

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Західноукраїнський національний університет
Факультет комп'ютерних інформаційних технологій
Кафедра інформаційно-обчислювальних систем і управління

МАТВІЙЧУК Микола Олександрович

**Модель інтеграції емоційних, семантичних та доменних
ознак для виявлення фейкових новин / A Model for
Integrating Emotional, Semantic, and Domain Features for
Fake News Detection**

спеціальність: 122 - Комп'ютерні науки
освітньо-професійна програма - Комп'ютерні науки

Кваліфікаційна робота

Виконав студент групи КНм-21
Матвійчук М.О

Науковий керівник:
к.т.н., доцент, Ліп'яніна-
Гончаренко Х.В.

Кваліфікаційну роботу
допущено до захисту:
«___» _____ 20___ р.
В.о. завідувача кафедри
_____ Н.М. Васильків

ТЕРНОПІЛЬ - 2024

Факультет комп'ютерних інформаційних технологій
Кафедра інформаційно-обчислювальних систем і управління
Освітній ступінь «магістр»
спеціальність: 122 – Комп'ютерні науки
освітньо-професійна програма – Комп'ютерні науки

ЗАТВЕРДЖУЮ
Завідувач кафедри
М.П. Комар
« ____ » _____ 20__ р.

**ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ СТУДЕНТУ
МАТВІЙЧУК Микола Олександрович**

(прізвище, ім'я, по батькові)

1. Тема кваліфікаційної роботи

Модель інтеграції емоційних, семантичних та доменних ознак для виявлення
фейкових новин / A Model for Integrating Emotional, Semantic, and Domain
Features for Fake News Detection

керівник роботи к.т.н., доцент, Ліп'яніна-Гончаренко Х.В.

затверджені наказом по університету від 12 грудня 2023 року № 753.

2. Строк подання студентом закінченої кваліфікаційної роботи 4 грудня 2024 р.

3. Вихідні дані до кваліфікаційної роботи: завдання на кваліфікаційну роботу студента, наукові статті, технічна література.

4. Основні питання, які потрібно розробити

- Проаналізувати поширення фейкових новин у соціальних медіа, враховуючи сучасні тенденції та виклики.
- Оглянути існуючі рішення щодо виявлення фейкових новин, зокрема методи машинного навчання та глибокого навчання.
- Розробити архітектуру моделі для інтеграції емоційних, семантичних та доменних ознак з метою виявлення фейкових новин.
- Моделювати контекстуальну релевантність у новинних текстах для покращення точності класифікації.
- Дослідити вплив емоційних ознак на поширення фейкових новин та їх інтеграцію в модель.
- Розробити методи видобування доменних ознак для покращення міждоменної релевантності в новинних текстах.
- Провести експерименти з оцінюванням точності, повноти та F1-міри на основі різних наборів даних, включаючи абляційний аналіз.

5. Перелік графічного матеріалу у роботі

- Загальна архітектурна схема моделі DESF.

6. Консультанти розділів кваліфікаційної роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		Завдання видав	Завдання прийняв

7. Дата видачі завдання 12 грудня 2023 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів кваліфікаційної роботи	Примітка
1	Затвердження теми кваліфікаційної роботи, ознайомлення з літературними джерелами та складання плану роботи.	до 01.01. 2024 р.	
2	Написання 1 розділу кваліфікаційної роботи	до 01.03. 2024 р.	
3	Написання 2 розділу кваліфікаційної роботи	до 20.05.2024 р.	
4	Написання 3 розділу кваліфікаційної роботи	до 28.10. 2024 р.	
5	Представлення попереднього варіанту кваліфікаційної роботи, перевірка та внесення змін керівником	до 11.11.2024 р.	
6	Опрацювання зауважень та представлення завершеного варіанту кваліфікаційної роботи. Підготовка супроводжуючих документів.	до 25.11.2024 р.	
7	Перевірка кваліфікаційної роботи на оригінальність тексту.	до 30.11.2024 р.	
8	Оформлення кваліфікаційної роботи та отримання допуску до захисту	до 04.12.2024 р.	
9	Подання кваліфікаційної роботи до захисту на засіданні атестаційної комісії.	до 14.12. 2024 р.	

Студент _____ Матвійчук М.О
підпис

Керівник роботи _____ к.т.н., доцент, Ліп'яніна-Гончаренко Х.В.
підпис

РЕЗЮМЕ

Кваліфікаційна робота на тему «Модель інтеграції емоційних, семантичних та доменних ознак для виявлення фейкових новин» на здобуття освітнього ступеня «Магістр» зі спеціальності 122 «Комп'ютерні науки» освітньої програми «Комп'ютерні науки» написана обсягом в 64 сторінок і містить 8 ілюстрацій, 9 таблиць, 1 додатки та 36 використаних джерела.

Метою даної кваліфікаційної роботи є розробка моделі інтеграції емоційних, семантичних та доменних ознак для підвищення точності та ефективності виявлення фейкових новин у багатодомених та багатомовних інформаційних потоках.

Методи дослідження: аналіз літератури, використання моделей глибокого навчання, таких як BiLSTM, механізми уваги для роботи з багатодоменими даними, аналіз емоційних ознак, експериментальне тестування моделей на реальних наборах даних.

Результати дослідження: запропоновано нову модель DESF, яка поєднує семантичні, емоційні та доменні ознаки для точного виявлення фейкових новин. Експерименти на наборах даних Weibo21 та FakeNewsNet підтвердили переваги моделі порівняно з існуючими підходами. Високі показники точності (до 92.7% на Weibo21) демонструють ефективність DESF у боротьбі з дезінформацією.

Практичне значення роботи: результати можуть бути застосовані для автоматизованого виявлення фейкових новин у соціальних мережах, покращення систем моніторингу інформаційного простору, забезпечення надійності новинних платформ.

Ключові слова: ФЕЙКОВІ НОВИНИ, ЕМОЦІЙНІ ОЗНАКИ, ГЛИБОКЕ НАВЧАННЯ, BiLSTM, УВАГА, СЕМАНТИЧНІ ОЗНАКИ.

ABSTRACT

The qualification thesis titled "Model for Integrating Emotional, Semantic, and Domain Features for Fake News Detection" for obtaining the Master's degree in the specialty 122 "Computer Science" of the educational program "Computer Science" consists of 64 pages, including 8 illustrations, 9 tables, 1 appendix, and 36 references.

The purpose of this thesis is to develop a model for integrating emotional, semantic, and domain features to enhance the accuracy and efficiency of fake news detection in multi-domain and multilingual information streams.

Research methods: literature analysis, application of deep learning models such as BiLSTM, attention mechanisms for multi-domain data, emotional feature analysis, and experimental testing of models on real-world datasets.

Research results: A new model, DESF, was proposed, combining semantic, emotional, and domain features for precise fake news detection. Experiments on datasets such as Weibo21 and FakeNewsNet demonstrated the model's superiority over existing approaches. High accuracy rates (up to 92.7% on Weibo21) validate DESF's effectiveness in combating disinformation.

Practical significance: The results can be applied to automate fake news detection in social media, improve monitoring systems for the information space, and enhance the reliability of news platforms.

Keywords: FAKE NEWS, EMOTIONAL FEATURES, DEEP LEARNING, BiLSTM, ATTENTION, SEMANTIC FEATURES.

ЗМІСТ

Вступ.....	7
1. Аналіз предметної області і постановка задачі дослідження.....	11
1.1 Аналіз поширення фейкових новин у соціальних медіа	11
1.2 Огляд існуючих рішень.....	13
1.3 Вибір перспективного шляху і постановка задачі дослідження	16
Висновки до розділу 1	19
2. Модель інтеграція емоційних, семантичних та доменних ознак для виявлення фейкових новин.....	21
2.1 Архітектура моделі інтеграція емоційних, семантичних та доменних ознак для виявлення фейкових новин.....	21
2.2 Моделювання контекстуальної релевантності в новинних текстах	22
2.3 Вплив емоційних ознак на поширення фейкових новин	24
2.4 Видобування доменних ознак для покращення міждоменної релевантності в новинних текстах	25
2.5 Метрики оцінювання моделей класифікації та їх застосування у виявленні фейкових новин	28
Висновки до розділу 2.....	30
3 Програмно-технологічне забезпечення	32
3.1 Дослідження набору даних	32
3.2 Аналіз результатів експериментів.....	33
3.3 Експеримент із абляцією.....	40
Висновки до розділу 3.....	44
Висновки	46
Список використаних джерел	48
Додаток А Апробація отриманих результатів	53

ВСТУП

Сучасне інформаційне середовище зазнає значного впливу зростання кількості користувачів інтернету та соціальних медіа, що призводить до швидкого поширення як достовірної інформації, так і дезінформації. Станом на липень 2024 року кількість користувачів інтернету досягла 5.45 мільярда, а соціальних мереж — понад 5.17 мільярда, що становить більше половини світового населення. Така швидка цифровізація суспільства створює нові виклики у боротьбі з дезінформацією, особливо в контексті поширення фейкових новин через платформи, як-от Facebook, TikTok та YouTube. Тому дослідження методів ефективного виявлення фейкових новин є надзвичайно актуальним завданням для сучасної науки.

Фейкові новини мають здатність викликати значні соціальні, політичні та економічні наслідки, маніпулюючи думкою громадськості та створюючи соціальну напруженість. Особливо це відчутно під час кризових ситуацій, таких як пандемія COVID-19 або політичні конфлікти, коли дезінформація може призвести до хибних уявлень та поведінкових патернів у суспільстві. Алгоритми соціальних мереж сприяють швидкому розповсюдженню таких новин, підсилюючи ефект "ехо-камери" та підтверджуючи існуючі упередження користувачів. Враховуючи масштабні негативні наслідки дезінформації, актуальність розробки ефективних моделей для виявлення фейкових новин постійно зростає.

Наукова спільнота постійно вдосконалює методи виявлення фейкових новин, однак більшість існуючих рішень зосереджені на текстових ознаках і не враховують багатовимірність інформаційних потоків. Використання емоційних, доменних та семантичних ознак може значно підвищити ефективність таких моделей, оскільки фейкові новини часто використовують емоційно насичену лексику та специфічні доменні контексти для маніпуляції. Відсутність досліджень, які б у комплексі використовували ці три типи ознак для

класифікації новин, обумовлює необхідність розробки нових підходів до виявлення дезінформації.

Запропонована модель DESF, яка поєднує доменні, емоційні та семантичні ознаки, є новим кроком у напрямку підвищення точності класифікації фейкових новин. Вона дозволяє не лише більш точно оцінювати емоційний контент новин, але й брати до уваги контекст домену, що є важливим для коректного розуміння новинного тексту. Актуальність такого підходу зумовлена необхідністю розширення можливостей алгоритмів штучного інтелекту для боротьби з фейковими новинами, особливо у контексті багатомовних і багатодоменних інформаційних потоків.

Мета дослідження полягає в розробці моделі інтеграції емоційних, семантичних та доменних ознак для ефективного виявлення фейкових новин у багатодоменних та багатомовних інформаційних потоках.

Досягнення цієї мети зумовило потребу теоретичних розробок, визначення та послідовного вирішення таких завдань:

1. Проаналізувати поширення фейкових новин у соціальних медіа, враховуючи сучасні тенденції та виклики.
2. Оглянути існуючі рішення щодо виявлення фейкових новин, зокрема методи машинного навчання та глибокого навчання.
3. Розробити архітектуру моделі для інтеграції емоційних, семантичних та доменних ознак з метою виявлення фейкових новин.
4. Моделювати контекстуальну релевантність у новинних текстах для покращення точності класифікації.
5. Дослідити вплив емоційних ознак на поширення фейкових новин та їх інтеграцію в модель.
6. Розробити методи видобування доменних ознак для покращення міждоменої релевантності в новинних текстах.
7. Провести експерименти з оцінюванням точності, повноти та F1-міри на основі різних наборів даних, включаючи абляційний аналіз.

Об'єктом дослідження є процес виявлення фейкових новин у соціальних медіа з використанням методів машинного навчання та глибокого навчання.

Предметом дослідження є інтеграція емоційних, семантичних та доменних ознак для покращення точності класифікації фейкових новин.

Методи дослідження включають аналіз та узагальнення літературних джерел для вивчення існуючих підходів до виявлення фейкових новин, розробку архітектури моделі для інтеграції емоційних, семантичних та доменних ознак, експериментальне моделювання та тестування моделі на наборах даних Weibo21 і FakeNewsNet, а також абляційні експерименти для визначення внеску кожного компоненту моделі. Методи оцінювання ефективності базуються на метриках точності, повноти, F1-міри та їх порівнянні з результатами інших моделей.

Наукова новизна отриманих результатів полягає в розробці моделі, яка вперше поєднує емоційні, семантичні та доменні ознаки для виявлення фейкових новин, перевершуючи існуючі рішення за точністю і F1-мірою завдяки синергії цих ознак, що дозволяє досягти кращих результатів порівняно з мультимодальними та доменними підходами, представленими в попередніх дослідженнях.

Практичне значення отриманих результатів полягає в можливості застосування розробленої моделі для автоматизованого виявлення фейкових новин у реальних умовах соціальних медіа, що дозволяє значно підвищити точність та ефективність класифікації новинного контенту. Модель може бути використана для моніторингу інформаційного простору, боротьби з дезінформацією та покращення якості новинних платформ, зокрема для більш точного і швидкого виявлення неправдивих повідомлень у різних доменах новин, таких як політика, економіка, здоров'я та суспільство.

Структура та обсяг роботи. Кваліфікаційна робота складається зі вступу, трьох основних розділів, висновків, списку літератури та додатків.

Апробація результатів дослідження. Основні теоретичні положення роботи й практичні результати дослідження доповідалися й обговорювалися: VI

Міжнародній мультидисциплінарній студентській науковій конференції «Тренди та перспективи розвитку мультидисциплінарних досліджень», яка відбулася 11 жовтня 2024 року у місті Кременчук, Україна; III Міжнародній науковій конференції «Інтелектуальний ресурс сьогодення: наукові задачі, розвиток та запитання», яка відбулася 20 вересня 2024 року у місті Одеса, Україна.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ І ПОСТАНОВКА ЗАДАЧІ ДОСЛІДЖЕННЯ

1.1 Аналіз поширення фейкових новин у соціальних медіа

Станом на липень 2024 року, кількість користувачів Інтернету в усьому світі досягла 5.45 мільярда осіб, що становить близько 67.1% всього населення планети. У той же час кількість користувачів соціальних мереж перевищила 5.17 мільярда, що становить більше половини світового населення. Це демонструє постійне зростання проникнення інтернету та використання соціальних медіа, хоча в деяких регіонах, таких як Центральна Африка, рівень проникнення залишається нижче 50% через економічні та інфраструктурні обмеження.

Що стосується поширення фейкових новин, це стало глобальною проблемою через зростання використання соціальних платформ, таких як Facebook, TikTok, та YouTube. Соціальні мережі сприяють швидкому розповсюдженню дезінформації через алгоритми, які підсилюють "ефект ехо-камери," що призводить до обмеженого сприйняття альтернативних поглядів і підтвердження вже існуючих упереджень.

Дослідження показують, що на соціальних платформах особливо активно поширюються фейкові новини, які емоційно маніпулюють читачами, викликаючи негативні наслідки для суспільства. Таким чином, виявлення та боротьба з фейковими новинами залишається ключовою проблемою в інформаційному просторі у 2024 році.

