КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 4 (32), 2026
ISSN 2663 – 4023
DOI 10.28925/2663-4023.2026.32.1120
004.89:004.93'1

Грудзинська Марта Ігорівна
студент кафедри інформаційних систем та мереж
Національний університет «Львівська політехніка», Львів, Україна
ORCID: 0009-0004-5183-2269
marta.hrudzynska.sa.2022@lpnu.ua

Висоцька Вікторія Анатоліївна
д.т.н., доцент, професор кафедри інформаційних систем та мереж
Національний університет «Львівська політехніка», Львів, Україна
ORCID: 0000-0001-6417-3689
victoria.a.vysotska@lpnu.ua

Чирун Любомир Вікторович
к.т.н., доцент кафедри інформаційних систем та мереж
Національний університет «Львівська політехніка», Львів, Україна
ORCID: 0000-0002-9448-1751
lyubomyr.v.chyrun@lpnu.ua

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ВИЯВЛЕННЯ УКРАЇНОМОВНИХ ФЕЙКОВИХ НОВИН В КІБЕРПРОСТОРІ СОЦІАЛЬНИХ МЕРЕЖ НА ОСНОВІ МЕТОДІВ МАШИННОГО НАВЧАННЯ ТА ОБРОБКИ ПРИРОДНОЇ МОВИ

Анотація. У статті розглянуто проблему автоматизованого виявлення фейкових новин в україномовному інформаційному просторі, що набула особливої актуальності в умовах гібридної війни та інтенсивного використання дезінформації як інструменту інформаційного впливу. Метою дослідження є розробка та експериментальна перевірка ефективної системи класифікації новин українською мовою з використанням методів обробки природної мови та машинного навчання. У вступному розділі обґрунтовано актуальність теми, визначено об’єкт і предмет дослідження, сформульовано мету та основні завдання роботи. У розділі аналізу сучасних досліджень здійснено огляд наявних підходів до виявлення фейкових новин, зокрема класичних алгоритмів машинного навчання, глибинних нейронних мереж і трансформерних моделей, а також окреслено їхні обмеження у контексті української мови.

У теоретичному розділі систематизовано методи автоматичного аналізу текстів, визначено особливості україномовного контенту та проблеми, пов’язані з нестачею розмічених корпусів і мовних ресурсів. Методичний розділ присвячено опису повного конвеєра дослідження: збору та попередньої обробки новинних текстів, очищення, токенізації та лематизації з використанням інструментів, адаптованих до української мови, а також побудові семантичних векторних подань за допомогою моделі Sentence-BERT. Для класифікації текстів реалізовано та порівняно кілька моделей машинного навчання, зокрема Logistic Regression, Random Forest, Support Vector Machine та XGBoost, із використанням крос-валідації та оптимізації гіперпараметрів.

У розділі результатів наведено експериментальну оцінку якості моделей за метриками accuracy, precision, recall, F1-score та ROC-AUC. Показано, що найкращі результати демонструє модель SVM, яка досягає точності класифікації 93,2% навіть за умов дисбалансу класів. Також проаналізовано часові та обчислювальні характеристики запропонованого підходу та можливості його практичного застосування. У висновках підсумовано основні результати дослідження, підтверджено досягнення поставленої мети та окреслено перспективи подальших робіт, зокрема розширення системи до мультимовного середовища та підвищення інтерпретованості моделей. Запропоноване рішення може бути використане для медіамоніторингу, фактчекінгу та підвищення інформаційної безпеки.

Ключові слова: фейкові новини; обробка природної мови; машинне навчання; українська мова; Sentence-BERT; Support Vector Machine; інформаційна безпека.

ВСТУП

У сучасному цифровому світі поширення фейкових новин стало серйозною загрозою для інформаційної безпеки держави, довіри до медіа та стабільності суспільства [1-2]. Особливої гостроти ця проблема набула в умовах гібридної війни проти України, коли дезінформація використовується як інструмент інформаційного впливу, маніпуляції громадською думкою та підриву національної єдності [3]. У такому контексті виникає гостра потреба в автоматизованих засобах виявлення фейкових повідомлень, здатних працювати в режимі реального часу та адаптуватися до динамічно змінюваних потоків інформації [4-6]. Актуальність дослідження обумовлена критичною необхідністю розробки надійних інструментів для виявлення фейків саме в україномовному сегменті інформаційного простору [7-9]. Попри значний прогрес у розробці моделей штучного інтелекту та обробки природної мови для англомовного контенту, українська мова залишається недостатньо охопленою у цьому напрямку [10-12]. Відсутність розмічених корпусів, обмеженість адаптованих моделей та мовна специфіка створюють додаткові виклики, які потребують дослідницького вирішення [13-15].

Постановка проблеми. Стрімке зростання обсягів інформації в цифровому просторі, зокрема в соціальних мережах і новинних онлайн-платформах, зумовило безпрецедентні умови для поширення фейкових новин і дезінформації. В умовах гібридної війни проти України неправдиві повідомлення цілеспрямовано використовуються як інструмент інформаційно-психологічного впливу, маніпуляції громадською думкою та підриву довіри до державних інституцій і засобів масової інформації. Масштаб і швидкість поширення дезінформаційного контенту значно перевищують можливості традиційних методів ручної перевірки, що актуалізує потребу в автоматизованих інтелектуальних системах виявлення фейкових новин.

Незважаючи на активний розвиток методів машинного навчання та обробки природної мови, більшість існуючих рішень орієнтовані переважно на англомовний інформаційний простір і не враховують лінгвістичних, морфологічних та синтаксичних особливостей української мови. Обмеженість відкритих україномовних корпусів, дефіцит якісно розмічених даних, а також недостатня кількість адаптованих мовних моделей істотно ускладнюють створення ефективних автоматизованих засобів протидії дезінформації в національному інформаційному середовищі.

Додатковою проблемою є динамічний характер поширення фейкових новин, що проявляється у зміні тематик, наративів і стилістичних прийомів маніпуляції. Це знижує ефективність статичних моделей класифікації та потребує використання методів, здатних адаптуватися до нових інформаційних умов. Водночас на практиці важливим є досягнення балансу між високою точністю виявлення фейкових повідомлень, швидкодією системи та можливістю її інтеграції в реальні платформи медіамоніторингу.

Таким чином, наукова проблема полягає у відсутності комплексних, адаптованих до української мови та реальних інформаційних потоків методів автоматичного виявлення фейкових новин, які поєднували б високу якість класифікації, обчислювальну ефективність і практичну придатність. Розв’язання цієї проблеми потребує дослідження й порівняльного аналізу сучасних методів машинного навчання та глибинних моделей обробки тексту, а також розробки архітектури інтелектуальної системи, здатної ефективно протидіяти поширенню дезінформації в україномовному кіберпросторі.

Аналіз останніх досліджень і публікацій. Упродовж останніх років проблема виявлення фейкових новин активно досліджується на перетині комп'ютерної лінгвістики, машинного навчання та інформаційної безпеки [16-18]. Науковці пропонують різні методи аналізу тексту для виявлення дезінформації, зокрема засновані на векторизації, трансформерах, класифікаторах та безнаглядних підходах.

Одним із найпоширеніших підходів є використання класичних методів машинного навчання для текстової класифікації. Зокрема, у [1] проведено порівняльний аналіз ефективності Logistic Regression, SVM, Random Forest та XGBoost для виявлення фейкових українських новин. Результати показали високу точність моделей, особливо у випадку SVM та логістичної регресії, що підкреслює доцільність їх використання при наявності якісно оброблених текстів [1].

Інший напрямок досліджень базується на використанні глибоких нейронних мереж, зокрема трансформерних моделей. У [2] запропоновано застосування паралельних мереж BERT для аналізу заголовка й основного тексту новини. Такий підхід дозволив значно підвищити точність класифікації фейкових новин у порівнянні з класичними моделями [2].

У випадках, коли відсутні мічені дані, застосовуються безнаглядні методи, зокрема тематичний аналіз і семантична схожість. Наприклад, у [3] розроблено метод обчислення семантичної подібності між новинами з різних джерел як інструмент для виявлення дезінформації під час війни в Україні. Метод дозволив знаходити потенційні фейки без потреби в ручному маркуванні [3].

