджень, а також обґрунтовані висновки. В наукових публікаціях здобувачки не виявлено порушень принципів академічної доброчесності, висновки є оригінальними.

Таким чином, наукові результати, описані в дисертаційній роботі, повністю висвітлені у наукових публікаціях здобувачки.

**Недоліки та зауваження до дисертаційної роботи**

По дисертаційній роботі можна зробити наступні зауваження:

1. Матриця невідповідностей (рис. 2.6) фіксує помітну кількість помилок, попри загальну F1≈0.94

2. У Розділі 4 описано оптимізації BERTopic, але якість тем оцінюється через візуальні кластери й вручну відібрані ключові слова.

3. Здобувачка деталізує, що зміщувала n\_neighbors і min\_dist UMAP до «найкращих» значень 15 та 0,1, але не показує, як змінювалась когерентність тем при інших комбінаціях.

4. Розділ 3.2.2 розбиває корпус твітів російської пропаганди випадковою пропорцією 80 / 10 / 10, не враховуючи часову послідовність постів. За умов інформаційних кампаній помітний «концепт-дрейф», тож випадкове перетинання старих і нових повідомлень у тренуванні й тесті може штучно завищити метрики.

5. У роботі присутні граматичні та стилістичні помилки.

Вважаю, що висловлені зауваження не є визначальними і не зменшують загальну наукову новизну та практичну значимість результатів та не впливають на позитивну оцінку дисертаційної роботи.

Висновок про дисертаційну роботу

Вважаю, що дисертаційна робота здобувачки ступеня доктора філософії Падалко Галини Анатоліївни на тему “Моделі, методи та інформаційна технологія виявлення і аналізу текстової дезінформації та пропаганди в соціальних мережах” виконана на високому науковому рівні, не порушує принципів академічної доброчесності та є закінченим науковим дослідженням, сукупність теоретичних та практичних результатів якого розв'язує наукове завдання, що має істотне значення для галузі знань 12 Інформаційні технології. Дисертаційна робота за актуальністю, практичною цінністю та науковою новизною повністю відповідає вимогам чинного законодавства України, що передбачені в Наказі МОН України Nº40 від 12.01.2017 «Про затвердження вимог щодо оформлення дисертації ( зі змінами)» та «Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти , наукової установи про присудження ступеня доктора філософії », затвердженого Постановою КМУ від 12.01.2022 Nº44.

Здобувачка Падалко Галина Анатоліївна заслуговує на присудження ступеня доктора філософії в галузі знань 12 Інформаційні технології за спеціальністю 122 Комп’ютерні науки.

Рецензент:

кандидат технічних наук, доцент,

доцент кафедри

інформаційно-комунікаційних технологій

ім. О. О. Зеленського

Національного аерокосмічного університету

«Харківський авіаційний інститут» Олексій РУБЕЛЬ