Фейкові новини, такі як "ефірна олія пригнічує вірусні інфекції", "п'ять яєць на день запобігають застуді" і "мешканці виходять на вулицю голити голови під час пандемії", стали поширюватися все частіше через перенасичення інтернету інформацією. Фейкові новини, які швидко розповсюджуються, часто містять довільну та запальну мову, майстерно оформлені як "клікбейт", характеризуються "відсутністю деталей" і "підбурюванням емоцій", що може ввести людей в оману та мати серйозні негативні наслідки для суспільства.

Поширення фейкових новин поглиблюється через "ефект ехо-камери" в соціальних мережах. Тому виявлення фейкових новин стало важливою темою в галузі обробки інформації [1].

Виявлення фейкових новин часто визначається як задача класифікації, і виділення ознак є ключовим етапом для покращення продуктивності [2]. Тому дослідження взаємозв'язку між різними характеристиками також стало предметом постійного вивчення серед науковців. Джін та інші [3] вперше запропонували att-RNN, яка інтегрує різні патерни ознак, включаючи текстові, візуальні та портрети користувачів, і коригує їх за допомогою механізму уваги. Хаттар та інші [4] запропонували Мультимодальний Варіаційний Автоенкодер (MVAE), який навчає користувачів знаходити кореляції між різними патернами в конкретному твіті. Модель, запропонована Цзеню та ін. [5], не тільки досліджує семантичний зв'язок між текстовим контентом і супровідними зображеннями, але й використовує механізм уваги для виділення важливих частин новинних статей.

Технологія асоціативних правил широко використовується в соціальних медіа завдяки своїй здатності узагальнювати інформацію. Вихід цього алгоритму легко пояснити та застосувати в різних галузях. Делгадо та інші [6] запропонували одне з перших досліджень щодо вирішення проблеми текстових асоціативних правил, яке визначило текстові транзакції та застосувало нечіткі асоціативні правила до них. Інше цікаве дослідження, проведене Феном та ін. [7], досліджувало взаємозв'язок між настроєм, кольором та соціальними мітками через асоціативні правила. У цьому дослідженні асоціативні правила використовувалися для кодування зображень, зв'язуючи кольори пікселів з конкретними емоціями. Шейдаєві та інші [8] запропонували асоціативний алгоритм класифікації, який включав етапи вибору ознак для відбору ключових характеристик і кластеризації на основі класових міток. Рохідін та інші [9] запропонували метод класифікації тексту, що поєднує асоціативний метод із нечіткими м'якими множинами, що дало високі результати класифікації. Цай та

інші [10] розробили мультидоменну нейро-емоційну модель класифікації, яка показала взаємозв'язок між доменами та емоціями. Ці дослідження демонструють, що пов'язуючи об'єктні характеристики у соціальних медіа та використовуючи різні дані для покращення продуктивності класифікації, можна краще розуміти задачі класифікації.

1.2 Огляд існуючих рішень

1.2.1 Виявлення фейкових новин

Виявлення фейкових новин стало актуальною сферою досліджень останніх років, спрямованою на боротьбу з поширенням дезінформації в інтернеті. Традиційні методи машинного навчання, такі як Наївний Байєс, Метод опорних векторів (SVM) та Random Forest, широко застосовувалися для цієї мети. Дослідники використовували різні характеристики, такі як текстові ознаки, соціальні медіа та семантичні ознаки, для покращення точності виявлення. У міру прогресу досліджень методи глибокого навчання, такі як CNN та RNN, також досягли певних результатів.

Джі та інші використовували графи знань для виявлення фейкових новин, створюючи базу знань, яка включала зв'язки в поширенні інформації, знання про сутності та часові рамки подій, і поєднали це з текстовими ознаками для аналізу [11]. Ліу та колеги більше уваги приділяли злиттю мультимодальних ознак, таких як текст, зображення та відео, що дозволяло отримати більш повну інформацію і підвищити точність моделі [12].

Ураховуючи постійний розвиток творців фейкових новин, Лі та інші застосовували Генеративно-змагальні мережі (GAN) для підвищення здатності моделей протистояти різним видам дезінформації [13]. Саху та інші аналізували поведінку Facebook-акаунтів, використовуючи новинний контент та інші ознаки з допомогою глибокого навчання [14]. Чжан та інші запропонували модель на основі глибокого навчання для швидкого виявлення фейкових новин у

кіберфізичних соціальних службах, використовуючи китайський текст як базовий об'єкт [15].

Незважаючи на виклики у виявленні фейкових новин, дослідники постійно працюють над покращенням точності та надійності моделей.

1.2.2 Злиття ознак

Злиття ознак передбачає комбінування кількох ознак або наборів ознак для створення більш інформативного та дискримінативного подання даних. Це дозволяє покращити загальну продуктивність моделі шляхом використання додаткової інформації, яку надають різні ознаки [16]. Поширені методи злиття ознак включають конкатенацію, де ознаки об'єднуються для формування довшого вектора ознак; зважену суму, коли різні ознаки комбінуються з вагами, що відображають їхню важливість; та складніші методи, такі як злиття на основі ядра або глибокого навчання. Завдяки злиттю ознак можна поєднувати кілька ознак для створення більш дискримінативного подання даних, що відіграє ключову роль у покращенні продуктивності прогностичних моделей і отриманні змістовної інформації з сирих даних.

Ван та інші [17] запропонували нову детальну мультимодальну мережу злиття ознак (FMFN) для повного поєднання текстових і візуальних ознак для виявлення фейкових новин. Лі та інші [18] запропонували нову структуру під назвою семантично-посилена мультимодальна мережа злиття, яка краще вловлює взаємні ознаки подій, що підвищує точність виявлення фейкових новин. Го та інші [19] запропонували нейронну мережу з взаємною увагою (MANN), яка може вивчати взаємозв'язки між різними модальностями.

На основі різних джерел інформації в мережі, фейкові новини можна поділити на кілька доменів (політика, історія, культура тощо), кожен з яких має свої характеристики [20]. Зазвичай достовірна інформація більш об'єктивна та нейтральна за настроєм, тоді як фейкові новини часто поширюються в перебільшеній, пропагандистській манері з вищим рівнем емоційного

напруження. За допомогою емоційних ознак можна аналізувати інтенсивність емоцій, полярність настрою та інші характеристики тексту, щоб визначити, чи є новина фейковою. Семантичні ознаки можуть виявити семантичні зв'язки між текстами. Однак окремі емоційні або семантичні ознаки можуть бути обмеженими через точність анотації та витягання даних, тому злиття ознак покращує точність виявлення. Інноваційні пари різних методів зображені у таблиці 1.1.

Таблиця 1.1 - Порівняння інноваційних аспектів різних методів

Метод	Інноваційний аспект
Ji et al. [11]	Використання графів знань
Liu et al. [12]	Злиття мультимодальних ознак, включаючи текст, зображення та відео
Li et al. [13]	Генеративно-змагальні мережі (GAN)
Sahoo et al. [14]	Використання кількох ознак, пов'язаних з акаунтами Facebook та новинним контентом
Zhang et al. [15]	Нейронна обчислювальна структура на основі згорток
Wang et al. [17]	Нова детальна мультимодальна мережа злиття ознак (FMFN)
Li et al. [18]	Семантично-посилена мультимодальна мережа злиття
Guo et al. [19]	Нейронна мережа з взаємною увагою (MANN)
Наша DESF	Мультидоменна ознака, семантична ознака, емоційна ознака, злиття ознак, увага

Таким чином, аналіз існуючих рішень у сфері виявлення фейкових новин та злиття ознак демонструє, що поєднання різних типів ознак є важливим фактором для покращення точності та надійності моделей. Використання графів знань, мультимодальних ознак, а також генеративно-змагальних мереж дозволяє виявляти дезінформацію на основі більш повної та глибокої інформації. Емоційні та семантичні ознаки відіграють ключову роль у розумінні контексту та інтенсивності настрою, що допомагає відрізнити правдиві новини від фейкових. Злиття цих ознак забезпечує більш точне подання даних та дозволяє моделям краще розуміти та класифікувати новинний контент. Запропонована в цій роботі модель DESF, яка поєднує мультидоменні, семантичні та емоційні ознаки із використанням механізму уваги, показує високий потенціал для подальшого вдосконалення технологій виявлення фейкових новин.

1.3 Вибір перспективного шляху і постановка задачі дослідження

Актуальність дослідження виявлення фейкових новин у сучасних умовах глобалізації інформаційних потоків стає дедалі важливішою, оскільки соціальні медіа відіграють значну роль у формуванні громадської думки та впливають на поведінку людей. Згідно зі статистикою, станом на липень 2024 року кількість користувачів інтернету досягла 5.45 мільярда осіб, а кількість користувачів соціальних мереж перевищила 5.17 мільярда. У такому середовищі поширення фейкових новин стає глобальною проблемою, що призводить до дезінформації, маніпуляцій та підриву довіри до інформаційних джерел.

Однією з ключових проблем є швидкість поширення фейкових новин у соціальних медіа. Завдяки алгоритмам платформ, які підсилюють "ефект ехо-камери," користувачі отримують обмежений доступ до альтернативних поглядів, що сприяє поширенню дезінформації. Це явище особливо гостро відчувається під час кризових подій, коли недостовірна інформація поширюється швидко і може викликати паніку, економічні втрати та політичну нестабільність. Фейкові новини часто експлуатують емоції людей, що робить їх важким для виявлення за допомогою традиційних методів обробки інформації.

Зі зростанням обсягів дезінформації виникає потреба у вдосконаленні методів виявлення фейкових новин, які поєднують аналіз різних типів ознак — текстових, семантичних, емоційних та доменних. Сучасні дослідження вказують на те, що методи, які використовують лише одну ознаку (наприклад, текстові або візуальні дані), не забезпечують достатньої точності. Тому злиття різних типів ознак стає необхідним кроком для покращення якості класифікації новинного контенту.

Сучасні підходи, такі як використання генеративно-змагальних мереж (GAN) та моделей глибокого навчання, вже показали свій потенціал у виявленні фейкових новин. Проте ці моделі мають свої обмеження, зокрема у здатності

адаптуватися до різних доменів новинного контенту. У цьому контексті особливе значення має розробка нових моделей, які можуть одночасно аналізувати емоційні, семантичні та доменні ознаки, забезпечуючи тим самим точніше виявлення фейкових новин.

Актуальність проблеми також пов'язана з викликами, які виникають у контексті мовної та культурної специфіки новинного контенту. Наприклад, китайська та англійська мови мають значні відмінності в структурі речень, граматиці та семантиці, що створює додаткові труднощі для автоматичних систем виявлення фейкових новин. Врахування цих аспектів є важливим для створення універсальних моделей, які можуть працювати з новинним контентом різних мов.

Іншим аспектом актуальності є те, що сучасні моделі для виявлення фейкових новин повинні бути стійкими до змін контенту та адаптуватися до нових форм дезінформації. Фейкові новини постійно еволюціонують, і автори дезінформації використовують нові техніки для обходу систем виявлення. Тому необхідні моделі, здатні до самонавчання та вдосконалення на основі нових даних.

Економічні наслідки поширення фейкових новин також підкреслюють актуальність дослідження. Компанії, організації та уряди вкладають значні ресурси в боротьбу з дезінформацією, але без ефективних інструментів для автоматизованого виявлення фейкових новин ці зусилля залишаються неефективними. Тому необхідні інноваційні рішення, які дозволять не лише виявляти фейкові новини, а й прогнозувати їх поширення.

Також важливо зазначити соціальний аспект проблеми. Фейкові новини можуть впливати на здоров'я населення, сприяти поширенню шкідливих медичних порад, особливо під час пандемій та інших криз. Таким чином, актуальність розробки ефективних моделей для виявлення фейкових новин у медичній та соціальній сферах є надзвичайно високою.

У результаті проведеного дослідження актуальність проблеми виявлення фейкових новин, з урахуванням мультидоменних, емоційних та семантичних аспектів, є обґрунтованою і спрямованою на вирішення однієї з найважливіших інформаційних загроз сучасного світу.

Мета дослідження полягає в розробці моделі інтеграції емоційних, семантичних та доменних ознак для ефективного виявлення фейкових новин у багатодоменних та багатомовних інформаційних потоках.

Досягнення цієї мети зумовило потребу теоретичних розробок, визначення та послідовного вирішення таких завдань:

1. Проаналізувати поширення фейкових новин у соціальних медіа, враховуючи сучасні тенденції та виклики.
2. Оглянути існуючі рішення щодо виявлення фейкових новин, зокрема методи машинного навчання та глибокого навчання.
3. Розробити архітектуру моделі для інтеграції емоційних, семантичних та доменних ознак з метою виявлення фейкових новин.
4. Моделювати контекстуальну релевантність у новинних текстах для покращення точності класифікації.
5. Дослідити вплив емоційних ознак на поширення фейкових новин та їх інтеграцію в модель.
6. Розробити методи видобування доменних ознак для покращення міждоменної релевантності в новинних текстах.
7. Провести експерименти з оцінюванням точності, повноти та F1-міри на основі різних наборів даних, включаючи абляційний аналіз.

Актуальність дослідження виявлення фейкових новин зумовлена глобалізацією інформаційних потоків та значною роллю соціальних медіа у формуванні громадської думки. Враховуючи масштаби поширення дезінформації, виклики мовної та культурної специфіки, швидкість адаптації фейкових новин до нових умов, а також їхні економічні та соціальні наслідки, проблема є однією з ключових у сучасному інформаційному середовищі. У

цьому контексті визначено мету дослідження — розробка моделі інтеграції емоційних, семантичних і доменних ознак для виявлення фейкових новин у багатодоменних і багатомовних потоках. Для досягнення цієї мети поставлено низку завдань, які охоплюють аналіз сучасних тенденцій, огляд існуючих рішень, створення нових методів та експериментальну оцінку їхньої ефективності, що забезпечує комплексний підхід до вирішення проблеми.

Висновки до розділу 1

1. У результаті аналізу поширення фейкових новин в соціальних медіа виявлено, що зростання кількості користувачів Інтернету і соціальних мереж безпосередньо впливає на швидкість поширення дезінформації. Соціальні мережі, завдяки своїм алгоритмам, сприяють ефекту "ехо-камери", який обмежує доступ до альтернативних точок зору і посилює вплив фейкових новин. Фейкові новини використовують емоційні тригери та провокаційну лексику, що підсилює їхній вплив на громадськість і збільшує негативні наслідки для суспільства.

2. Огляд існуючих рішень у сфері виявлення фейкових новин показав, що традиційні методи машинного навчання, такі як Наївний Байєс та Метод опорних векторів, є ефективними, але з розвитком технологій глибокого навчання, таких як CNN та RNN, їхні можливості значно розширюються. Використання графів знань, мультимодальних ознак і генеративно-змагальних мереж підвищує здатність моделей до точного виявлення дезінформації, оскільки дозволяє моделювати складні зв'язки між текстовими, візуальними та контекстуальними ознаками.

3. Запропонована в цьому дослідженні модель DESF, яка об'єднує мультидоменні, семантичні та емоційні ознаки з використанням механізму уваги, показала перспективні результати у виявленні фейкових новин. Злиття ознак дозволяє поліпшити точність класифікації, а також краще розуміти

взаємозв'язки між різними типами даних, що робить цю модель більш адаптивною і стійкою до різних видів дезінформації.

2 МОДЕЛЬ ІНТЕГРАЦІЯ ЕМОЦІЙНИХ, СЕМАНТИЧНИХ ТА ДОМЕННИХ ОЗНАК ДЛЯ ВИЯВЛЕННЯ ФЕЙКОВИХ НОВИН

2.1 Архітектура моделі інтеграція емоційних, семантичних та доменних ознак для виявлення фейкових новин

Виявлення фейкових новин часто моделюється як задача бінарної класифікації, де на вході моделі є текст новин, а на виході — мітка, яка вказує, чи є новина правдивою, чи фейковою. У цій роботі пропонується модель DESF, яка ефективно поєднує емоційні, доменні та текстові ознаки. Модель спрямована на визначення емоційної полярності важливих слів у різних доменах, захоплення специфічної для домену інформації в тексті та злиття емоційних ознак доменних слів із семантичними ознаками для отримання більш детальної текстової інформації.