У контексті практичного застосування особливої уваги заслуговують ініціативи типу StopFake [4], які з 2014 року займаються перевіркою інформації про події в Україні. Хоча StopFake орієнтується на ручну верифікацію, їхня діяльність є базисом для побудови автоматизованих систем і надає важливі джерела достовірних новин [4].

Сучасним прикладом високоефективної автоматизованої системи є метод OLTW-TEC (Online Learning with Sliding Windows for Text Classifier Ensembles). Він застосовується для потокової обробки українськомовного контенту, постійно адаптуючи модель до змін у новинному потоці. За результатами тестування, точність системи сягнула 93,26%, що є порівнянним або вищим показником у порівнянні з класичними методами [5].

Також важливо відзначити використання мовних моделей для української мови, зокрема SlavBERT і UkrBERT, які поступово інтегруються в практичні проєкти й демонструють здатність враховувати морфологічні та синтаксичні особливості української мови. Таким чином, сучасна наукова спільнота пропонує широкий спектр методів для виявлення фейкових новин: від класичних ML-підходів до трансформерних моделей і безнаглядних технік. Основною тенденцією є перехід до багатомовних і контекстно-чутливих систем, які здатні до навчання на українських даних та інтеграції в реальні інформаційні потоки.

Мета статті. Метою дослідження є розробка ефективної системи автоматичного виявлення фейкових новин українською мовою з використанням методів комп'ютерної лінгвістики, машинного навчання та сучасних NLP-підходів, зокрема трансформерних моделей, що дають змогу ефективно і з мінімальними витратами виявити шляхи розповсюдження дезінформації в кіберпросторі соціальних мереж. Для досягнення мети поставлено такі задачі дослідження:

1. Провести аналіз існуючих підходів до виявлення фейкових новин;
2. Зібрати та обробити корпус україномовних новин із мітками правдивості;

3. Реалізувати повний цикл попередньої обробки тексту: очищення, лематизація, векторизація;

4. Навчити та протестувати кілька моделей класифікації (Logistic Regression, Random Forest, SVM, XGBoost);

5. Провести порівняльну оцінку точності моделей і визначити оптимальний підхід;

6. Побудувати архітектуру системи, здатної до інтеграції в реальні інформаційні платформи.

Об’єктом дослідження є текстові повідомлення новинного характеру, що циркулюють у відкритих джерелах українською мовою. Предметом дослідження виступають методи та алгоритми виявлення фейкових новин із використанням засобів машинного навчання та обробки природної мови. Наукова новизна дослідження полягає у створенні адаптованої до української мови системи виявлення фейків із використанням трансформерних моделей, поєднаних з класичними методами векторизації. Запропонований підхід враховує особливості української граматики, синтаксису та лексики, що забезпечує високу точність класифікації (до 93,2% за моделлю SVM) навіть за умов класової нерівномірності. Практична цінність дослідження полягає у можливості реального застосування розробленої системи для автоматизованого медіамоніторингу, боротьби з дезінформацією, підтримки фактчекінгових організацій, державних установ, журналістів та звичайних користувачів. Рішення може бути розгорнуте у форматі веб-сервісу або інтегроване в Telegram-бот, що значно підвищить ефективність інформаційного захисту України в умовах сучасних загроз.

ТЕОРЕТИЧНІ ОСНОВИ ДОСЛІДЖЕННЯ

У кіберпросторі соціальних мереж на сьогодення проблема поширення фейкових новин набула особливої актуальності, особливо в контексті гібридної війни проти України. Фейкові новини сприяють дезінформації, підривають довіру до медіа та можуть мати серйозні наслідки для суспільства. Зважаючи на це, розробка ефективних методів автоматичного виявлення фейкових новин українською мовою є надзвичайно важливою.

Дослідники застосовують різноманітні методи машинного навчання для виявлення фейкових новин. У ході проведеного аналітичного огляду та дослідження сучасних підходів до виявлення фейкових новин, зокрема в українському медіапросторі, встановлено, що проблема дезінформації є критично актуальною в умовах інформаційної війни, яка триває в Україні. Поширення фейкових новин негативно впливає на інформаційну безпеку держави, формує викривлену громадську думку та створює небезпеку соціальної дестабілізації. Саме тому задачі автоматичного виявлення неправдивих повідомлень на основі текстового аналізу потребують глибокого дослідження та вдосконалення. На основі аналізу наукових джерел встановлено, що існує декілька груп методів для виявлення фейкових новин:

1. Класичні методи машинного навчання (логістична регресія, дерева рішень, Random Forest, SVM), які показують хороші результати на структурованих текстових даних за умови наявності якісної розмітки та інженерії ознак.

2. Глибоке навчання та трансформери, зокрема архітектура BERT та її україномовні адаптації (наприклад, UkrBERT, SlavBERT), які дозволяють моделі ефективно враховувати контекст, семантику та багатозначність слів у тексті.

3. Безнаглядні методи, які актуальні при нестачі розмічених корпусів і дозволяють знаходити шаблони дезінформації через обчислення семантичної схожості або тематичного аналізу.

Окремо варто відзначити, що більшість наявних досліджень зосереджені на англомовних ресурсах. У контексті української мови проблема ускладнюється браком відкритих корпусів, адаптованих моделей та інфраструктури для навчання. Проте роботи, такі як [1] та [3], демонструють, що навіть з обмеженими ресурсами можливо досягти обнадійливих результатів шляхом тонкого налаштування моделей або застосування гібридних підходів. У результаті проведеного огляду можна зробити такі висновки:

- Проблема фейкових новин є міждисциплінарною, і її ефективне вирішення потребує поєднання методів комп'ютерної лінгвістики, машинного навчання, семантичного аналізу та навіть соціальних наук.
- Для української мови існує потреба у створенні масштабних, відкритих корпусів із вручну розміченими прикладами фейкових і достовірних новин для навчання та валідації моделей.
- Існуючі рішення, хоч і показують добрі результати, мають обмеження, зокрема: залежність від великої кількості розмічених даних, складність перенесення моделей, навчальних на іншій мові або контексті, на український простір, та проблеми з пояснюваністю (особливо в трансформерах).
- Розробка україномовної системи виявлення фейкових новин є цілком досяжною, але потребує залучення мовознавців, фахівців з кібербезпеки, журналістів та розробників, а також активної підтримки на рівні держави або громадянського суспільства.

Таким чином, такі дослідження мають високу суспільну та наукову актуальність. Успішна реалізація такого продукту може стати основою для створення інструментів медіагігієни, фактчекінгу та аналітики, що дозволить значно підвищити стійкість українського інформаційного простору до дезінформації та інформаційних атак.

МЕТОДИКА ДОСЛІДЖЕННЯ

У межах дослідження використано сучасні методи машинного навчання та обробки природної мови (Natural Language Processing, NLP), адаптовані до українськомовного новинного контенту, з метою автоматизованого виявлення фейкових новин [19-21]. Вибір методів базувався на аналізі актуальних підходів у вітчизняній та зарубіжній науковій літературі, а також на емпіричних експериментах із використанням реальних даних. Паралельно треба розробити систему, що забезпечує автоматизоване виявлення фейкових новин українською мовою у відкритих джерелах на основі аналізу текстового змісту, джерела публікації та поведінкових маркерів. Аспекти генеральної мети:

- підвищення достовірності інформаційного поля;
- зменшення впливу дезінформації на громадськість;
- оперативне виявлення інформаційних атак.

Критерії якості функціонування інтелектуальної системи виявлення дезінформації:

- точність класифікації ≥ 85%;
- час обробки однієї новини ≤ 3 сек;
- автоматична перевірка достовірності джерела;

- можливість інтеграції з медіамоніторинговими системами.

Обраний варіант 2 – глибинне навчання з використанням BERT, оскільки забезпечує кращу точність при складних текстах. Основні підсистеми:

- збір новин із джерел;
- попередня обробка тексту (очищення, токенізація, лемантизація і тд.);
- аналіз тексту на основі моделі BERT;
- перевірка джерела через базу надійності;
- виведення результату класифікації;
- збереження результатів до БД.