Модель DESF складається з п'яти основних частин:

1. Модуль видобування доменних ознак: Виділяє ознаки різних доменів через кодування дев'яти доменів, таких як політика, військова справа, культура, суспільство тощо.

2. Модуль видобування емоційних ознак: Аналізує емоційну тенденцію тексту за допомогою моделі Senta-BiLSTM та використовує механізм уваги для злиття доменних і емоційних ознак.

3. Модуль видобування семантичних ознак: Використовує Bi-LSTM для вилучення семантичних ознак із тексту.

4. Модуль злиття ознак: Повністю інтегрує семантичні ознаки з ознаками доменної емоційної релевантності для захоплення глобальних залежностей тексту.

5. Шар прогнозування: Класифікує злиті ознаки за допомогою MLP і функції Softmax, а також обчислює втрати.

Загальна архітектура моделі, представлена на рисунку 2.1, включає: вхідний шар, шар представлення ознак, шар уваги, шар злиття ознак та вихідний шар.

Такий підхід дозволяє ефективно комбінувати кілька різних типів ознак, що суттєво підвищує точність моделі у виявленні фейкових новин.

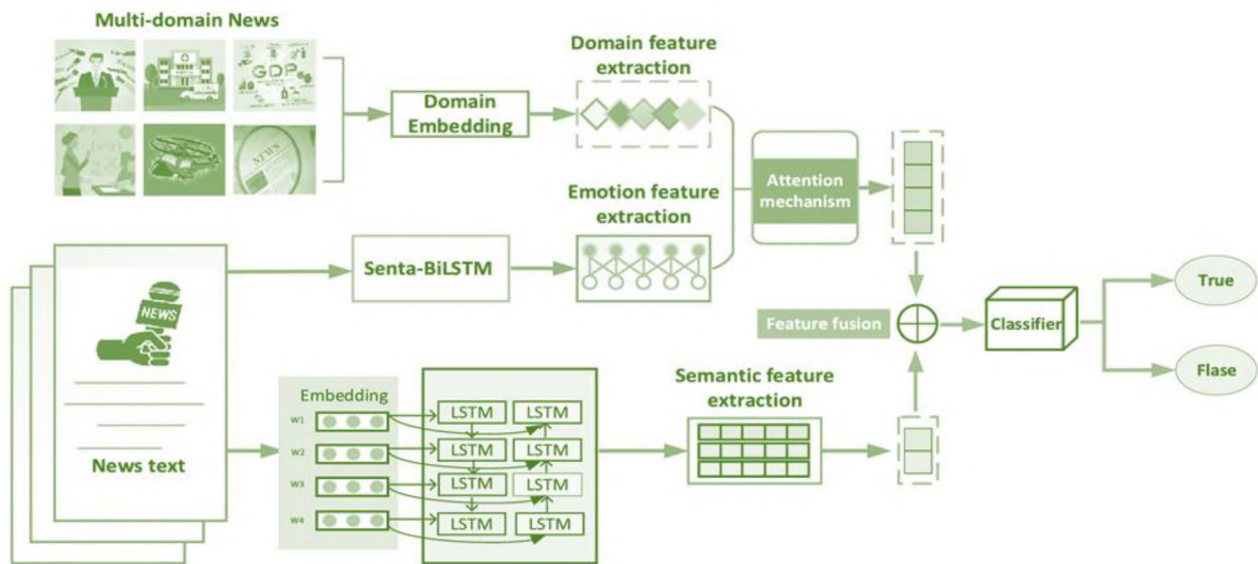


Рисунок 2.1 - Загальна архітектурна схема моделі DESF

2.2 Моделювання контекстуальної релевантності в новинних текстах

Припустимо, що N представляє текст новин, що складається з T слів, де w означає слово, тоді текст новин можна виразити як $N = \{w_1, w_2, w_3, \dots, w_T\}$. Перед введенням даних відбувається очищення тексту, яке включає видалення дубльованих даних, стоп-слів та інших непотрібних елементів, щоб покращити якість даних і усунути зайвий шум.

Для тексту новин часто важливо звертати увагу на контекстуальну релевантність, особливо коли мова йде про тексти з різних доменів. Поєднання контексту може допомогти краще визначити семантику. Наприклад, у сфері електроніки, якщо коментується продуктивність батареї, прикметник "швидка", якщо стоїть перед словом "зарядка", вказує на те, що батарея має хороші

показники і швидко заряджається. Але якщо він стоїть перед "витратою енергії", це свідчить про низьку продуктивність із коротким часом автономної роботи.

У традиційних моделях рекурентних нейронних мереж (RNN) і LSTM інформація може поширюватися тільки вперед, через що стан у момент часу t залежить лише від інформації, отриманої до цього моменту. Щоб забезпечити наявність контекстуальної інформації на кожному етапі, у цій роботі використовується поєднання двонапрявленої рекурентної нейронної мережі (BiRNN) та одиниць LSTM, що відоме як Bi-LSTM.

Bi-LSTM може одночасно обробляти дані з двох напрямків, використовуючи дві незалежні приховані шари, і поєднувати вихідні дані як вхідні для вихідного шару [21]. Це дозволяє зберігати зв'язок між словами оригінального тексту та повністю враховувати контекст, що ефективно знижує семантичну неоднозначність.

Принцип роботи моделі LSTM формулюється таким чином:

$$\begin{aligned}
 i_t &= \text{sigmoid}(W_i \times [h_{t-1}, x_t] + b_i) \\
 f_t &= \text{sigmoid}(W_f \times [h_{t-1}, x_t] + b_f) \\
 o_t &= \text{sigmoid}(W_o \times [h_{t-1}, x_t] + b_o) \\
 c_t &= f_t \odot c_{t-1} + i_t \odot \tanh(W_c \times [h_{t-1}, x_t] + b_c) \\
 h_t &= o_t \odot \tanh(c_t) \quad ,
 \end{aligned} \tag{2.1}$$

де x_t - вхідний вектор у момент часу t ;

i_t - вхідні двері;

f_t - забуваючі двері;

o_t - вихідні двері;

W_i - матриця ваг для вхідних дверей;

W_f - матриця ваг для забуваючих дверей;

W_o - матриця ваг для вихідних дверей;

b_i – вектор зміщення для вхідних дверей;
 b_f – вектор зміщення для забуваючих дверей;
 b_o – вектор зміщення для вихідних дверей;
 c_t – блок пам'яті поточного моменту;
 $t - 1$ – момент після поточного моменту t ;
 W_c – матриця ваг для поточної інформації;
 b_c – вектор зміщення для поточної інформації;
 $\sigma(\cdot)$ та $\tanh(\cdot)$ – активаційні функції;
 h_t — це вектор виходу.

Отже, вихід можна виразити наступним чином:

$$\begin{aligned}
 \vec{h}_t &= LSTM\left(\vec{h}_{t-1}, x_t\right) \\
 \overleftarrow{h}_t &= LSTM(\overleftarrow{h}_{t+1}, x_t) \\
 y_t &= \begin{bmatrix} \vec{h}_t & \overleftarrow{h}_t \end{bmatrix}
 \end{aligned} \tag{2.2}$$

де $\vec{h}, \overleftarrow{h}$ – вектори ознак, отримані прямою та зворотною моделями LSTM;
 y_t – вектор семантичних ознак, отриманих моделлю Bi-LSTM у поточний момент часу.

2.3 Вплив емоційних ознак на поширення фейкових новин

У реальних сценаріях видавці новин часто публікують твіти з сильними емоціями, щоб привернути трафік, викликати резонанс у аудиторії та сприяти широкому поширенню. Фейкові новини часто містять такі характеристики, як

"перебільшені техніки" і "провокаційна лексика", що впливають на суб'єктивні судження читачів. Тому емоційні ознаки відіграють важливу роль.

Після вибору набору даних відповідного розміру та різноманітних категорій, у цій роботі використовується модель Senta-BiLSTM для аналізу емоційної тенденції тексту [22]. Завдяки захопленню далеких залежностей у тексті та емоційної інформації в контексті, модель досягає точного визначення полярності настрою тексту. У конкретних експериментах модель спочатку налаштовується шляхом коригування параметрів для адаптації до завдання та набору даних. Далі модель використовується для аналізу новинних текстів, передбачаючи ймовірність позитивного і негативного настрою тексту. У результаті отримується подання емоційної полярності тексту новин. У цій роботі N використовується для позначення новинного тексту, а y_e використовується для позначення вектора емоційних ознак, тоді:

$$y_e = Senta_BiLSTM(N), \quad (2.3)$$

2.4 Видобування доменних ознак для покращення міждоменної релевантності в новинних текстах

Проблема зміни домену та міждоменної релевантності у різних новинних доменах створює значні виклики для прогнозувальних можливостей моделі. Мета видобування доменних ознак полягає в ідентифікації джерел, які представляють ключову інформацію даних, роблячи модель дискримінативною, стабільною та інтерпретованою.

Після видобування семантичних та емоційних ознак тексту в цій роботі використовується кодування домену як додаткову ознаку для того, щоб зробити процес видобування ознак більш деталізованим. Під час експериментів

використовуються вручну класифіковані доменні мітки, і для кожного домену отримується закодоване векторне подання домену.

$$D = \{D_1, \dots, D_n\} \quad , \quad (2.4)$$

де D – вектор домену.

Те саме емоційне слово може мати різну емоційну тенденцію в різних доменах текстів. У цій роботі використовується механізм уваги для захоплення взаємозв'язку "домен-емоція". Призначаючи різні ваги емоційним словам у різних доменах, модель може точніше розуміти емоційну інформацію в тексті.

В експерименті застосовується адитивний механізм уваги для динамічного налаштування зв'язку між доменними ознаками та емоційними ознаками шляхом навчання вектора ваги уваги. Спочатку розраховується оцінка уваги, і її формула виглядає так:

$$\text{text score}(D, y_e) = \text{text tanh}(W_D D + W_{y_e} y_e) \quad , \quad (2.5)$$

де W_D та $W_{\{y_e\}}$ є матрицями ваг для навчання.

На основі оцінки уваги обчислюється вага уваги α за наступною формулою:

$$\alpha = \text{text softmax}(\text{text score}(D, y_e)) \quad , \quad (2.6)$$

Зважене підсумовування доменних ознак за допомогою ваг уваги дає кінцеве подання:

$$textfinalrepresentation = f_r = sum(\alpha_n D_n), \quad (2.7)$$

де n — кількість доменів;

f_r — кінцеве подання.

Посилені доменні ознаки D , емоційні ознаки y_e і семантичні ознаки y_t зливаються через адаптивний гейт, щоб утворити кінцеву зливу ознаку C . Цей процес захоплює глобальні залежності тексту і адаптивно зосереджується на ключових ознаках, знижуючи кількість зайвої інформації, введеної під час конкатенації моделі, і покращуючи здатність до розуміння текстів новин. Це можна виразити за допомогою такої формули:

$$C = \alpha D + \beta y_e + y_t, \quad (2.8)$$

У формулі додаються параметри, що навчаються, α та β , причому параметр перед семантичними ознаками y_t позначається як 1.

Злиті ознаки подаються на вхід багат шарової перцептронної мережі (MLP), після чого з'єднуються з шаром Softmax для передбачення фейкових новин, як це показано у формулі:

$$y = \text{softmax}(W * MLP(C) + b), \quad (2.9)$$

де W і b — параметри шару з'єднання і терми зсуву відповідно.

Для обчислення втрат у цій роботі застосовується функція втрат на основі перехресної ентропії для тренування, як це показано у формулі:

$$L = -[y \log \hat{y} + (1 - y) \log(1 - \hat{y})], \quad (2.10)$$

де y — справжнє значення мітки;

$\hat{y}$ — передбачуване значення ймовірності.

2.5 Метрики оцінювання моделей класифікації та їх застосування у виявленні фейкових новин

Метрики оцінювання включають точність (accuracy), точність класифікації (precision), повноту (recall) та F1-міру. Їх значення наступні:

Точність (Accuracy) представляє частку правильно класифікованих зразків від загальної кількості зразків.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN}, \quad (2.11)$$

де TP — кількість справжніх позитивних зразків;

TN — кількість справжніх негативних зразків;

FP — кількість хибно позитивних зразків;

FN — кількість хибно негативних зразків.

Точність класифікації (Precision) представляє частку правильно класифікованих позитивних зразків від усіх позитивних зразків, класифікованих моделлю.

$$Precision = \frac{TP}{TP+FP}, \quad (2.12)$$

Повнота (Recall) представляє частку правильно класифікованих позитивних зразків від загальної кількості позитивних зразків.

$$Recall = \frac{TP}{TP+FN} , \quad (2.13)$$

F1-міра представляє зважене гармонійне середнє між точністю та повнотою, де P означає точність, а R — повноту.

$$F_1 = \frac{2*P*R}{P+R} , \quad (2.14)$$

Ці метрики використовуються для оцінки якості моделей класифікації, особливо у задачах з дисбалансом класів.

Щоб повністю перевірити ефективність DESF, у цій роботі використовуються два типи базових моделей для порівняльної перевірки та аналізу експериментальних результатів.

Перший тип: Вибираємо наступні базові моделі з точки зору врахування доменних ознак:

— FuDFEND [25]: Підхід до виявлення фейкових новин на основі нечітких доменів, який використовує нейронні мережі для моделювання процесу нечіткого висновку та створює нечітку доменну мітку для кожної новини.

— MDFEND [24]: Метод багатодоменового виявлення на основі багатопрофільної структури, який агрегує подання кількох "експертів" через "домени гейти".

— SLFEND [26]: Ця модель використовує м'які мітки для видобування багатодомених ознак новин і отримання остаточного загального подання ознак.

— M3FEND [27]: Модель використовує підхід багатовидового моделювання для одночасного моделювання семантичної, стилістичної та емоційної інформації новин, а також використовує метод взаємодії між видами для більш повного моделювання інформаційних взаємодій.

— CFPFND [28]: Метод на основі злиття ознак між різними характеристиками, який вловлює відмінності ознак між новинами з різних доменів, досліджує взаємозв'язки між новинами та контролює стратегію злиття ознак моделі в різних доменах з різних вимірів.

Другий тип: Наступні базові моделі вибрані з точки зору злиття ознак:

— BERT [29]: Використовує попередньо натреновану модель BERT для вилучення текстових ознак новин і їх класифікації.

— BiGRU [30]: Багатошарова структура на основі GRU, яка додає другий шар GRU для захоплення взаємодії високорівневих ознак між різними часовими кроками.

— SpotFake [31]: Мультимодальна структура, яка використовує як текстові, так і візуальні ознаки статті для виявлення фейкових новин.

— EANN [32]: Ця модель використовує дискримінатор подій для вимірювання відмінностей між різними подіями та навчання інваріантних ознак подій.

Висновки до розділу 2

1. Запропонована модель DESF демонструє високу ефективність у задачах виявлення фейкових новин завдяки поєднанню емоційних, семантичних та доменних ознак. Поєднання цих ознак дозволяє не лише враховувати емоційний контекст новинного контенту, але й розуміти специфіку домену, що допомагає досягти більш точних результатів класифікації. Використання Bi-LSTM для вилучення семантичних ознак з тексту дозволяє моделі захоплювати довгострокові залежності у текстах, що є критичним для аналізу складних новинних текстів. Механізм уваги, що використовує емоційно-доменні зв'язки, допомагає зосередитися на найважливіших аспектах тексту, що підвищує точність виявлення фейкових новин.

2. Аналізуючи вплив емоційних ознак, можна дійти висновку, що вони відіграють важливу роль у виявленні фейкових новин. Емоційні тригери, які використовуються у фейкових новинах, мають на меті маніпулювати емоціями читачів, сприяючи їх поширенню. Використання моделі Senta-BiLSTM дозволяє точно визначати полярність емоцій у тексті, що допомагає відрізнити фейковий контент від реального. Врахування емоційного компонента особливо важливе для аналізу новин із різних доменів, таких як політика, культура та суспільство, де емоційний вплив має значний вплив на поширення контенту.

3. Модель DESF демонструє високу здатність до адаптивної класифікації фейкових новин у різних доменах завдяки використанню доменних ознак. Виділення доменних ознак дозволяє моделі краще розуміти специфіку контенту з різних тематичних областей, що робить її стійкою до міждоменної релевантності та зміни контексту новин. Важливо зазначити, що застосування адаптивного механізму уваги у поєднанні з мультидоменними ознаками дозволяє моделі точніше класифікувати новини в умовах змін домену, що робить її універсальною для використання в різних тематичних контекстах.

3 ПРОГРАМНО-ТЕХНОЛОГІЧНЕ ЗАБЕЗПЕЧЕННЯ

3.1 Дослідження набору даних

Щоб запобігти перенавчанню моделі, у цій роботі введено Dropout для забезпечення випадковості під час навчання, встановивши параметр Dropout на рівні 0.5. Швидкість навчання встановлена на рівні 0.001 для тренування моделі, і проводяться 200 ітерацій навчання. L2-регуляризація застосовується до повністю зв'язаного шару з параметром 0.01, а розмір встановлюється на 50. Для забезпечення ефективності було обрано оптимізатор Adam, який автоматично регулює швидкість навчання та зменшення ваг. Для зменшення нестабільності під час навчання моделі використовуються різні випадкові числа для ініціалізації моделі, а середнє значення результатів обчислюється на основі трьох запусків.

У цій роботі для перевірки ефективності моделі DESF було обрано набори даних FakeNewsNet [23] та Weibo21 [24]. FakeNewsNet є сховищем даних, яке містить контент новин, соціальний контекст та просторово-часову інформацію, охоплюючи переважно два домени: політику та розваги. Набір даних Weibo21 є реорганізацією загальнодоступного набору даних Weibo, здійсненою лабораторією Інституту обчислювальної техніки Китайської академії наук. Він включає інформацію про фейкові новини, офіційно розкриті платформою Weibo з грудня 2014 по березень 2021 року. Кожен елемент даних містить багатовимірну інформацію, таку як текстовий контент, коментарі та часові мітки, охоплюючи загалом 9 тематичних доменів.

У роботі проводиться збір і організація наборів даних, і остаточна статистична інформація для англomовного набору даних представлена в таблиці 3.1, а для китайського набору даних — у таблиці 3.2.

Таблиця 3.1 - Статистика наборів даних FakeNewsNet

Домен	PolitiFact	GossipCop	Всього
Реальні	447	16,804	17,251
Фейкові	379	5067	5446

Таблиця 3.2 - Статистика наборів даних Weibo21

Домен	Наука	Військова справа	Освіта	Стихійні лиха	Політика
Реальні	143	121	243	185	306
Фейкові	93	222	248	591	546
Загалом	236	343	491	776	852
Домен	Охорона здоров'я	Фінанси	Розваги	Суспільство	Всього
Реальні	485	959	1000	1198	4460
Фейкові	515	362	440	1471	4488
Загалом	1000	1321	1440	2669	9128

У цій роботі набір даних ділиться на навчальний, валідаційний та тестовий набори в співвідношенні 6:2:2. Використовуються 9 доменів із китайського набору даних і 2 домени з англійського набору даних для максимального використання ефекту посилення доменних ознак. Кожному новинному елементу присвоюється мітка з конкретним доменом.

3.2 Аналіз результатів експериментів

У цьому експерименті було використано дев'ять базових моделей та модель DESF для навчання на китайських і англійських наборах даних. Після навчання моделі були протестовані на тестовому наборі, а метрикою оцінювання обрана F1-міра. Результати, отримані на основі різних базових моделей, представлені далі.

З точки зору того, як моделі поєднують доменні ознаки для виявлення, результати показані в таблицях 3.3 і 3.4.

Щоб полегшити більш наочне порівняння результатів моделі DESF з іншими базовими моделями на Weibo, у цій роботі використовується комбінована діаграма для візуалізації змін і відмінностей у кривих даних, як показано на рисунку 3.1

Таблиця 3.3 - Порівняння продуктивності DESF та інших базових методів на Weibo21

Метод	Освіта	Стихійні лиха	Охорона здоров'я	Фінанси	Політика	Військова справа	Наука	Суспільство	Розваги
MDFEND	0.883	0.878	0.943	0.875	0.883	0.936	0.833	0.894	0.909
M3FEND	0.884	0.882	0.952	0.899	0.888	0.928	0.838	0.904	0.929
FuDFEND	0.892	0.906	0.942	0.890	0.901	0.947	0.813	0.917	0.916
SLFEND	0.928	0.891	0.924	0.923	0.937	0.923	0.883	0.927	0.925
CFPFEND	0.918	0.896	0.957	0.905	0.897	0.971	0.862	0.915	0.928
DESF	0.923	0.913	0.959	0.927	0.920	0.958	0.887	0.929	0.931

Таблиця 3.4 - Загальний огляд продуктивності різних моделей

Метод	Точність	Повнота	F1-міра
MDFEND	0.906	0.880	0.893
M3FEND	0.898	0.902	0.900
FuDFEND	0.901	0.905	0.903
SLFEND	0.910	0.916	0.918
CFPFEND	0.904	0.930	0.917
DESF	0.929	0.925	0.927

Рисунок 3.1 показує, що запропонована в цій роботі модель DESF перевершує інші моделі в багатьох доменах, особливо в таких сферах, як охорона здоров'я, економіка, наука, суспільство, стихійні лиха та розваги, демонструючи високу здатність до узагальнення і загальну стійкість. Аналізуючи дані по окремих доменах, модель DESF досягла F1-міри 0.959 у домені охорони здоров'я, що на 0.002 вище, ніж друга за продуктивністю модель CFPEND у цьому домені. У домені стихійних лих модель DESF досягла F1-міри 0.913, що на 0.007 вище, ніж F1-міра другої найкращої моделі FuDFEND, у той час як F1-міри інших чотирьох моделей у цьому домені коливаються близько 0.89. В економічному домені модель DESF показала хороший результат із F1-мірою 0.927, що на 5.9% вище за MDFEND та на 0.4% вище за модель SLFEND. У доменах науки,

суспільства та розваг F1-міри моделі DESF становлять відповідно 0.887, 0.929 і 0.931, що на 0.5%, 0.2% і 0.2% вище відповідних других найкращих моделей. У військовому домені DESF досягла F1-міри 0.958, що на 0.013 нижче за CFPFND, тоді як FuDFEND та MDFEND також досягли хороших результатів із F1-мірами 0.947 та 0.936.

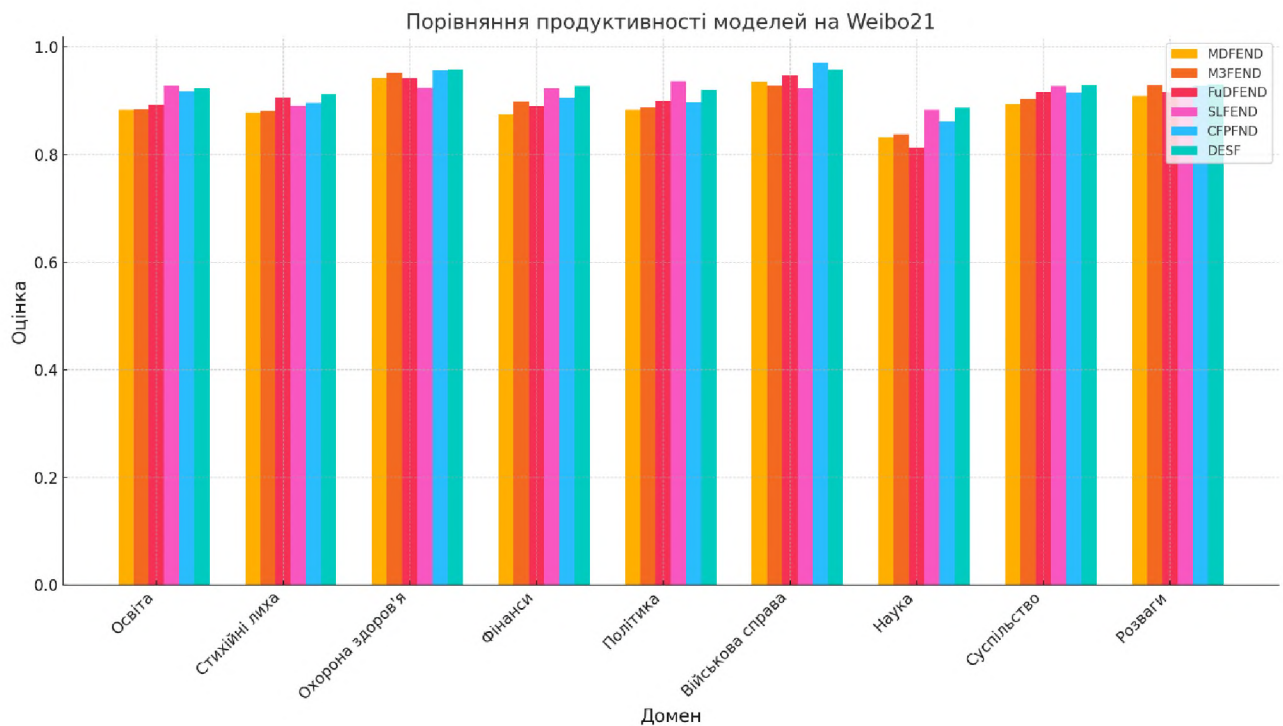


Рисунок 3.1 - Порівняння результатів різних моделей на наборах даних Weibo21 (оцінка F1)

У політичному домені найкращий результат показала модель SLFEND, тоді як DESF, поступившись лише SLFEND, досягла F1-міри 0.920. У домені освіти загалом було 491 новинний елемент, з яких 243 — реальні новини, а 248 — фейкові. Хоча розмір вибірки не є великим, модель DESF добре виявляє фейкові новини, досягнувши F1-міри 0.923 у цьому домені. З точки зору дев'яти доменів, включених у цей набір даних, зелена лінія, що представляє модель DESF, загалом рівномірно розподілена, що свідчить про відносну стабільність цієї моделі та її здатність підтримувати відносно точну швидкість виявлення.

Набір даних FakeNewsNet в основному складається з двох доменів: PolitiFact та GossipCop. Як показано в Таблиці 3.5, запропонована модель DESF перевершує інші п'ять моделей. У наборі даних PolitiFact DESF досягла F1-міри 0.862, що на 0.001 більше в порівнянні з другою за результатами моделлю MDFEND, яка мала F1-міру 0.861.

Таблиця 3.5 - Порівняння продуктивності різних методів на FakeNewsNet

Метод	PolitiFact	GossipCop	Всього
MDFEND	0.861	0.806	0.822
M3FEND	0.856	0.822	0.820
FuDFEND	0.859	0.822	0.819
SLFEND	0.857	0.823	0.822
CFPFND	0.860	0.826	0.825
DESF	0.862	0.828	0.825

У наборі даних GossipCop DESF досягла F1-міри 0.828, що на 0.2% вище за другу за результатами модель CFPFND, що свідчить про хорошу здатність до виявлення фейкових новин. Таким чином, у цілому запропонована в цій роботі модель DESF демонструє певний прогрес.

Порівняно з результатами на китайському наборі даних, модель DESF показує кращу продуктивність у виявленні фейкових новин на китайському наборі даних, ніж на англійському. Це може бути зумовлено значними відмінностями в граматиці, словниковому запасі та структурах речень між китайською та англійською мовами. Ці відмінності можуть призводити до варіацій у продуктивності моделі під час обробки двох мов.

Наприклад, китайські слова часто мають багатше значення та контекстуальні залежності, тоді як в англійській мові більша увага приділяється порядку слів і граматичним структурам. Результати для PolitiFact показані на рисунку 3.2, а результати для GossipCop — на рисунку 3.3.

З точки зору того, чи модель поєднує злиття ознак для виявлення фейкових новин, експериментальні результати представлені далі.