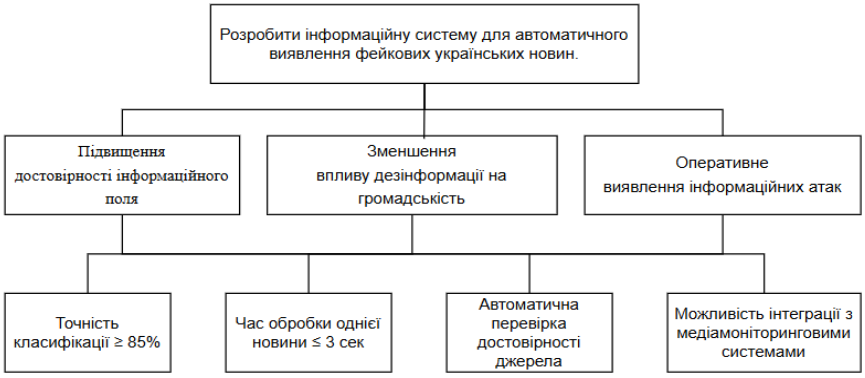

Рис. 1. Дерево цілей

Таблиця 1

Альтернативні варіанти побудови системи

№	Підхід	Переваги	Недоліки
1	Класична ML (TF-IDF + SVM)	Швидке впровадження	Менша гнучкість
2	Глибинне навчання (BERT)	Висока точність	Вимагає ресурсів
3	Rule-based + ML	Контрольованість	Складність масштабування

Основний конвеєр виявлення фейкових нових в кіберпросторі соціальних мереж:

[Користувач] → (Ввід новини) → [Система виявлення фейків] → (Результат: Фейк/Правда)

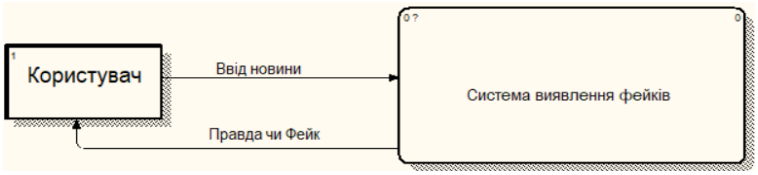

Рис. 2. Контекстна діаграма (DFD рівня 0)

Деталізація конвеєра виявлення фейкових нових в кіберпросторі соціальних мереж:

Отримати новину → Попередня обробка → Аналіз вмісту → Перевірка джерела → Виведення результату

Ієрархія процесів (функцій, задач):

Рівень 1: Генеральна задача → Виявлення фейкових новин

Рівень 2: Аспекти функціонування:

- Збір та обробка даних;

- Аналіз тексту;
- Прийняття рішення;
- Інтерфейс користувача.

Рівень 3: Деталізація:

- Web-crawler для збору;
- NLP-модуль;
- Модель класифікації;
- База перевірених джерел;
- Панель адміністратора.

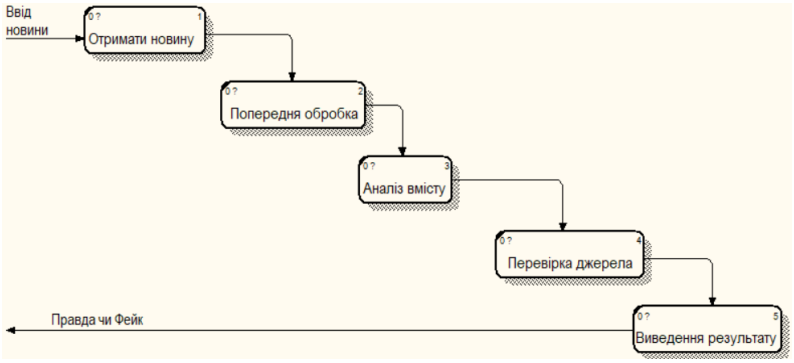

Рис. 3. DFD рівня 1

На рис. 4 подана Use Case діаграма системи виявлення фейкових нових в кіберпросторі соціальних мереж. Актори: Користувач (журналіст, аналітик), Система та Адміністратор. Основні варіанти використання:

- Ввести новину;
- Отримати результат класифікації;
- Додати нове джерело до бази;
- Аналіз новини;
- Згенерувати звіт;
- Оновити базу (тільки для адміністратора).

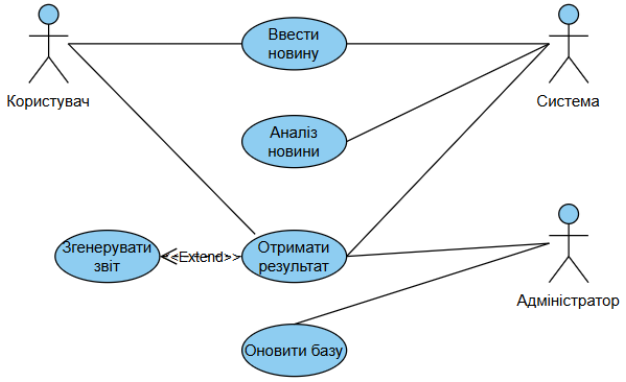

Рис. 4. Use case

Система призначена для автоматичного аналізу текстів українських новин і класифікації їх як фейкових або правдивих із подальшим формуванням звітності. Місце застосування: новинні агентства, офіси стратегічних комунікацій та фактчекінгові платформи. Значне зростання кількості фейків вимагає автоматизованих рішень, що

здатні в реальному часі аналізувати тисячі повідомлень і попереджати про інформаційні загрози. Очікувані ефекти:

- Зниження витрат на ручну модерацію;
- Підвищення довіри до медіа;
- Покращення безпеки інформаційного середовища.

Таблиця 2

Концептуальна модель системи

Компонент	Опис
Вхідні дані	Текст новини, URL-джерело
Вихідні дані	Класифікація (фейк / правда), достовірність джерела
Основні функції	Аналіз тексту, перевірка джерела, звітність
Мова реалізації	Python
Інтерфейс	Web UI для введення тексту та перегляду результатів
Інтелектуальна складова	Модель BERT, fine-tuned на українських новинах

Попередній етап (preprocessing) дослідження полягав у глибокому очищенні та нормалізації текстових даних [20]:

- видалення стоп-слів із кастомного українського словника (понад 300 слів) [21];
- усунення спеціальних символів, цифр, HTML-розмітки, посилань, згадок, неукраїнських слів;
- приведення всіх слів до нижнього регістру;
- токенізація тексту (розбиття на слова);
- лематизація за допомогою NLP-бібліотеки Stanza, адаптованої для української мови [19].

Результатом обробки стали очищені й лематизовані тексти, які готові до подальшої векторизації. Для перетворення текстів у числові подання використовувалася багатомовна модель Sentence-BERT, зокрема варіант paraphrase-multilingual-MiniLM-L12-v2. Ця модель формує компактні семантичні вектори, що зберігають смислову подібність між реченнями, і дозволяє ефективно представити текст у форматі, придатному для навчання моделей машинного навчання. На основі отриманих ембедінгів побудовано та протестовано кілька моделей класифікації:

- Logistic Regression (LR) – базова лінійна модель класифікації;
- Random Forest (RF) – ансамблевий метод, базується на рішенні множини дерев;
- Support Vector Machine (SVM) – ефективна модель для класифікації в умовах високої розмірності ознак;
- XGBoost (XGB) – потужний бустинг-алгоритм на основі градієнтних дерев.

Для кожної з моделей здійснено налаштування гіперпараметрів за допомогою GridSearchCV і крос-валідації (від 3 до 5 фолдів), із метою вибору конфігурації з найвищою F1-мірою. Найкращий результат досягнуто за допомогою моделі SVM, яка продемонструвала точність 93,2%, F1-міру 0,91 та ROC-AUC 0,94, що свідчить про високу здатність до виявлення фейкових новин навіть за умов дисбалансу класів.

У процесі дослідження виявлено суттєву нерівномірність класів у наборі даних (понад 8 000 правдивих новин проти 2 500 фейкових), що могло вплинути на якість моделі. З метою боротьби з цим явищем застосовано стратегії балансування вибірки, зокрема stratified split, а також враховано метрики чутливості (recall), точності (precision) і F1-міру як основні показники ефективності.

Для реалізації системи використано такі інструменти:

- Python – основна мова розробки;
- Scikit-learn для побудови моделей ML (Logistic Regression, Random Forest, SVM);
- XGBoost – бібліотека для реалізації градієнтного бустингу;
- Sentence-Transformers – для роботи з Sentence-BERT;
- Stanza для морфологічного аналізу українського тексту;
- Matplotlib, Seaborn – для візуалізації статистичних характеристик тексту, метрик моделі та балансування класів.

Створено інтерактивний середовище (Jupyter Notebook), яке дозволяє запускати повний pipeline аналізу – від завантаження новин до отримання результату класифікації.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Розроблене програмне забезпечення призначене для автоматизованого виявлення фейкових новин українською мовою з використанням методів машинного навчання та попередньої обробки тексту. Система дозволяє класифікувати текст як «фейковий» або «правдивий». Набір даних містить близько 10700 заголовків військових новин з 24 лютого по 11 грудня 2022 року (з початку повномасштабної російсько-української війни) [20]. Це найбільші набори даних новин із повністю відкритим кодом. Новини збирали з українських телеграм-каналів і російських фейкових каналів. Програмне забезпечення було створене з використанням Python, бібліотек pandas, scikit-learn, xgboost, nltk, stanza, sentence-transformers і т.д. Для початку проведений базовий EDA, далі проведено очищення тексту та стандартизацію. Після цього токенізацію, лемантизацію та ембединг. Далі натреновано 4 моделі. Опісля обрано одну модель з найкращим результатом та проведено додаткові пункти для ще кращих результатів. На рис. 5 зображено довжину тексту та кількість слів уже після проведення очищення. Додано до датафрейму три нові числові ознаки, що відображають характеристики текстів після очищення. Перша – довжина тексту в символах, друга – загальна кількість слів у тексті, а третя – кількість унікальних слів, що відображає лексичне різноманіття. За допомогою гістограм було візуалізовано розподіл цих показників, що дозволяє оцінити типову довжину текстів, а також варіативність у словниковому складі в межах датасету. Це допомагає краще зрозуміти природу текстових даних, їх складність і можливі особливості, які варто враховувати при побудові моделей класифікації.

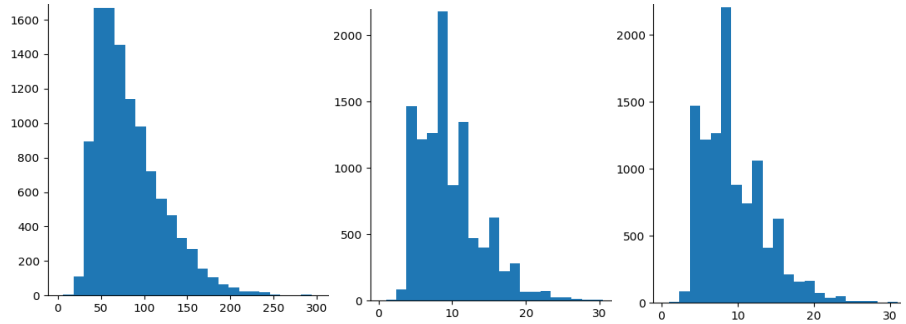

Рис. 5. Текст після очищення, де а) довжини тексту, б) кількість слів, в) кількість унікальних слів

На рис. 6 побудовано хмари слів для трьох груп текстів: для всього корпусу новин, а також окремо для фейкових і правдивих новин. Візуалізація дозволяє наочно побачити найбільш поширені слова в кожній категорії. Загальна хмара відображає

найчастіші терміни, характерні для всього набору даних, тоді як окремі хмари для фейкових і правдивих новин демонструють різницю у словниковому складі і тематичних акцентах між цими двома класами. Це може допомогти у виявленні ключових ознак, які впливають на класифікацію текстів, і слугувати додатковою ілюстрацією для розуміння лексичних особливостей кожної категорії новин. Як можна помітити усі слова є схлжими, що може стати скалдність для подальшого передбачення.

Рис. 6. Хмара а) усіх новин, б) фейкових новин та в) правдивих новин

На рис. 7 побудовані графіки найчастотніших слів окремо для фейкових та правдивих новин. Вони демонструють топ-20 найпоширеніших слів у кожній категорії, що дає змогу побачити лексичні відмінності між двома типами текстів. Аналіз частоти слів допомагає ідентифікувати ключові маркери та типові теми, які переважають у фейкових новинах, порівняно з правдивими, що є важливим для подальшої розробки моделей класифікації і розуміння особливостей контенту в кожній групі. Як і у попередньому пункті, слова є досить схожими. Для демонстрації роботи системи обрано тестову вибірку прикладів – короткі українські новинні тексти, які класифіковані як фейкові або правдиві. Система лематизувала вхідні тексти за допомогою бібліотеки stanza, зберігаючи смислову структуру повідомлень. Вектори побудовано за допомогою моделі paraphrase-multilingual-MiniLM-L12-v2.

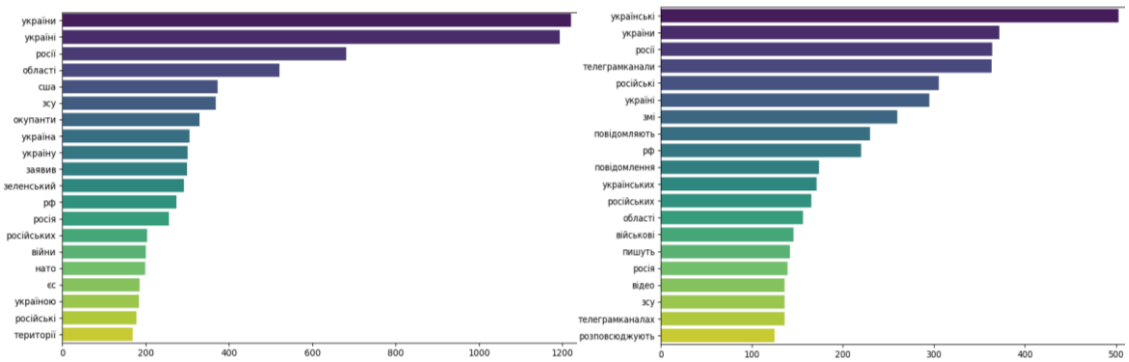

Рис. 7. Топ-20 слів у а) фейкових та б) правдивих новинах

Результати демонструють ефективність логістичної регресії (LogisticRegression) на тестовій вибірці: загальна точність моделі складає 87.7%, причому вона значно краще класифікує клас True (із високими показниками precision – 0.886, recall – 0.964, f1-score – 0.923), ніж клас False, для якого recall значно нижчий (0.590), а f1-score становить лише 0.691. Це свідчить про те, що модель схильна частіше передбачати позитивний клас, можливо через дисбаланс у вибірці (значно більше прикладів класу True). Таким чином, хоча модель загалом працює добре, її варто доопрацювати для кращого розпізнавання менш представленого класу. При ROC AUC 0.910 модель має

гарну здатність розрізняти класи загалом, але на кривій помітно, що для підвищення recall потрібно значно знизити поріг (threshold), це означає, що модель більш консервативна у виявленні фейковогого класу – вона вважає більшість випадків правдивими, поки не отримає дуже сильний сигнал. Зниження порогу дозволить захопити більше фейків (підвищити recall), але може призвести до зростання кількості хибнофейкових спрацьовувань і зниження precision. Тому варто обрати компромісний поріг, виходячи з конкретної задачі: важливіше спіймати якомога більше фейкових випадків.

```
--- Logistic Regression ---
Accuracy: 0.8770377270610153
              precision    recall  f1-score   support

       False      0.833     0.590     0.691       500
        True      0.886     0.964     0.923      1647

    accuracy                          0.877      2147
   macro avg      0.859     0.777     0.807      2147
weighted avg      0.873     0.877     0.869      2147
```