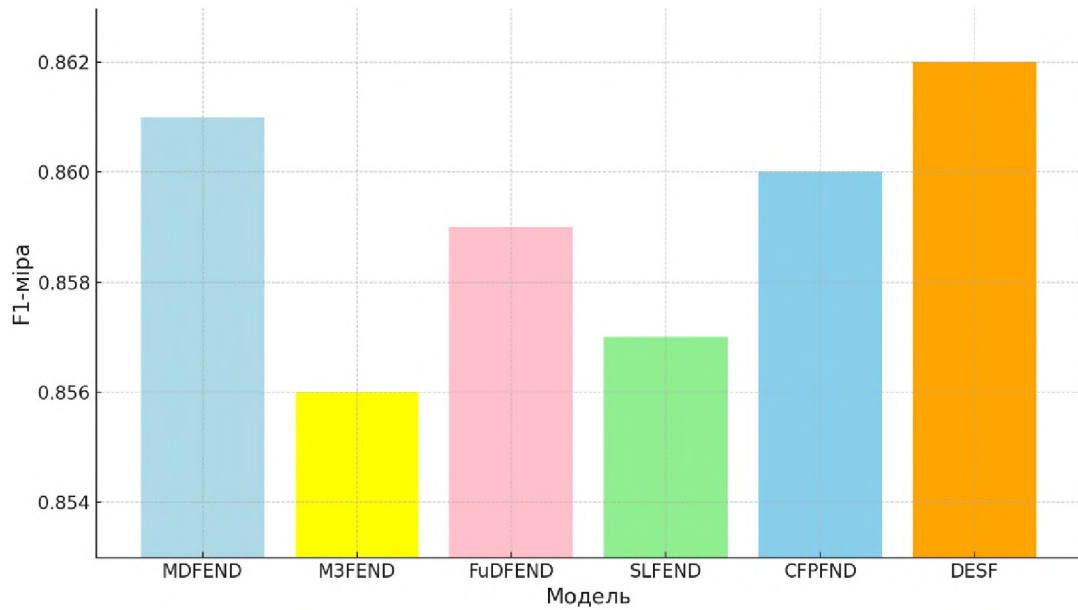


Рисунок 3.2 - Результати на PolitiFact

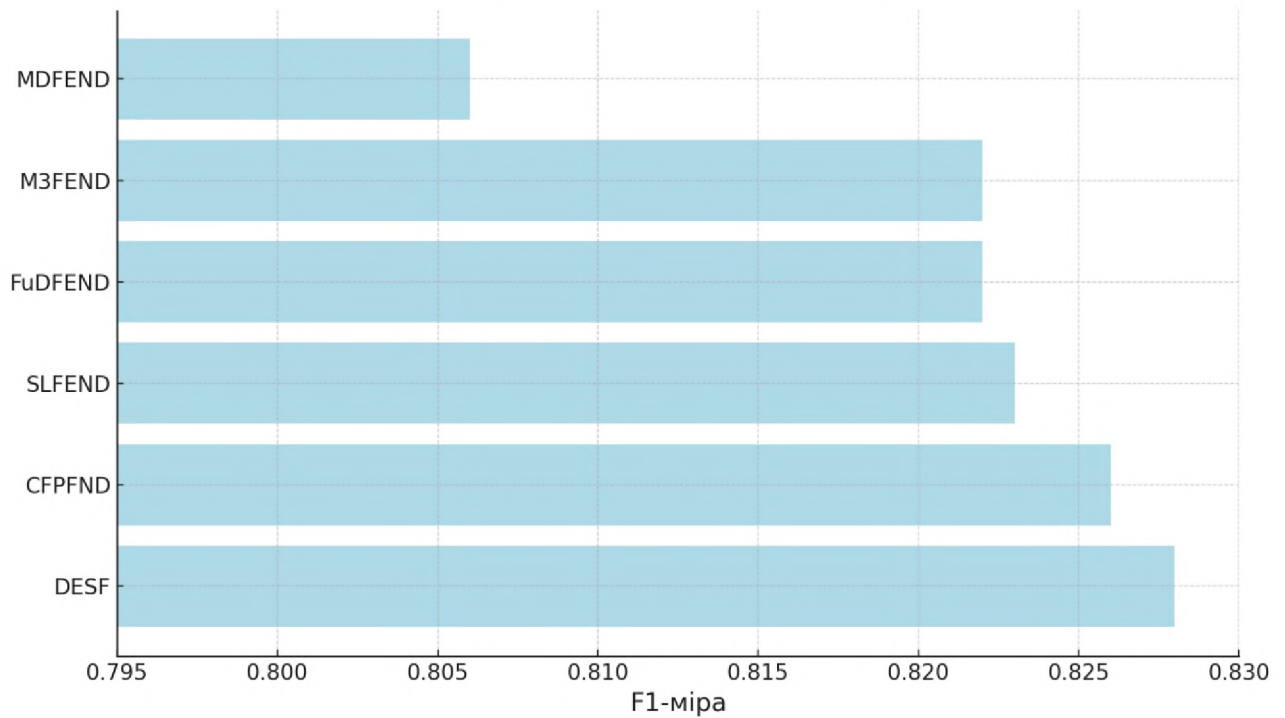


Рисунок 3.3 - Результати на наборі даних GossipCop

Таблиця 3.6 показує, що DESF демонструє чудову продуктивність у порівнянні з іншими базовими моделями. Серед дев'яти доменних наборів даних DESF досягла F1-міри 0.958, що на 4.9% вище за другу за результатами модель EANN та на 6.3% вище за SpotFake. У домені освіти DESF досягла F1-міри 0.923, що на 6.3% вище за модель EANN та на 13.4% вище за BIGRU. Більше того, у

сферах стихійних лих, економіки, охорони здоров'я, розваг, політики та суспільства DESF досягла F1-міри, що перевищує 0.9, демонструючи хорошу продуктивність. Це свідчить про те, що методи злиття ознак відіграють важливу роль у процесі виявлення фейкових новин, оскільки поєднання різних типів ознак за допомогою різних моделей і алгоритмів може давати різні результати. DESF поєднує взаємозв'язок між доменними емоційними ознаками та семантичними ознаками, досягаючи хороших результатів у виявленні на обох наборах даних.

Таблиця 3.6 - Порівняння продуктивності різних методів на Weibo21

Метод	Освіта	Стихійні лиха	Охорона здоров'я	Фінанси	Розваги	Політика	Військова справа	Наука	Суспільство
BiGRU	0.814	0.794	0.887	0.829	0.863	0.836	0.872	0.727	0.849
EANN	0.868	0.894	0.935	0.879	0.919	0.869	0.913	0.826	0.904
BERT	0.833	0.851	0.909	0.875	0.877	0.921	0.907	0.777	0.858
SpotFake	0.857	0.893	0.964	0.847	0.903	0.819	0.901	0.932	0.898
DESF	0.923	0.913	0.959	0.927	0.931	0.920	0.958	0.887	0.929

Таблиця 3.7 показує продуктивність п'яти моделей: BiGRU, EANN, BERT, SpotFake та DESF на FakeNewsNet. Порівняння показало, що DESF досягла найкращої продуктивності в домені PolitiFact, отримавши F1-міру 0.862, що на 3.6% вище, ніж у другої за результатами моделі BiGRU, і на 5.8% вище, ніж у моделі SpotFake. У домені GossipCop DESF досягла другої за результатами продуктивності. Загалом, модель DESF із F1-мірою 0.845 на 1.8% перевищила другу за результатами модель та на 4.8% перевищила EANN, демонструючи високу конкурентоспроможність та здатність до виявлення. Для наочного відображення відмінностей у продуктивності моделей у цій роботі використовуються Рисунки 3.4 та 3.5 для показу змін кривих.

Рисунки 3.4 та 3.5 показують порівняння результатів різних моделей. Чітко видно, що запропонована в цій роботі модель DESF має очевидні переваги на наборі даних FakeNewsNet і краще виконує задачу виявлення фейкових новин.

Таблиця 3.7 - Порівняння продуктивності різних методів на FakeNewsNet

Метод	PolitiFact	GossipCop	Всього
BiGRU	0.832	0.819	0.826
EANN	0.826	0.786	0.806
BERT	0.827	0.833	0.830
SpotFake	0.815	0.820	0.818
DESF	0.862	0.828	0.845

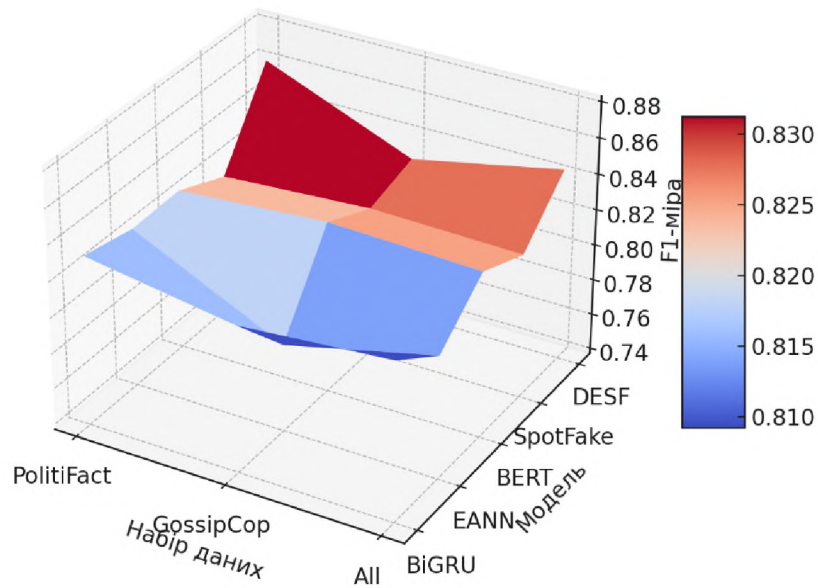


Рисунок 3.4 - Продуктивність між різними моделями

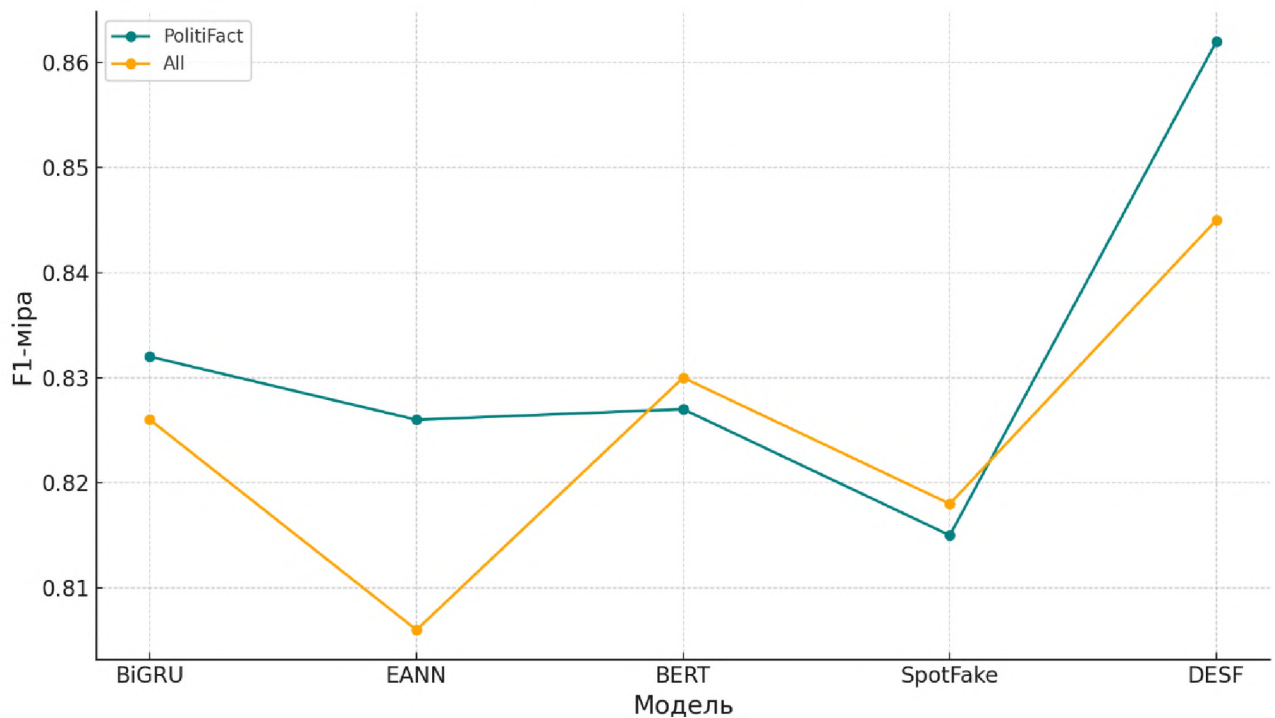


Рисунок 3.5 - Продуктивність на PolitiFact

Аналіз результатів експериментів показав, що запропонована модель DESF демонструє високі показники продуктивності в порівнянні з іншими базовими методами, досягаючи найкращих F1-міри у більшості доменів як на китайських, так і на англійських наборах даних. Модель DESF вирізняється стабільністю та універсальністю завдяки інтеграції семантичних, емоційних та доменних ознак. У порівнянні з альтернативними моделями, DESF показала перевагу у виявленні фейкових новин у таких доменах, як охорона здоров'я, економіка, стихійні лиха та суспільство, демонструючи загальну F1-міру 0.927 на наборі даних Weibo21 і 0.845 на FakeNewsNet. Незважаючи на відмінності в граматиці та структурі речень між китайською та англійською мовами, модель DESF забезпечила високу точність в обох мовних середовищах, що підтверджує її ефективність і перспективність для подальшого впровадження в автоматизованих системах виявлення фейкових новин.

3.3 Експеримент із абляцією

Для дослідження внеску кожного компонента в запропоновану модель, у цій роботі проводиться абляційне дослідження, щоб оцінити вплив різних ознак на алгоритм виявлення фейкових новин. Зокрема, експерименти налаштовані наступним чином: доменні ознаки, емоційні ознаки та семантичні ознаки по черзі вилучаються, щоб спостерігати зміни в продуктивності після їхнього видалення.

— Експеримент I: Видаляється етап видобування доменних ознак, і в експерименті поєднуються тільки емоційні та семантичні ознаки.

— Експеримент II: Видаляється етап видобування емоційних ознак, і в експерименті поєднуються тільки доменні та семантичні ознаки.

— Експеримент III: Видаляється етап видобування семантичних ознак, і в експерименті поєднуються тільки емоційні та доменні ознаки.

— Експеримент IV: Видаляється етап видобування емоційних ознак, пов'язаних з доменом, і використовуються лише семантичні ознаки для виявлення фейкових новин.

Ці чотири експерименти проводяться окремо на наборах даних Weibo21 та FakeNewsNet, щоб виміряти продуктивність і доцільність різних модулів, і порівнюються з моделлю DESF, запропонованою в цій роботі. Результати експериментів показані в таблиці 3.8.

Таблиця 3.8 - Результати абляційних експериментів, виконаних моделлю DESF на двох наборах даних

Набір даних	Метод	Точність	Точність (Фейкові новини)	Повнота (Фейкові новини)	F1-міра (Фейкові новини)	Точність (Реальні новини)	Повнота (Реальні новини)	F1-міра (Реальні новини)
Weibo21	I	0.834	0.839	0.851	0.845	0.829	0.647	0.727
	II	0.821	0.825	0.829	0.827	0.682	0.790	0.732
	III	0.817	0.819	0.824	0.821	0.547	0.698	0.613
	IV	0.789	0.792	0.796	0.794	0.685	0.740	0.711
	DESF	0.896	0.899	0.913	0.906	0.891	0.784	0.834
FakeNewsNet	I	0.823	0.803	0.834	0.818	0.832	0.849	0.840
	II	0.827	0.799	0.863	0.828	0.829	0.837	0.833
	III	0.835	0.732	0.798	0.764	0.846	0.852	0.849
	IV	0.799	0.657	0.721	0.689	0.801	0.805	0.803
	DESF	0.838	0.840	0.852	0.845	0.842	0.850	0.846

Аналіз показує, що модель DESF, запропонована в цій роботі, яка поєднує доменні емоційні ознаки та семантичні ознаки для виявлення фейкових новин, повністю використовує переваги злиття ознак та має абсолютні переваги. На наборі даних Weibo21 модель досягла точності 0.896, F1-міри 0.906 для виявлення фейкових новин і F1-міри 0.727 для виявлення реальних новин. На наборі даних FakeNewsNet точність становить 0.838, F1-міра для виявлення фейкових новин — 0.845, а для виявлення реальних новин — 0.834.

У Експерименті I, коли було видалено доменні ознаки і використовувалися лише емоційні та семантичні ознаки на наборі даних Weibo21, точність знизилася на 7.4% у порівнянні з моделлю DESF, а F1-міра знизилася на 7.2%.

На наборі даних FakeNewsNet точність знизилася на 1.8%, а F1-міра — на 3.3% у порівнянні з DESF, що свідчить про важливість доменних ознак.

У Експерименті II, коли були видалені емоційні ознаки і використовувалися лише доменні та семантичні ознаки, точність на Weibo21 становила 0.821, а F1-міра — 0.827, що означає зниження на 9.1% та 9.6% відповідно у порівнянні з DESF. На FakeNewsNet точність становила 0.827, а F1-міра — 0.828, що означає зниження на 1.3% та 2.1% відповідно у порівнянні з DESF, що свідчить про важливість емоційних ознак для виявлення фейкових новин.

Порівнюючи Експерименти I та II, можна побачити, що на наборі Weibo21 продуктивність після видалення емоційних ознак знижується швидше, ніж після видалення доменних ознак, що вказує на те, що емоційні ознаки відіграють більш важливу роль.

У Експерименті III, коли були видалені семантичні ознаки, а для виявлення фейкових новин використовувалися лише емоційні та доменні ознаки, точність знизилася на 9.7%, а F1-міра — на 10.4% у порівнянні з моделлю DESF, що свідчить про значне зниження продуктивності в порівнянні з Експериментами I та II. Це вказує на те, що семантичні ознаки допомагають у розумінні та оцінці новин, і важливо повністю оволодіти семантичною інформацією під час виявлення фейкових новин для полегшення розуміння текстового змісту новин.

Аналізуючи результати Експерименту IV, в якому використовувалися лише семантичні ознаки без механізму уваги для захоплення емоційних ознак, пов'язаних із доменом, продуктивність була найнижчою серед усіх п'яти моделей. У порівнянні з DESF, точність знизилася на 13.6%, а F1-міра — на 14.1% на наборі Weibo21, і точність на наборі FakeNewsNet знизилася на 4.9%. Це демонструє, що основний модуль, запропонований у цій роботі (видобування емоційних ознак, пов'язаних із доменом), позитивно впливає на завдання виявлення фейкових новин. Візуальне порівняння результатів абляційних експериментів показано на рисунках 3.6 та 3.7.

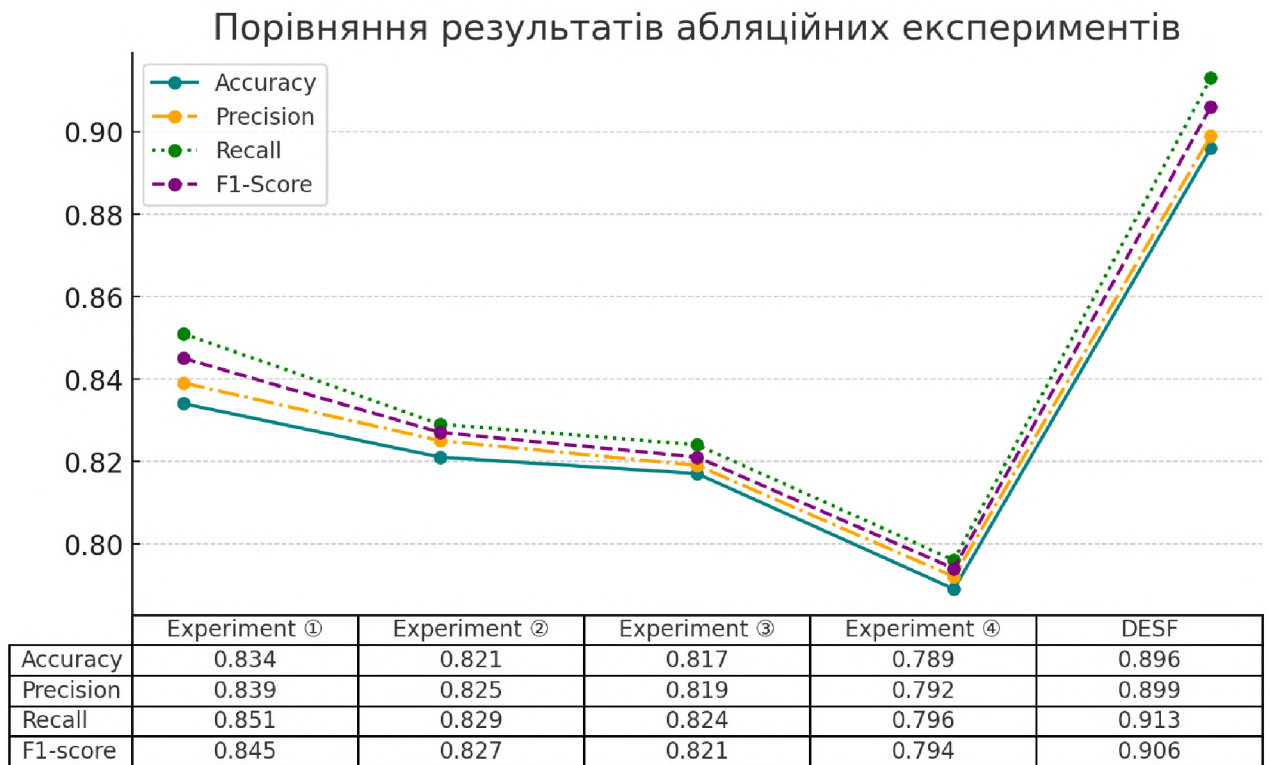


Рисунок 3.6 - Експерименти з абляцією на Weibo21

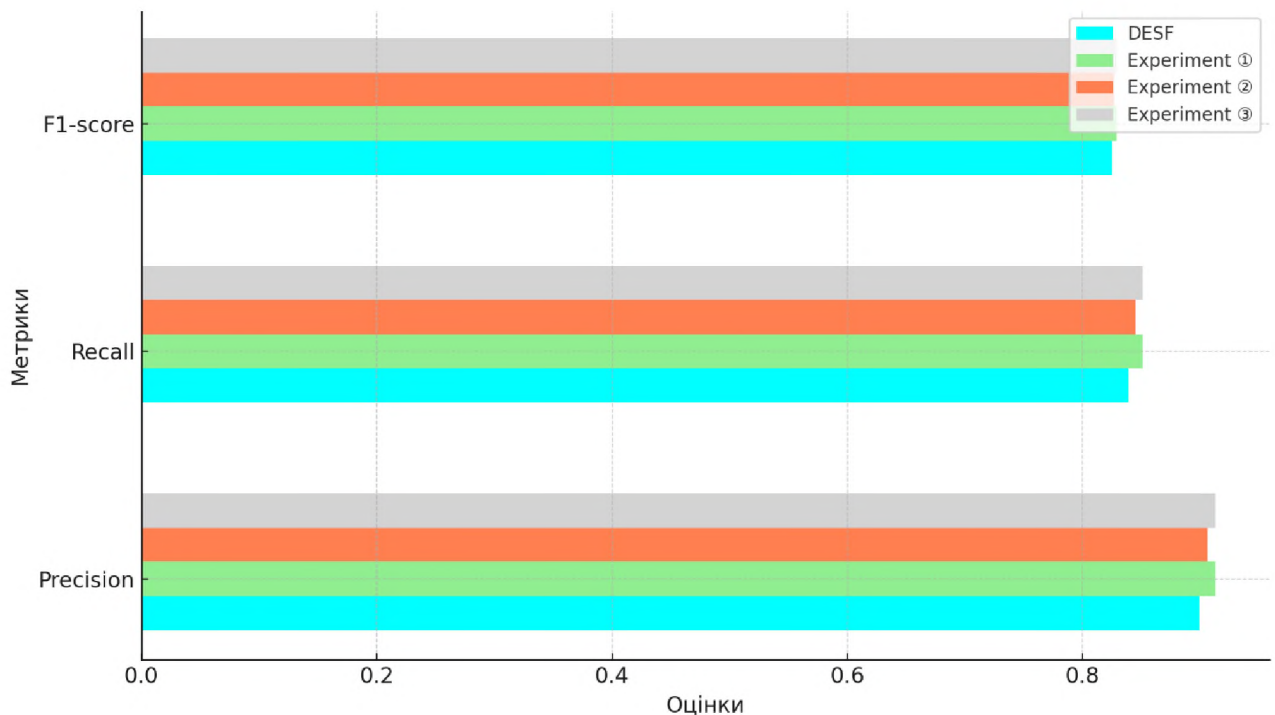


Рисунок 3.7 - Експерименти з абляцією у FakeNewsNet

Отримані результати абляційних експериментів чітко демонструють важливість кожного з компонентів моделі DESF для успішного виявлення фейкових новин. Видалення будь-якої з ключових ознак (доменних, емоційних або семантичних) значно знижує точність і F1-міру моделі, що підкреслює синергію цих компонентів у злитті ознак. Зокрема, емоційні та семантичні ознаки виявилися найважливішими, оскільки їх видалення спричиняло найбільше падіння продуктивності моделі, особливо на наборі даних Weibo21. Таким чином, комбіноване використання доменних, емоційних і семантичних ознак у моделі DESF забезпечує стабільну та високу ефективність у виявленні фейкових новин, що підтверджується результатами на обох наборах даних Weibo21 та FakeNewsNet.

Висновки до розділу 3

1. У процесі навчання моделі для запобігання перенавчанню було успішно застосовано метод Dropout з рівнем 0.5, а також L2-регуляризацію для контролю ваг. Це дозволило забезпечити стабільність та загальну ефективність моделі під час тренування, що підтверджується високими результатами моделі на двох різних наборах даних — FakeNewsNet та Weibo21. Використання цих методів у поєднанні з оптимізатором Adam гарантувало ефективне управління навчанням та зменшення нестабільності під час навчання моделі.

2. Проведене дослідження показало, що модель DESF демонструє високу здатність до узагальнення в різних тематичних доменах. Набори даних Weibo21 та FakeNewsNet охоплюють широкий спектр тематичних областей, і результати моделі показують її стабільність та точність у таких доменах, як охорона здоров'я, фінанси, розваги та інші. Модель досягла найкращих показників у доменах охорони здоров'я (F1-міра 0.959), економіки (F1-міра 0.927), що свідчить про її здатність ефективно працювати з міждоменними даними.

3. Аналіз абляційних експериментів підтвердив важливість кожного компонента моделі DESF у досягненні високих результатів. Видалення будь-якої з ключових ознак (доменних, емоційних або семантичних) призводило до значного зниження точності та F1-міри, що підтверджує їхню синергетичну роль. Зокрема, емоційні та семантичні ознаки виявилися найбільш критичними, оскільки їх видалення спричиняло найбільше падіння продуктивності, особливо на наборі даних Weibo²¹. Це демонструє, що поєднання цих ознак є ключовим для успішного виявлення фейкових новин.

ВИСНОВКИ

1. У ході дослідження була розроблена модель DESF, яка поєднує три ключові типи ознак: доменні, емоційні та семантичні. Завдяки їхньому злиттю, модель забезпечує високу точність у виявленні фейкових новин. На наборі даних Weibo21 модель DESF досягла точності 0.896 і F1-міри 0.906 для виявлення фейкових новин. У домені FakeNewsNet модель також показала високі результати, з точністю 0.838 і F1-мірою 0.845 для фейкових новин.

2. Модель DESF перевершила кілька існуючих методів, таких як BiGRU, EANN, BERT та SpotFake, на обох наборах даних. Зокрема, у домені PolitiFact DESF досягла F1-міри 0.862, що на 3.6% вище, ніж у BiGRU, та на 5.8% вище, ніж у SpotFake. У домені GossipCop DESF показала F1-міру 0.828, що на 0.2% перевищило результати моделі CFPFND.

3. Абляційні експерименти показали, що видалення будь-якого з компонентів моделі призводить до зниження продуктивності. Зокрема, видалення доменних ознак знизило точність моделі на 7.4% на Weibo21, а видалення емоційних ознак призвело до зниження точності на 9.1%. Це свідчить про важливість кожного з компонентів моделі для її загальної продуктивності.

4. Модель DESF продемонструвала високу здатність до узагальнення на різних тематичних доменах. У домені охорони здоров'я вона досягла F1-міри 0.959, що на 0.002 більше за другу за результатами модель. У домені економіки F1-міра DESF становила 0.927, що на 5.9% вище, ніж у MDFEND. Це свідчить про можливість моделі адаптуватися до різних типів даних.

5. Аналізуючи результати моделі на англійському та китайському наборах даних, було виявлено, що DESF показує кращі результати на китайському наборі Weibo21. Зокрема, на цьому наборі модель досягла точності 0.896 і F1-міри 0.906. Це може бути зумовлено різницею в граматичних структурах та особливостях обробки китайської та англійської мов.

6. У порівнянні з іншими моделями, DESF демонструє стабільний прогрес у завданні виявлення фейкових новин. Високі результати на двох різних наборах даних підтверджують, що інтеграція доменних, емоційних та семантичних ознак є ефективною стратегією для підвищення точності та здатності до узагальнення моделей класифікації новин.

7. Результати дослідження показують, що модель DESF має значний потенціал для подальшого вдосконалення. Комбіноване використання різних типів ознак забезпечує її високу продуктивність, проте можливості моделі можуть бути ще більше розширені шляхом додавання нових типів ознак або покращення механізмів уваги. Це відкриває перспективи для подальших досліджень і застосування моделі в інших сферах.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Phan, H. T., Nguyen, N. T., & Hwang, D. (2023). Fake news detection: A survey of graph neural network methods. *Applied Soft Computing*, 139, 110235. <https://doi.org/10.1016/j.asoc.2023.110235>
2. Capuano, N., Fenza, G., Loia, V., & Di Nota, F. (2023). Content-based fake news detection with machine and deep learning: A systematic review. *Neurocomputing*, 530, 91-103. <https://doi.org/10.1016/j.neucom.2023.02.005>
3. Jin, Z., Cao, J., Guo, H., Zhang, Y., & Luo, J. (2017). Multimodal fusion with recurrent neural networks for rumor detection on microblogs. *Proceedings of the 25th ACM International Conference on Multimedia*, 795-816. <https://doi.org/10.1145/3123266.3123454>
4. Khattar, D., Goud, J. S., Gupta, M., & Varma, V. (2019). MVAE: Multimodal variational autoencoder for fake news detection. *Proceedings of the 2019 World Wide Web Conference*, 2915-2921. <https://doi.org/10.1145/3308558.3313552>
5. Zeng, J., Zhang, Y., & Ma, X. (2021). Fake news detection for epidemic emergencies via deep correlations between text and images. *Sustainable Cities and Society*, 66, 102652. <https://doi.org/10.1016/j.scs.2020.102652>
6. Delgado, M., Manín, N., Martín-Bautista, M. J., Sánchez, D., & Vila, M. A. (2005). Mining fuzzy association rules: An overview. *Soft Computing for Information Processing and Analysis*, 351-373. <https://doi.org/10.1007/3-540-32365-1>
7. Feng, H., Lesot, M. J., & Detyniecki, M. (2010). Using association rules to discover color-emotion relationships based on social tagging. *Lecture Notes in Computer Science*, 6276, 544-553. https://doi.org/10.1007/978-3-642-15387-7_55
8. Sheydaei, N., Saraee, M., & Shahgholian, A. (2015). A novel feature selection method for text classification using association rules and clustering. *Journal of Information Science*, 41(1), 3-15. <https://doi.org/10.1177/0165551514550143>
9. Rohidin, D., Samsudin, N. A., & Deris, M. M. (2022). Association rules of fuzzy soft set based classification for text classification problem. *Journal of King Saud*

University – Computer and Information Sciences, 34(3), 801-812.
<https://doi.org/10.1016/j.jksuci.2020.03.014>

10. Cai, Y., & Wan, X. (2019). Multi-domain sentiment classification based on domain-aware embedding and attention. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, 4904-4910.
<https://doi.org/10.18653/v1/D19-1496>

11. Ji, S., Pan, S., Cambria, E., Marttinen, P., & Yu, P. S. (2021). A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE Transactions on Neural Networks and Learning Systems*, 33(2), 494-514.
<https://doi.org/10.1109/TNNLS.2021.3070843>

12. Liu, J. S., Feng, K., Pan, J., Deng, J. Z., & Wang, L. N. (2020). MSRD: Multimodal network rumor detection method. *Computer Research and Development*, 57(11), 2328-2336. https://doi.org/10.1007/978-3-030-47250-4_20

13. Li, A., Dan, Z. P., Dong, F. M., & Liu, L. W., Feng, Y. (2020). Rumor detection method based on improved generative adversarial network. *Journal of Chinese Information Science*, 34(9), 78-88. https://doi.org/10.1007/978-981-13-8139-0_8

14. Sahoo, S. R., & Gupta, B. B. (2021). Multiple features based approach for automatic fake news detection on social networks using deep learning. *Applied Soft Computing*, 100, 106983. <https://doi.org/10.1016/j.asoc.2020.106983>

15. Zhang, Q., Guo, Z., Zhu, Y., Vijayakumar, P., Castiglione, A., & Gupta, B. B. (2023). A deep learning-based fast fake news detection model for cyber-physical social services. *Pattern Recognition Letters*, 168, 31-38.
<https://doi.org/10.1016/j.patrec.2023.02.026>

16. Luo, Z., Li, Q., & Zheng, J. (2021). Deep feature fusion for rumor detection on Twitter. *IEEE Access*, 9, 126065-126074.
<https://doi.org/10.1109/ACCESS.2021.3111790>