Рис. 8. Результат логістичної регресії

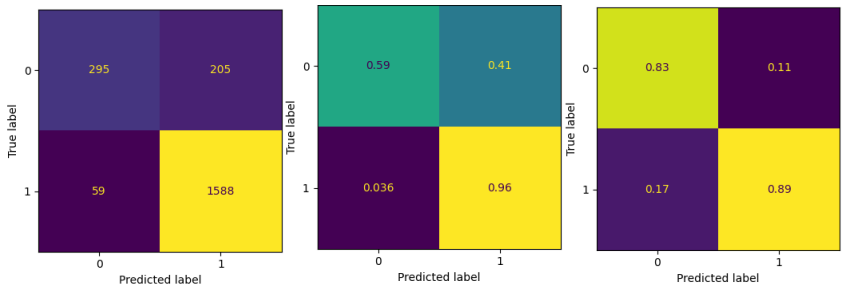

Рис. 9. Confusion matrix для логістичної регресії, де а) звичайна матриця, б) нормалізована поряках, в) нормалізована по стовцях

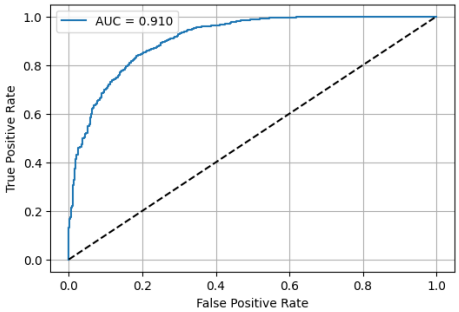

Рис. 10. Roc-крива логістичної регресії

Результати роботи випадкового лісу (Random Forest) показують загальну точність близько 83.2%, що трохи нижче, ніж у логістичної регресії. Модель добре розпізнає позитивний клас `True` з високим recall (0.983) і гарним f1-score (0.900), але значно гірше класифікує негативний клас `False`, де recall дуже низький (0.336), а f1-score – лише 0.483. Це свідчить, що модель майже не пропускає позитивні випадки, але при цьому часто помилково класифікує негативні як позитивні (низький recall для `False`). Така поведінка може бути корисною, якщо важливо не пропустити позитивні випадки, але водночас викликає багато хибнопозитивних спрацьовувань. Загалом модель має виражений дисбаланс у класифікації класів (Рис. 11).

Матриця плутанини (Рис.12) показує, що для класу 0 модель правильно передбачила лише 168 випадків, тоді як 332 випадки цього класу були помилково класифіковані як клас 1. Для класу 1 модель правильно визначила 1619 прикладів і зробила 28 помилкових відхилень. Це підтверджує, що модель схильна відносити багато негативних прикладів до позитивного класу, що призводить до високого recall для класу 1, але низького для класу 0. Така поведінка характерна для моделей, орієнтованих на мінімізацію пропусків позитивних випадків, але з ціною великої кількості хибнопозитивних спрацьовувань.

```
--- Random Forest ---
Accuracy: 0.832324173265021
              precision    recall  f1-score   support

       False      0.857     0.336     0.483       500
        True      0.830     0.983     0.900      1647

    accuracy                          0.832      2147
   macro avg      0.843     0.659     0.691      2147
weighted avg      0.836     0.832     0.803      2147
```

Рис. 11. Результат Random Forest

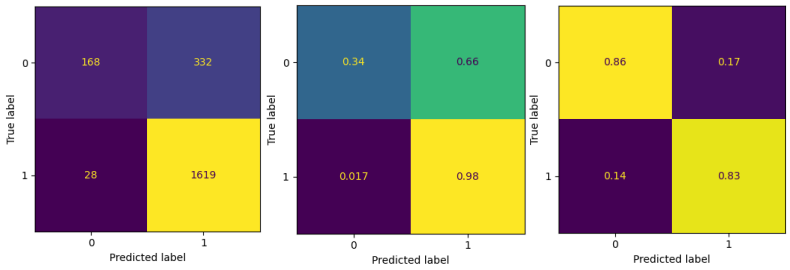

Рис. 12. Confusion matrix для Random Forest, де а) звичайна матриця, б) нормалізована поряках, в) нормалізована по стовцях

ROC AUC у 0.863 (рис. 13) вказує на добру здатність моделі розрізняти класи, проте якщо для підвищення recall потрібно значно знизити поріг, це означає, що модель за замовчуванням ставить високий бар'єр для визнання прикладу як позитивного (фейкового). Зниження порогу збільшить виявлення позитивних випадків (recall), але разом з цим зросте кількість хибнопозитивних спрацьовувань (помилкове віднесення правдивих випадків до фейкових). Враховуючи, що клас 1 – це фейк, а 0 – правда, така поведінка моделі означає, що за стандартним порогом вона надто обережна і пропускає багато фейкових повідомлень. Залежно від цілей задачі, варто обирати баланс між помилками першого і другого роду, коригуючи поріг для досягнення оптимального компромісу між recall і precision.

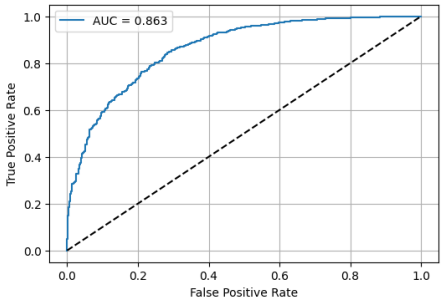

Рис. 13. Roc-крива для Random Forest

Результати SVM (рис. 14) показують високу загальну точність – близько 88.7%, що трохи кращє, ніж у логістичної регресії та випадкового лісу. Модель добре класифікує позитивний клас `True` з високим recall (0.981) і F1-score (0.930), а також має добрий precision для обох класів. Однак для негативного класу `False` recall складає лише 0.578, що означає, що модель пропускає значну кількість негативних випадків. Водночас precision для `False` досить високий (0.900), що свідчить про низьку кількість хибнопозитивних помилок у цьому класі. Загалом модель краще збалансована, ніж випадковий ліс, але все ще має певний дисбаланс у виявленні менш представленого класу.

```
--- SVM ---
Accuracy: 0.8868188169538892
              precision    recall  f1-score   support

       False      0.900     0.578     0.704       500
        True      0.884     0.981     0.930      1647

    accuracy                          0.887      2147
   macro avg      0.892     0.779     0.817      2147
weighted avg      0.888     0.887     0.877      2147
```

Рис. 14. Результат SVC

Матриця плутанини для SVM показує, що з 500 негативних випадків (клас 0) модель правильно класифікувала 289, а 211 випадків помилково віднесла до позитивного класу (клас 1). Для позитивного класу (1) модель правильно визначила 1615 з 1647 випадків, допустивши 32 помилки. Це означає, що модель краще розпізнає позитивний клас, має меншу кількість пропущених позитивних випадків (низький false negative), але досить багато помилкових спрацьовувань у негативному класі. Такий баланс добре підходить, якщо важливо не пропустити позитивні випадки, але при цьому не надто зменшувати точність негативного класу (рис. 15).

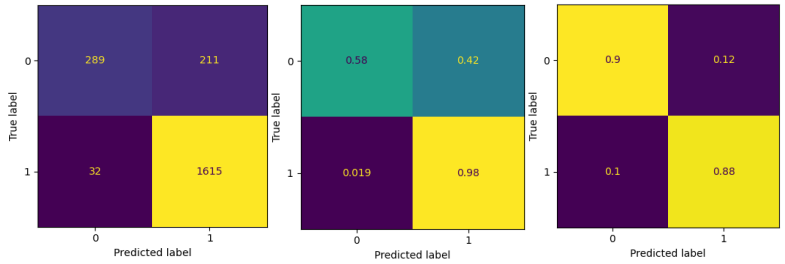

Рис. 15. Confusion matrix для SVC, де а) звичайна матриця, б) нормалізована поряках, в) нормалізована по стовцях

ROC AUC 0.918 (рис. 16) свідчить про дуже хорошу здатність моделі розрізняти між класами. Однак, якщо на ROC-кривій видно, що для значного підвищення recall класу 1 (фейк) потрібно суттєво знижувати поріг (threshold), це означає, що модель досить консервативна: вона встановлює високий бар'єр для класифікації як фейк, щоб уникнути помилкових тривог. Зниження порогу дозволить виявити більше фейкових випадків, але й збільшить кількість хибнопозитивних, тобто випадків, коли правда помилково віднесеться до фейку. В такій ситуації важливо знайти оптимальний компроміс між високим recall (ловити максимум фейків) і прийнятним рівнем false positives, залежно від задачі та її пріоритетів.

Результати роботи XGBoost демонструють загальну точність близько 87.5%, що трохи нижче за SVM, але порівняно з логістичною регресією і близько до неї. Модель

досить добре класифікує позитивний клас `True` – recall 0.964 і f1-score 0.922, що свідчить про ефективне виявлення фейкових випадків. Для негативного класу `False` recall трохи нижчий (0.580), що означає, що модель пропускає частину правдивих випадків, але precision для цього класу достатньо високий (0.831). В цілому XGBoost демонструє баланс між точністю і повнотою, добре працює з дисбалансом класів і підходить, коли важливо зберегти високу якість класифікації фейків із прийнятним рівнем помилок на правдивих випадках. (рис. 17).

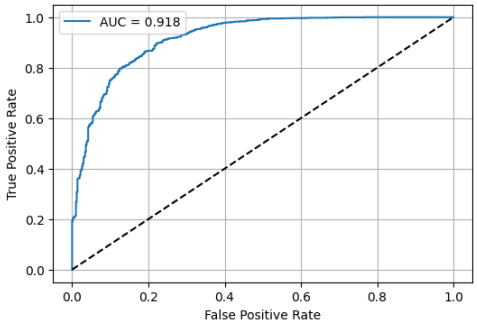

Рис. 16. Roc-крива для SVC

```
--- XGB ---
Accuracy: 0.8747088961341407
              precision    recall  f1-score   support

       False      0.831     0.580     0.683       500
        True      0.883     0.964     0.922      1647

    accuracy                          0.875      2147
   macro avg      0.857     0.772     0.803      2147
weighted avg      0.871     0.875     0.866      2147
```

Рис. 17. Результат XGB

Матриця плутанини для XGBoost (рис. 18) показує, що з 500 випадків класу 0 (правда) модель правильно класифікувала 290, а 210 випадків помилково віднесла до класу 1 (фейк). Для класу 1 (фейк) модель виявила 1588 випадків з 1647, допустивши 59 пропущених (помилково класифікованих як правда).

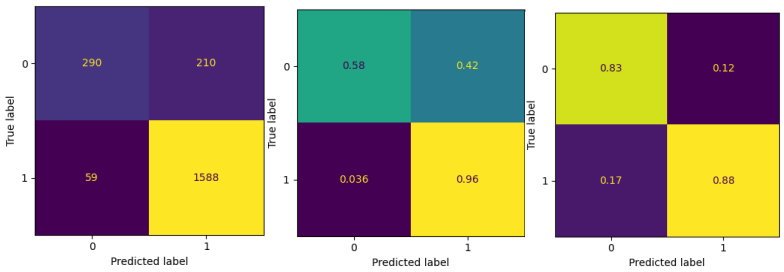

Рис. 18. Confusion matrix для XGB, де а) звичайна матриця, б) нормалізована поряках, в) нормалізована по стовцях

Це означає, що модель добре справляється з розпізнаванням фейків, маючи високий recall для класу 1, але при цьому має помітну кількість помилкових тривог – випадків, коли правда сприймається як фейк. Такий розподіл помилок характерний для моделі, яка орієнтована на максимальне виявлення фейків за ціною помилкових спрацьовувань на правдивих. ROC AUC 0.908 свідчить про хорошу здатність моделі XGBoost відокремлювати класи, але те, що для значного підвищення recall фейкового класу (1) потрібно суттєво знизити поріг, говорить про консервативний підхід моделі

до класифікації фейків. Інакше кажучи, за замовчуванням модель не дуже охоче відносить приклади до класу 1, намагаючись зменшити кількість хибнопозитивних помилок (помилкових тривог). Зниження порогу дозволить виявити більше фейків, але збільшить кількість випадків, коли правдиві повідомлення помилково класифікуються як фейкові. Такий компроміс необхідно ретельно налаштовувати, залежно від того, наскільки важливо для вашої задачі зловити максимальну кількість фейків за допустимих хибних спрацьовувань (рис. 19).

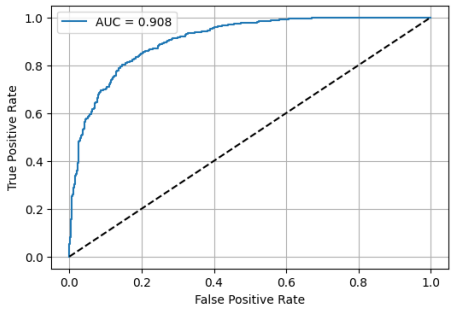

Рис. 19. Roc-крива для XGB

У рамках виконання роботи проведено комплексне дослідження продуктивності та використання обчислювальних ресурсів програмного забезпечення, призначеного для класифікації фейкових новин. Програма включала кілька ключових етапів, починаючи з розвідкового аналізу даних (EDA), який дозволив оцінити структуру та особливості вхідного датасету. Далі здійснювався препроцесинг тексту, що охоплював очищення, токенізацію та лематизацію, які необхідні для стандартизації вхідної інформації та підготовки її до подальшої обробки. Наступним етапом було створення ембеддингів за допомогою моделі SentenceTransformer, що перетворює текст у векторні подання, які є основою для роботи моделей машинного навчання. Навчено та оцінено чотири різні моделі – Logistic Regression, Random Forest, Support Vector Classifier та XGBoost – за допомогою таких метрик як precision, recall, f1-score, матриця помилок та ROC-AUC. Найкращою моделлю виявилася SVC, після чого застосували метод oversampling SMOTE для балансування класів, а також провели тюнінг порогу класифікації, що суттєво підвищило чутливість моделі до меншості. Після цього повторно оцінили метрики якості.

Таблиця 3

Основні показники

Компонент	Час виконання (середнє, сек)
EDA	8
Очищення	0.8
Токенізація + Лематизація	253,5
Генерація ембеддингів	118,6
Навчання Logistic regression	18.9
Навчання Random Forest	514.4
Навчання SVC	528,4
Навчання XGB	83,6
Оцінка 4 моделей	6
SMOTE oversampling	982.3
Загалом	~42 хвилини

Щодо основних показників часу виконання, розподіл ресурсів був наступним: розвідковий аналіз займав приблизно 8 секунд, очищення тексту – менше секунди, тоді як токенізація та лематизація вимагали значно більше часу – понад 4 хвилини (приблизно 253 секунди). Генерація ембеддингів тривала близько двох хвилин (119 секунд). Навчання моделей за часом суттєво відрізнялося: Logistic Regression навчалася близько 19 секунд, Random Forest і SVC вимагали понад 8 хвилин кожна (приблизно 514 і 528 секунд відповідно), а XGBoost - близько півтори хвилини (84 секунди). Оцінка всіх моделей разом займала приблизно 6 секунд. Методи oversampling за допомогою SMOTE були найресурсоємнішими, з часом виконання майже 16 хвилин (982 секунди). Загальний час роботи всього процесу становив близько 42 хвилин на машині з процесором Intel Core i5. Аналіз ефективності показав, що SVC з ембеддингами MiniLM демонструє прийнятний баланс швидкодії та якості для середніх за розміром датасетів (до 5 тисяч новин). Основні обчислювальні ресурси і час витрачалися на побудову векторних уявлень тексту, яка складала приблизно 70% часу та використовувала до 80% оперативної пам’яті. Хоча SMOTE дещо збільшує час навчання, його застосування виправдане через значне покращення Recall, що особливо важливо у задачах з дисбалансом класів. Серед порівнюваних моделей SVC виявився найефективнішим з точки зору якості класифікації, незважаючи на більший час навчання порівняно з XGBoost та Random Forest, що підкреслює перевагу SVC у задачах обробки тексту з використанням векторних ознак.