17. Wang, J., Mao, H., & Li, H. (2022). FMFN: Fine-grained multimodal fusion networks for fake news detection. *Applied Sciences*, 12(3), 1093. <https://doi.org/10.3390/app12031093>
18. Li, S., Yao, T., Li, S., & Yan, L. (2022). Semantic-enhanced multimodal fusion network for fake news detection. *International Journal of Intelligent Systems*, 37(12), 12235-12251. <https://doi.org/10.1002/int.23084>
19. Guo, Y. (2023). A mutual attention based multimodal fusion for fake news detection on social networks. *Applied Intelligence*, 53(12), 15311-15320. <https://doi.org/10.1007/s10489-022-04266-w>
20. Silva, A., Luo, L., Karunasekera, S., & Leckie, C. (2021). Embracing domain differences in fake news: Cross-domain fake news detection using multi-modal data. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(1), 557-565. <https://doi.org/10.1609/aaai.v35i1.16134>
21. Mridha, M. F., Keya, A. J., Hamid, M. A., Monowar, M. M., & Rahman, M. S. A. (2021). A comprehensive review on fake news detection with deep learning. *IEEE Access*, 9, 156151-156170. <https://doi.org/10.1109/ACCESS.2021.3129329>
22. Liang, F., Shu, Y. F., Lei, M. Y., Zhou, M., Kang, J. Y., & Guo, Q. L. (2024). False news detection based on adaptive fusion of multimodal features. *Journal of Computational Science*, 53, 228-234. <https://doi.org/10.1016/j.jocs.2023.11.012>
23. Shu, K., Mahudeswaran, D., Wang, S., Lee, D., & Liu, H. (2020). FakeNewsNet: A data repository with news content, social context, and spatiotemporal information for studying fake news on social media. *Big Data*, 8(3), 171-188. <https://doi.org/10.1089/big.2020.0062>
24. Nan, Q., Cao, J., Zhu, Y., Wang, Y., & Li, J. (2021). MDFEND: Multi-domain fake news detection. *Proceedings of the 2021 IEEE International Conference on Big Data*, 3343-3347. <https://doi.org/10.1109/BigData.2021.9671863>
25. Liang, C., Zhang, Y., Li, X., Zhang, J., & Yu, Y. (2022). FuDFEND: Fuzzy-domain for multi-domain fake news detection. *Natural Language Processing and Chinese Computing*, 13552, 45-57. https://doi.org/10.1007/978-3-031-17189-5_6

26. Wang, D., Zhang, W., Wu, W., & Guo, X. (2023). Soft-label for multi-domain fake news detection. *IEEE Access*, 11, 98596-98606. <https://doi.org/10.1109/ACCESS.2023.3313602>
27. Zhu, Y. (2023). Memory-guided multi-view multi-domain fake news detection. *IEEE Transactions on Knowledge and Data Engineering*, 35(7), 7178-7191. <https://doi.org/10.1109/TKDE.2022.3185151>
28. Wang, Z. Q., Chen, T., Zhang, B. Y., Zhang, M. L., Sun, C. Y., & Zhang, W. S. (2024). Multi-domain fake news detection based on cross-feature perception fusion. *Computer Systems and Applications*, 33(3), 264-272. <https://doi.org/10.15888/j.cnki.csa.009439>
29. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 4171-4186. <https://doi.org/10.18653/v1/N19-1414>
30. Ma, J. (2016). Detecting rumors from microblogs with recurrent neural networks. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 3818-3824. <https://doi.org/10.18653/v1/D16-1395>
31. Singhal, S., Shah, R. R., Chakraborty, T., Kumaraguru, P., & Satoh, S. (2019). SpotFake: A multi-modal framework for fake news detection. *Proceedings of the 2019 ACM International Conference on Multimedia Retrieval*, 39-47. <https://doi.org/10.1145/3323873.3325040>
32. Wang, Y. (2018). EANN: Event adversarial neural networks for multi-modal fake news detection. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 849-857. <https://doi.org/10.18653/v1/D18-1084>
33. Загальні методичні рекомендації з підготовки, оформлення, захисту та оцінювання кваліфікаційних робіт здобувачів вищої освіти першого (бакалаврського) і другого (магістерського) рівнів. Тернопіль: ЗУНУ, 2024. 83 с.

34. Комар М.П., Саченко А.О., Васильків Н.М., Загородня Д.І. Методичні рекомендації до виконання кваліфікаційної роботи з освітньо-професійної програми «Комп'ютерні науки» спеціальності 122 «Комп'ютерні науки» за другим (магістерським) рівнем вищої освіти. – Тернопіль: ЗУНУ, 2024. – 32 с.

Додаток А
Апробація отриманих результатів

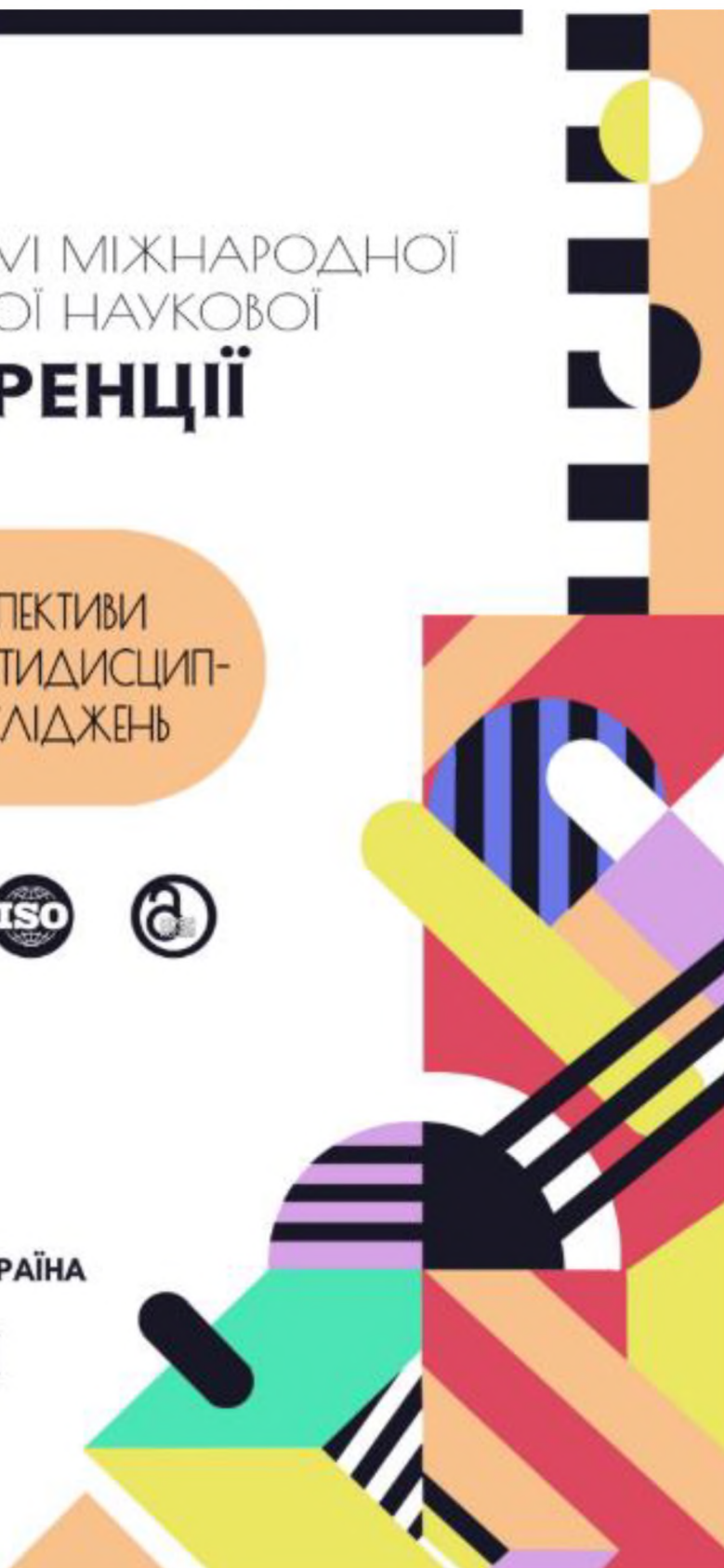
МАТЕРІАЛИ VI МІЖНАРОДНОЇ
СТУДЕНТСЬКОЇ НАУКОВОЇ
КОНФЕРЕНЦІЇ

ТРЕНДИ ТА ПЕРСПЕКТИВИ
РОЗВИТКУ МУЛЬТИДИСЦИП-
ЛІНАРНИХ ДОСЛІДЖЕНЬ



М. КРЕМЕНЧУК, УКРАЇНА

**11 ЖОВТНЯ
2024 РІК**



УДК 082:001

Т 66



Голова оргкомітету: Коренюк І.О.

Верстка: Зрада С.І.

Дизайн: Бондаренко І.В.

Рекомендовано до видання Вченою Радою Інституту науково-технічної інтеграції та співпраці. Протокол № 57 від 10.10.2024 року.



Конференцію зареєстровано Державною науковою установою «УкрІНТЕІ» в базі даних науково-технічних заходів України та бюлетені «План проведення наукових, науково-технічних заходів в Україні» (Посвідчення №320 від 12.06.2024).

Матеріали конференції знаходяться у відкритому доступі на умовах ліцензії Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

Т 66

Тренди та перспективи розвитку мультидисциплінарних досліджень: матеріали VI Міжнародної студентської наукової конференції, м. Кременчук, 11 жовтня, 2024 рік / ГО «Молодіжна наукова ліга». — Вінниця: ТОВ «УКРЛОГОС Груп», 2024. — 198 с.

ISBN 978-617-8440-34-3

DOI 10.62732/liga-inter-11.10.2024

Викладено матеріали учасників VI Міжнародної мультидисциплінарної студентської наукової конференції «Тренди та перспективи розвитку мультидисциплінарних досліджень», яка відбулася 11 жовтня 2024 року у місті Кременчук, Україна.

УДК 082:001

ISBN 978-617-8440-34-3

© Колектив учасників конференції, 2024

© ГО «Молодіжна наукова ліга», 2024

© ТОВ «УКРЛОГОС Груп», 2024

СЕКЦІЯ 11.**АГРАРНІ НАУКИ ТА ПРОДОВОЛЬСТВО**

ЕКОСИСТЕМНІ ПОСЛУГИ ЛІСОВОГО ГОСПОДАРСТВА

Луців В.В., Науковий керівник: Бондар О.Б. 95

СЕКЦІЯ 12.**ЕЛЕКТРОНІКА ТА ТЕЛЕКОМУНІКАЦІЇ**

ПРИВАТНІСТЬ ТА ЕТИЧНІ ПИТАННЯ ВІДЕОПОСТЕРЕЖЕННЯ У ГРОМАДСЬКИХ МІСЦЯХ

Зеленський Д.М. 97

СЕКЦІЯ 13.**КОМП'ЮТЕРНА ТА ПРОГРАМНА ІНЖЕНЕРІЯ**

LOGIC-FREE ВЕКТОРНИЙ КОМП'ЮТИНГ

Лук'янов Я.І., Науковий керівник: Хаханова Г.В. 99

THE EVOLUTION OF SOFTWARE ENGINEERING: FROM CODE MONKEYS TO ARCHITECTS

Almakadma M.I., Pershyna A.A. 103

СЕКЦІЯ 14.**ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ТА СИСТЕМИ**

ВИКОРИСТАННЯ ІНСТРУМЕНТІВ AUTOML ДЛЯ ОПТИМІЗАЦІЇ ОБРОБКИ ДАНИХ

Власенко Є.С., Науковий керівник: Слюсар В.І. 105

ДОСЛІДЖЕННЯ МОЖЛИВОСТЕЙ ВИКОРИСТАННЯ СУЧАСНИХ ТЕХНОЛОГІЙ ДЛЯ ДВОФАКТОРНОЇ АУТЕНТИФІКАЦІЇ НА ОСНОВІ БІОМЕТРИЧНИХ ПАРАМЕТРІВ

Безмалій О.Р., Науковий керівник: Висоцька О.О. 107

ІНТЕГРАЦІЯ МЕСЕНДЖЕРА ТА GPT ДЛЯ РЕАЛІЗАЦІЇ ВІРТУАЛЬНОГО АСИСТЕНТА

Перевозкін А.С., Науковий керівник: Слюсар І.І. 110

ІНТЕЛЕКТУАЛЬНІ МЕТОДИ ВИЯВЛЕННЯ МАНІПУЛЯТИВНОГО КОНТЕНТУ

Колівушко Е., Лучка С., Матвійчук М., Науковий керівник: Лип'яніна-Гончаренко Х. 112

МЕТОД ПРОГНОЗУВАННЯ ЗАТРИМОК ТА ПЕРЕВИТРАТ В ІНЖЕНЕРНИХ ПРОЕКТАХ ЗА ДОПОМОГОЮ МОДЕЛЕЙ ГЛИБИННОГО НАВЧАННЯ

Сулима Б.Я. 114

ШТУЧНИЙ ІНТЕЛЕКТ - РУШІЙНА СИЛА ДЛЯ РОЗВИТКУ РІЗНОМАНІТНИХ ТЕХНОЛОГІЧНИХ СФЕР ДІЯЛЬНОСТІ ЛЮДИНИ

Черпаха М.О., Науковий керівник: Сердюк Н.М. 116

Тренди та перспективи розвитку мультидисциплінарних досліджень

Колівушко Едуард, здобувач вищої освіти факультету комп'ютерних інформаційних технологій
Західноукраїнський національний університет, Україна

Лучка Святослав, здобувач вищої освіти факультету комп'ютерних інформаційних технологій
Західноукраїнський національний університет, Україна

Матвійчук Микола, здобувач вищої освіти факультету комп'ютерних інформаційних технологій
Західноукраїнський національний університет, Україна

Науковий керівник: Ліп'яніна-Гончаренко Христина, канд. техн. наук, доцент, доцент кафедри інформаційно-обчислювальних систем та управління
Західноукраїнський національний університет, Україна

ІНТЕЛЕКТУАЛЬНІ МЕТОДИ ВИЯВЛЕННЯ МАНІПУЛЯТИВНОГО КОНТЕНТУ

З розвитком цифрових технологій та глобальної мережі Інтернет дезінформація і маніпулятивний контент стали невід'ємною частиною сучасного інформаційного простору. Швидке поширення фейкових новин і дипфейків негативно впливає на суспільну думку, політичні процеси та економічні відносини. За даними досліджень, 86% користувачів Інтернету щонайменше один раз стикалися з фейковими новинами [5]. Це підвищує необхідність розробки ефективних технологій для автоматичного виявлення такого контенту.

Основною складністю виявлення дезінформації є її постійна еволюція та складність у визначенні об'єктивної правдивості інформації. Додатковим викликом є те, що фейкові новини можуть бути дуже схожими на правдиві за своїм стилем і структурою, а дипфейки часто використовують високотехнологічні методи, що робить їх майже непомітними для людського ока [2]. Це вимагає впровадження інтелектуальних систем на базі машинного навчання для більш ефективного розпізнавання маніпуляцій.

Для автоматичного виявлення фейкових новин широко використовуються алгоритми машинного навчання, які здатні аналізувати текстові дані, виявляти закономірності та класифікувати контент на правдивий або маніпулятивний. Серед найбільш популярних методів можна виділити логістичну регресію, метод опорних векторів (SVM) та нейронні мережі. Наприклад, нейронні мережі показали високий рівень точності при класифікації текстів, аналізуючи синтаксичні та семантичні ознаки [3].

Текстові методи виявлення маніпулятивного контенту мають низку переваг, таких як швидка обробка та висока точність при виявленні патернів у тексті. Однак їхня ефективність знижується при роботі з неоднозначними чи контекстно залежними даними. Відеоаналітичні методи, такі як виявлення дипфейків, можуть виявити візуальні аномалії, проте вимагають значних обчислювальних ресурсів і складних моделей для точного розпізнавання підробленого контенту [1], [2].