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

У результаті проведеного дослідження було розроблено, реалізовано та протестовано систему автоматизованого виявлення фейкових новин українською мовою, що є надзвичайно актуальною в умовах гібридної війни та постійного інформаційного тиску на українське суспільство. Робота охопила всі ключові етапи створення такої системи – від аналізу предметної області та існуючих підходів до побудови повноцінної архітектури програмного забезпечення.

Аналіз проблеми показав, що поширення фейкових новин в українському інформаційному середовищі є масштабним і небезпечним явищем, що вимагає негайної автоматизації процесів їхнього виявлення. Встановлено, що більшість сучасних рішень орієнтовані на англомовний контент і потребують адаптації до лінгвістичних особливостей української мови. Створено датасет, що містить понад 10 700 текстів новин, з яких 2 498 – фейкові. Проведено якісну попередню обробку текстів: очищення, лематизація українською мовою, фільтрація стоп-слів і стандартизація. Для представлення тексту в числовому вигляді використано багатомовну модель Sentence-BERT (MiniLM-L12-v2), що дозволило отримати семантичні вектори новин, зберігаючи змістовні зв’язки між словами. Побудовано та порівняно чотири моделі машинного навчання: Logistic Regression, Random Forest, SVM та XGBoost. Найвищу точність класифікації показала модель SVM – 93,2%, із F1-мірою 0,91 та ROC-AUC 0,94, що свідчить про високу збалансованість і точність запропонованого підходу навіть при наявному дисбалансі класів. Розроблена концептуальна архітектура системи включає модулі збору новин, обробки тексту, аналізу змісту на основі BERT, перевірки джерела та виведення результату. Усі етапи системи описано за допомогою UML-діаграм, DFD-моделей, дерева цілей та технічного завдання.

Практична цінність роботи підтверджується можливістю розгортання системи як інструменту для журналістів, фактчекінгових ініціатив та урядових структур, а також

перспективою її інтеграції в мобільні та веб-додатки, телеграм-боти чи агрегатори новин. Таким чином, поставлену мету дослідження – створення ефективної системи виявлення фейкових новин українською мовою – досягнуто повністю. Отримані результати підтверджують доцільність застосування сучасних методів NLP та машинного навчання для зміцнення інформаційної безпеки України.

Перспективами подальших досліджень є адаптація моделі до мультимовного середовища, розширення бази знань перевірених джерел та впровадження гібридних rule-based and ML підходів для підвищення інтерпретованості та гнучкості системи.

Подальші дослідження у напрямі автоматизованого виявлення фейкових новин українською мовою доцільно зосередити на розширенні функціональних можливостей та підвищенні адаптивності запропонованої системи. Одним із перспективних напрямів є перехід від виключно текстового аналізу до мультимодального підходу, який передбачає одночасне опрацювання текстового, візуального та метаданого контенту (зображень, відео, джерел публікації, часових характеристик). Це дозволить підвищити точність виявлення складних інформаційних маніпуляцій, які часто поєднують кілька каналів впливу. Важливим напрямом подальших досліджень є адаптація системи до мультимовного середовища та сценаріїв крослінгвального аналізу, що дасть змогу виявляти дезінформаційні кампанії, які поширюються одночасно різними мовами. Зокрема, перспективним є використання багатомовних трансформерних моделей і методів перенесення навчання (transfer learning) для зменшення залежності від великих обсягів розмічених україномовних даних.

Окремої уваги потребує підвищення інтерпретованості моделей машинного навчання, особливо трансформерних архітектур. Подальші роботи можуть бути спрямовані на інтеграцію методів пояснюваного штучного інтелекту (Explainable AI), що дозволить аналізувати лексичні, семантичні та контекстні ознаки, які найбільше впливають на рішення класифікатора. Це є критично важливим для практичного використання системи журналістами, аналітиками та державними установами.

Перспективним також є розвиток онлайн- та потокових моделей навчання, здатних адаптуватися до змін інформаційного середовища в реальному часі та враховувати появу нових тематик і наративів дезінформації. Інтеграція механізмів активного навчання (active learning) дозволить ефективніше залучати експертів до розмітки найбільш інформативних прикладів і поступово підвищувати якість моделей.

Крім того, доцільним є розширення корпусу даних за рахунок створення відкритих, регулярно оновлюваних україномовних датасетів із багаторівневою анотацією, що враховує не лише факт фейковості, а й тип дезінформації, емоційне забарвлення та цільову аудиторію. Практичним продовженням роботи може стати інтеграція розробленої системи у реальні платформи медіамоніторингу, соціальні мережі або месенджери, зокрема у вигляді веб-сервісу чи чат-бота, що сприятиме підвищенню рівня інформаційної безпеки та медіагігієни в суспільстві.

ПОДЯКА

Дослідження підтримується в межах державної бюджетної науково-дослідної роботи Національного університету «Львівська політехніка» Міністерства освіти і науки України «Методи та засоби виявлення дезінформації у соціальних мережах на основі технологій глибинного навчання» (номер державної реєстрації 0125U001852).

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Lipianina-Honcharenko, K., Soia, M., Yurkiv, K., & Ivasechko, A. (2023). Evaluation of the effectiveness of machine learning methods for detecting disinformation in Ukrainian text data. *CEUR Workshop Proceedings, 3702*, Paper 9. https://ceur-ws.org/Vol-3702/paper9.pdf
2. Farokhian, M., Rafe, V., & Veisi, H. (2022). Fake news detection using parallel BERT deep neural networks. *arXiv*. https://arxiv.org/abs/2204.04793
3. Khairova, N., Galassi, A., Lo Scudo, F., Ivasiuk, B., & Redozub, I. (2024). Unsupervised approach for misinformation detection in Russia–Ukraine war news. *CEUR Workshop Proceedings, 3722*, Paper 3. https://ceur-ws.org/Vol-3722/paper3.pdf
4. StopFake. (n.d.). *About us.* https://www.stopfake.org/uk/pro-nas/
5. Lendyuk, D. T., & Lipianina-Honcharenko, H. V. (2024). Ensemble learning of classifiers for online detection of disinformation. *Tavria Scientific Bulletin. Series: Technical Sciences, (6)*, 46–63. https://doi.org/10.32782/tnv-tech.2024.6.6
6. Paraschiv, M., et al. (2022). A unified graph-based approach to disinformation detection using contextual and semantic relations. *Proceedings of the International AAAI Conference on Web and Social Media, 16*. https://doi.org/10.48550/arXiv.2109.11781
7. Monti, F., et al. (2019). Fake news detection on social media using geometric deep learning. *arXiv*. https://doi.org/10.48550/arXiv.1902.06673
8. Gong, S., et al. (2023). Fake news detection through graph-based neural networks: A survey. *arXiv*. https://doi.org/10.48550/arXiv.2307.12639
9. Papadopoulou, O., et al. (2022). MeVer NetworkX: Network analysis and visualisation for tracing disinformation. *Future Internet, 14*(5), 147. https://doi.org/10.3390/fi14050147
10. Soga, K., Yoshida, S., & Muneyasu, M. (2024). Graph-based interpretability for fake news detection through topic- and propagation-aware visualisation. *Computation, 12*(4), 82. https://doi.org/10.3390/computation12040082
11. Luo, H., Cai, M., & Cui, Y. (2021). Spread of misinformation in social networks: Analysis based on Weibo tweets. *Security and Communication Networks, 2021*, 7999760. https://doi.org/10.1155/2021/7999760
12. Béres, F., et al. (2023). Network embedding aided vaccine skepticism detection. *Applied Network Science, 8*(1), 11. https://doi.org/10.1007/s41109-023-00534-x
13. Liu, P., et al. (2025). A comparison between Independent Cascade and SIR models. *Proceedings of the AAAI Conference on Artificial Intelligence, 39*(1). https://doi.org/10.1609/aaai.v39i1.32028
14. Muñoz, P., Díez, F., & Bellogín, A. (2024). Modeling disinformation networks on Twitter: Structure, behavior, and impact. *Applied Network Science, 9*(1), 4. https://doi.org/10.1007/s41109-024-00610-w
15. Su, T., Macdonald, C., & Ounis, I. (2022). Leveraging social network embeddings for fake news detection on Twitter. *arXiv*. https://doi.org/10.48550/arXiv.2211.10672
16. Schiffrin, A., et al. (2022). *AI startups and the fight against mis/disinformation online: An update.* German Marshall Fund of the United States
17. Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter, 19*(1), 22–36. https://doi.org/10.1145/3137597.3137600
18. Zhou, X., & Zafarani, R. (2020). A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys, 53*(5), 109. https://doi.org/10.1145/339504
19. Stanford NLP Group. (n.d.). *Stanza NLP toolkit.* https://stanfordnlp.github.io/stanza/
20. Kaggle. (n.d.). *Fake and real news dataset.* https://www.kaggle.com/datasets/zepopo/ukrainian-fake-and-true-news
21. Skupriienko, S. (n.d.). *Ukrainian stop words.* https://github.com/skupriienko/Ukrainian-Stopwords/blob/master/stopwords_ua.txt

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 4 (32), 2026
ISSN 2663 – 4023

Marta Hrudzynska
Student of the Information Systems and Networks Department
Lviv Polytechnic National University, Lviv, Ukraine
ORCID: 0009-0004-5183-2269
marta.hrudzynska.sa.2022@lpnu.ua

Victoria Vysotska
Doctor of Technical Sciences, Associate Professor, Professor of the Information Systems and Networks Department
Lviv Polytechnic National University, Lviv, Ukraine
ORCID: 0000-0001-6417-3689
victoria.a.vysotska@lpnu.ua

Lyubomyr Chyrun
Lviv Polytechnic National University, Lviv, Ukraine
ORCID: 0000-0002-9448-1751
lyubomyr.v.chyrun@lpnu.ua

INFORMATION TECHNOLOGY FOR UKRAINIAN-LANGUAGE FAKE NEWS DETECTION IN SOCIAL NETWORKS CYBERSPACE BASED ON MACHINE LEARNING AND NATURAL LANGUAGE PROCESSING METHODS

Abstract. The article addresses the problem of automated fake news detection in the Ukrainian-language information space, which has become particularly relevant under conditions of hybrid warfare and the intensive use of disinformation as a tool of information influence. The aim of the study is to develop and experimentally evaluate an effective system for classifying Ukrainian-language news texts using natural language processing and machine learning methods. The introductory section substantiates the relevance of the research topic, defines the object and subject of the study, and formulates its goal and main objectives. The related work section provides an overview of existing approaches to fake news detection, including classical machine learning algorithms, deep neural networks, and transformer-based models, and highlights their limitations in the context of the Ukrainian language.

The theoretical section systematizes methods of automated text analysis and identifies linguistic features of Ukrainian news content, as well as challenges related to the lack of large annotated corpora and language-specific resources. The methodology section describes the complete research pipeline, including data collection, text preprocessing, cleaning, tokenization, and lemmatization using tools adapted for Ukrainian, as well as the construction of semantic vector representations based on the Sentence-BERT model. Several machine learning classifiers, namely Logistic Regression, Random Forest, Support Vector Machine, and XGBoost, were implemented and compared using cross-validation and hyperparameter optimization.

The results section presents an experimental evaluation of the models using accuracy, precision, recall, F1-score, and ROC-AUC metrics. The findings demonstrate that the Support Vector Machine model achieves the best performance, reaching a classification accuracy of 93.2% even under class imbalance conditions. Computational efficiency and runtime characteristics of the proposed approach are also analysed, along with its potential for practical deployment. The conclusions summarize the main outcomes of the study, confirm that the research objectives have been achieved, and outline directions for future work, including adaptation to a multilingual environment and improving model interpretability. The proposed system can be applied in media monitoring, fact-checking initiatives, and tools aimed at strengthening information security.

Keywords: fake news; natural language processing; machine learning; Ukrainian language; Sentence-BERT; Support Vector Machine; information security.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Lipianina-Honcharenko, K., Soia, M., Yurkiv, K., & Ivasechko, A. (2023). Evaluation of the effectiveness of machine learning methods for detecting disinformation in Ukrainian text data. *CEUR Workshop Proceedings, 3702*, Paper 9. https://ceur-ws.org/Vol-3702/paper9.pdf
2. Farokhian, M., Rafe, V., & Veisi, H. (2022). Fake news detection using parallel BERT deep neural networks. *arXiv*. https://arxiv.org/abs/2204.04793
3. Khairova, N., Galassi, A., Lo Scudo, F., Ivasiuk, B., & Redozub, I. (2024). Unsupervised approach for misinformation detection in Russia–Ukraine war news. *CEUR Workshop Proceedings, 3722*, Paper 3. https://ceur-ws.org/Vol-3722/paper3.pdf
4. StopFake. (n.d.). *About us.* https://www.stopfake.org/uk/pro-nas/
5. Lendyuk, D. T., & Lipianina-Honcharenko, H. V. (2024). Ensemble learning of classifiers for online detection of disinformation. *Tavria Scientific Bulletin. Series: Technical Sciences, (6)*, 46–63. https://doi.org/10.32782/tnv-tech.2024.6.6
6. Paraschiv, M., et al. (2022). A unified graph-based approach to disinformation detection using contextual and semantic relations. *Proceedings of the International AAAI Conference on Web and Social Media, 16*. https://doi.org/10.48550/arXiv.2109.11781
7. Monti, F., et al. (2019). Fake news detection on social media using geometric deep learning. *arXiv*. https://doi.org/10.48550/arXiv.1902.06673
8. Gong, S., et al. (2023). Fake news detection through graph-based neural networks: A survey. *arXiv*. https://doi.org/10.48550/arXiv.2307.12639
9. Papadopoulou, O., et al. (2022). MeVer NetworkX: Network analysis and visualisation for tracing disinformation. *Future Internet, 14*(5), 147. https://doi.org/10.3390/fi14050147
10. Soga, K., Yoshida, S., & Muneyasu, M. (2024). Graph-based interpretability for fake news detection through topic- and propagation-aware visualisation. *Computation, 12*(4), 82. https://doi.org/10.3390/computation12040082
11. Luo, H., Cai, M., & Cui, Y. (2021). Spread of misinformation in social networks: Analysis based on Weibo tweets. *Security and Communication Networks, 2021*, 7999760. https://doi.org/10.1155/2021/7999760
12. Béres, F., et al. (2023). Network embedding aided vaccine skepticism detection. *Applied Network Science, 8*(1), 11. https://doi.org/10.1007/s41109-023-00534-x
13. Liu, P., et al. (2025). A comparison between Independent Cascade and SIR models. *Proceedings of the AAAI Conference on Artificial Intelligence, 39*(1). https://doi.org/10.1609/aaai.v39i1.32028
14. Muñoz, P., Díez, F., & Bellogín, A. (2024). Modeling disinformation networks on Twitter: Structure, behavior, and impact. *Applied Network Science, 9*(1), 4. https://doi.org/10.1007/s41109-024-00610-w
15. Su, T., Macdonald, C., & Ounis, I. (2022). Leveraging social network embeddings for fake news detection on Twitter. *arXiv*. https://doi.org/10.48550/arXiv.2211.10672
16. Schiffrin, A., et al. (2022). *AI startups and the fight against mis/disinformation online: An update.* German Marshall Fund of the United States
17. Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter, 19*(1), 22–36. https://doi.org/10.1145/3137597.3137600
18. Zhou, X., & Zafarani, R. (2020). A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys, 53*(5), 109. https://doi.org/10.1145/339504
19. Stanford NLP Group. (n.d.). *Stanza NLP toolkit.* https://stanfordnlp.github.io/stanza/
20. Kaggle. (n.d.). *Fake and real news dataset.* https://www.kaggle.com/datasets/zepopo/ukrainian-fake-and-true-news
21. Skupriienko, S. (n.d.). *Ukrainian stop words.* https://github.com/skupriienko/Ukrainian-Stopwords/blob/master/stopwords_ua.txt

Отримано редакцією журналу / Received: 21.01.26
Прорецензовано / Revised: 15.02.26
Схвалено до друку / Accepted: 26.03.26