Моделі машинного навчання показують різну продуктивність залежно від типу

контенту. Для текстових даних методи, засновані на класифікації, такі як логістична регресія чи LSTM, досягають точності понад 90% при аналізі новинних статей [3]. Для відеоконтенту глибокі нейронні мережі, такі як CNN або GAN, успішно виявляють дипфейки, проте їхня точність може коливатися залежно від якості відео і складності підробки, досягаючи 80-90% точності у складних випадках [1].

Для підвищення точності виявлення маніпулятивного контенту варто комбінувати різні підходи, що дозволить аналізувати і текстову, і відеоінформацію. Наприклад, інтеграція текстових моделей для попередньої фільтрації новин з подальшим аналізом відео чи зображень може значно підвищити ефективність. Крім того, використання гібридних моделей, що поєднують машинне та глибоке навчання, може забезпечити більш гнучке розпізнавання різних типів контенту [4].

У результаті дослідження було встановлено, що машинне навчання є потужним інструментом для виявлення маніпулятивного контенту, проте його ефективність залежить від типу даних та складності підробки. Комбінування текстових і відеоаналітичних методів може суттєво підвищити точність і зменшити ймовірність помилкових спрацювань.

У майбутньому можна очікувати подальшого розвитку інтелектуальних систем виявлення дезінформації, зокрема завдяки вдосконаленню глибоких нейронних мереж та їхньої інтеграції з іншими алгоритмами. Особливий інтерес представляє використання мультимодальних підходів, що дозволяють аналізувати текст, аудіо та відео одночасно, що зробить виявлення маніпуляцій ще більш точним та оперативним.

Список використаних джерел:

1. H. Nguyen, J. Yamagishi, and I. Echizen. "Capsule-forensics: Using capsule networks to detect forged images and videos." ICASSP 2019.
2. T. Korshunov and S. Marcel. "Deepfakes: A New Threat to Face Recognition? Assessment and Detection." IEEE International Conference on Biometrics, 2020.
3. D. Shu, H. Mahmud, and P. Resnick. "Fake News Detection on Social Media: A Data Mining Perspective." ACM SIGKDD, 2019.
4. C. Conroy, V. Rubin, and Y. Chen. "Automatic deception detection: Methods for finding fake news." Proceedings of the Association for Information Science and Technology, 2015.
5. Statista. "Share of internet users worldwide who encountered fake news as of January 2021." [Online]. Available: www.statista.com.

ЗБІРНИК НАУКОВИХ ПРАЦЬ

З МАТЕРІАЛАМИ ІІІ МІЖНАРОДНОЇ НАУКОВОЇ КОНФЕРЕНЦІЇ

20 ВЕРЕСНЯ 2024 РІК

М. ОДЕСА, УКРАЇНА

**«ІНТЕЛЕКТУАЛЬНИЙ РЕСУРС СЬОГОДЕННЯ:
НАУКОВІ ЗАДАЧІ, РОЗВИТОК ТА ЗАПИТАННЯ»**



УДК 082:001
I-57



Організація, від імені якої випущено видання:

ГО «Міжнародний центр наукових досліджень»

Номер запису організації в Єдиному реєстрі громадських об'єднань: 1499141.

Голова оргкомітету: Сотник С.Г.

Верстка: Бабич Ю.В.

Дизайн: Бондаренко І.В.

Рекомендовано до видання Вченою Радою Інституту науково-технічної інтеграції та співпраці. Протокол № 54 від 19.09.2024 року.



Конференцію зареєстровано Державною науковою установою у сфері управління Міністерства освіти і науки «Український інститут науково-технічної експертизи та інформації» в базі даних науково-технічних заходів України на поточний рік та бюлетені «План проведення наукових, науково-технічних заходів в Україні» (Посвідчення № 348 від 12.06.2024).

Збірник наукових праць з матеріалами конференції видано офіційно суб'єктом видавничої справи зі **Свідоцтвом ДК № 7860 від 22.06.2023**

Матеріали конференції знаходяться у відкритому доступі на умовах ліцензії Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

I-57 **Інтелектуальний ресурс сьогодення: наукові задачі, розвиток та запитання:** збірник наукових праць з матеріалами III Міжнародної наукової конференції, м. Одеса, 20 вересня, 2024 р. / Міжнародний центр наукових досліджень. — Вінниця: ТОВ «УКРЛОГОС Груп, 2024. — 322 с.

ISBN 978-617-8440-13-8

DOI 10.62731/mcnd-20.09.2024

Викладено матеріали учасників III Міжнародної наукової конференції «Інтелектуальний ресурс сьогодення: наукові задачі, розвиток та запитання», яка відбулася 20 вересня 2024 року у місті Одеса.

УДК 082:001

© Колектив учасників конференції, 2024

© ГО «Міжнародний центр наукових досліджень», 2024

ISBN 978-617-8440-13-8

© ТОВ «УКРЛОГОС Груп», 2024

**СЕКЦІЯ X.
ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ТА СИСТЕМИ**

ОГЛЯД КОНЦЕПЦІЙ ТА МЕТОДІВ МАШИННОГО НАВЧАННЯ: ТЕОРІЯ,
ЗАСТОСУВАННЯ ТА ВИКЛИКИ
Кондратюк Г., Кравчук Б., Тарасюк С., Матвійчук М. 177

**СЕКЦІЯ XI.
ТРАНСПОРТ ТА ТРАНСПОРТНІ ТЕХНОЛОГІЇ**

INNOVATIVE TECHNOLOGIES AND REGULATORY MEASURES TO REDUCE
ENVIRONMENTAL RISKS IN THE SHIPPING INDUSTRY
Sagaydak O., Kucherenko V., Kotenko O., Prokhorov V. 181

**СЕКЦІЯ XII.
ФІЗИКО-МАТЕМАТИЧНІ НАУКИ**

ПЕРЕХІД МЕТАЛ-ІЗОЛЯТОР У МОНОКРИСТАЛАХ $Y_{1-x}Pr_xBa_2Cu_3O_{7-\delta}$ З РІЗНИМ
ВМІСТОМ ПРАЗЕОДИМУ
Камчатна С. М., Ярчук Д. Ф., Чепурін О. Г., Іноземцев М. М., Вовк Р. В. 187

**СЕКЦІЯ XIII.
СОЦІОЛОГІЯ ТА СТАТИСТИКА**

СІМ'Я ЯК СОЦІАЛЬНИЙ ІНСТИТУТ. ФУНКЦІЇ СІМ'І
Олійник М. В. 191

**СЕКЦІЯ XIV.
ФІЛОЛОГІЯ ТА ЖУРНАЛІСТИКА**

FROM BEATS TO WORDS: HOW MUSIC NORMALIZES CREATIVE AND IMPLICIT
ENGLISH LANGUAGE
Hulei T. 196

ФУНКЦІОНУВАННЯ НЕФОРМАЛЬНОЇ ЛЕКСИКИ У СУЧАСНИХ ПІСНЯХ
Борисова К. М. 205

АВТОР І ТЕКСТ У РОМАНІ У. ЕКО «ІМ'Я ТРОЯНДИ»
Кулик К. В. 209

ЗАСОБИ ПЕРЕКЛАДУ РЕАЛІЙ В ІСТОРИЧНОМУ ДЕТЕКТИВІ (НА МАТЕРІАЛІ
РОМАНУ К. ДЖ. СЕНСОМА «DISSOLUTION»)
Марченко В. В. 212

СЕКЦІЯ X. ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ТА СИСТЕМИ

ОГЛЯД КОНЦЕПЦІЙ ТА МЕТОДІВ МАШИННОГО НАВЧАННЯ: ТЕОРІЯ, ЗАСТОСУВАННЯ ТА ВИКЛИКИ

Кондратюк Георгій

здобувач вищої освіти факультету комп'ютерних інформаційних технологій
Західноукраїнський національний університет, Україна

Кравчук Богдан

здобувач вищої освіти факультету комп'ютерних інформаційних технологій
Західноукраїнський національний університет, Україна

Тарасюк Софія

здобувач вищої освіти факультету комп'ютерних інформаційних технологій
Західноукраїнський національний університет, Україна

Матвійчук Микола

здобувач вищої освіти факультету комп'ютерних інформаційних технологій
Західноукраїнський національний університет, Україна

Науковий керівник: Ліп'яніна-Гончаренко Христина

канд. тех. наук, доцент, доцент кафедри інформаційно-обчислювальних
систем та управління
Західноукраїнський національний університет, Україна

Машинне навчання є однією з ключових технологій, що трансформують різні галузі, від медицини до інженерії та промисловості. З розвитком обчислювальної потужності та зростанням обсягів даних, методи машинного навчання стають все більш складними та ефективними. Однак, щоб повністю зрозуміти та застосувати ці методи, необхідно володіти глибоким знанням теоретичних концепцій та практичних підходів. У цьому огляді зібрано ключові публікації, які досліджують різні аспекти машинного навчання, починаючи від основ статистичного навчання і закінчуючи використанням машинного

навчання в різних галузях, таких як медицина, телекомунікації та автоматизація. Метою цього огляду є узагальнення та систематизація знань у цій сфері для кращого розуміння сучасних методів машинного навчання.

У публікації "Machine Learning Methods: An Overview" [1] автор R. Muhamedyev надає загальний огляд методів машинного навчання, їх класифікацію та застосування. Стаття розглядає різні категорії алгоритмів, включаючи навчання з учителем і без учителя, а також надає практичні приклади їх використання. Особлива увага приділяється аналізу переваг та обмежень різних підходів, що робить цю роботу корисною для початківців у цій сфері.

A.F.A.H. Alnuaimi та T.H.K. Albaldawi у своїй роботі [2] представляють фундаментальні концепції статистичного навчання та класифікації в машинному навчанні. Автори висвітлюють основні методи та алгоритми, зокрема регресію та класифікацію, і пояснюють їх роль у вирішенні складних завдань. Ця стаття акцентує увагу на важливості статистичних підходів у машинному навчанні та їх впливі на точність моделей.

У статті "Logical Approaches to Machine Learning" [3] автор P. Flach досліджує логічні підходи до машинного навчання. Ця публікація підкреслює важливість логіки та математичних основ для створення ефективних алгоритмів навчання. Вона розглядає різні типи логічних моделей та їхню здатність вирішувати складні завдання, забезпечуючи глибше розуміння теоретичних аспектів машинного навчання.

У публікації [4] автори S. Albahra та співавтори досліджують застосування методів машинного навчання та штучного інтелекту в патології та лабораторній медицині. Стаття надає огляд ключових алгоритмів та їх застосувань у медичному аналізі, демонструючи, як ці методи можуть покращити діагностику та підвищити ефективність медичних процедур.

H.M. El Misilmani та T. Naous у своїй роботі [5] розглядають використання машинного навчання в проєктуванні антен. Вони описують різні концепції машинного навчання та алгоритми, які можуть бути застосовані для оптимізації антен. Стаття надає практичні приклади використання машинного навчання у галузі телекомунікацій, демонструючи його ефективність та потенціал для інженерних задач.

У публікації [6] автори D. Kreuzberger, N. Kühl, та S. Hirschl досліджують концепцію Machine Learning Operations (MLOps). Вони

надають визначення MLOps, пояснюють його архітектуру та важливість для ефективного розгортання моделей машинного навчання в промислових додатках. Ця стаття є ключовою для розуміння того, як інтегрувати машинне навчання у виробниче середовище.

J.G. Carbonell, R.S. Michalski, та T.M. Mitchell у своїй класичній роботі [7] надають фундаментальний огляд концепцій машинного навчання. Вони пропонують критерії для класифікації та порівняння методів машинного навчання, висвітлюючи різні аспекти, такі як представлення знань та алгоритми навчання. Ця стаття є однією з перших, яка систематизувала знання у цій галузі.

У публікації "An Overview of Machine Learning Techniques in Constraint Solving" [8] автори A. Popescu та інші розглядають, як машинне навчання може використовуватися для оптимізації процесів вирішення обмежень. Вони аналізують різні підходи та методи, зосереджуючись на використанні машинного навчання для автоматизації та покращення процесу вирішення задач з обмеженнями.

T.O. Ayodele у своїй праці [9] надає таксономічний аналіз машинного навчання, організовуючи його на основі різних підходів до навчання. Стаття розглядає різні категорії алгоритмів та підходів до машинного навчання, пропонуючи вичерпний огляд методів та їх застосувань у різних галузях.

У статті "An Overview of Privacy in Machine Learning" [10] автор E. De Cristofaro аналізує концепції конфіденційності в машинному навчанні. Ця робота досліджує, як забезпечити конфіденційність даних під час навчання моделей, і пропонує огляд методів, які спрямовані на захист даних. Стаття є важливим внеском у розуміння безпеки та етики в машинному навчанні.

Огляд літератури показує, що машинне навчання є багатогранною дисципліною, яка включає різноманітні методи, підходи та галузі застосування. Від статистичного та логічного навчання до спеціалізованих практичних застосувань, таких як медична діагностика або проєктування антен, машинне навчання демонструє значний потенціал для вирішення складних задач. При цьому важливим є не лише розвиток нових алгоритмів, але й розуміння таких аспектів, як конфіденційність даних, етика та інтеграція машинного навчання у виробничі середовища. Ці публікації допомагають розширити розуміння ключових концепцій машинного навчання та сприяють розвитку цієї

динамічної області, відкриваючи нові перспективи для подальших досліджень та інновацій.

Список використаних джерел:

1. R. Muhamedyev. Machine Learning Methods: An Overview. 2015. https://www.academia.edu/download/54752237/CMNT_2015_Machine_Learning_ns_87art02_ReviewPaper.pdf
2. A.F.A.H. Alnuaimi, T.H.K. Albaldawi. Concepts of Statistical Learning and Classification in Machine Learning: An Overview. 2024. https://www.bioconferences.org/articles/bioconf/pdf/2024/16/bioconf_iscku2024_00129.pdf
3. P. Flach. Logical Approaches to Machine Learning - An Overview. 1992. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=03b54d20a9d1b68f01c959be0d1266076803f54d>
4. S. Albahra, T. Gorbett, S. Robertson, G. D'Aleo. Artificial Intelligence and Machine Learning Overview in Pathology & Laboratory Medicine. 2023. <https://www.sciencedirect.com/science/article/pii/S0740257023000138>
5. H.M. El Misilmani, T. Naous. Machine Learning in Antenna Design: An Overview on Machine Learning Concept and Algorithms. 2019. https://hilalelmisilmani.wordpress.com/wp-content/uploads/2019/10/hpcs_acme_2019_paper_4.pdf
6. D. Kreuzberger, N. Köhl, S. Hirschl. Machine Learning Operations (MLOps): Overview, Definition, and Architecture. 2023. <https://ieeexplore.ieee.org/iel7/6287639/6514899/10081336.pdf>
7. J.G. Carbonell, R.S. Michalski, T.M. Mitchell. An Overview of Machine Learning. 1983. <https://www.sciencedirect.com/science/article/pii/B9780080510545500054>
8. Popescu, S. Polat-Erdeniz, A. Felfernig, M. Uta. An Overview of Machine Learning Techniques in Constraint Solving. 2022. <https://link.springer.com/content/pdf/10.1007/s10844-021-00666-5.pdf>
9. T.O. Ayodele. Machine Learning Overview. 2010. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=72bfa0a326e4385ac46916e840e9fdc73a98b9fb>
10. E. De Cristofaro. An Overview of Privacy in Machine Learning. 2020. <https://arxiv.org/pdf/2005.08679>