Вінницький національний технічний університет

Факультет менеджменту та інформаційної безпеки

Кафедра менеджменту та безпеки інформаційних систем

МАГІСТЕРСЬКА КВАЛІФІКАЦІЙНА РОБОТА
на тему:

Удосконалення методу виявлення фейкової інформації у соцмережах засобами штучного інтелекту на основі мережі LSTM

Виконав: студент 2-го курсу, гр. 2КІТС-22м
спеціальності 125 – Кібербезпека
Освітня програма – Кібербезпека
інформаційних технологій та систем
Пугач В.С.
Керівник: к.т.н., проф., проф. каф. МБІС
Азарова А.О.
«04» грудня 2023 р.

Опонент: к.т.н., доц., доц. каф. ОТ
Савицька Л.А.
«04» грудня 2023 р.

Допущено до захисту
Голова секції УБ кафедри МБІС

Юрій ЯРЕМЧУК
“04” грудня 2023 р.

Вінниця ВНТУ – 2023 рік

Вінницький національний технічний університет

Факультет менеджменту та інформаційної безпеки

Кафедра менеджменту та безпеки інформаційних систем

Рівень вищої освіти ІІ-й (магістерський)

Галузь знань 12 – Інформаційні технології

Спеціальність 125 – Кібербезпека

Освітньо-професійна програма – Кібербезпека інформаційних технологій та систем

ЗАТВЕРДЖУЮ

Голова секції УБ, кафедра МБІС

_____**Юрій ЯРЕМЧУК**

“20” вересня 2023 р.

З А В Д А Н Н Я

на магістерську кваліфікаційну роботу студенту

Пугачу Володимиру Сергійовичу

1. Тема роботи «Удосконалення методу виявлення фейкової інформації у соцмережах засобами штучного інтелекту на основі мережі LSTM».
Керівник роботи к.т.н., проф., проф. каф. МБІС Азарова А.О., затверджені наказом вищого навчального закладу від “18” вересня 2023 року № 247
2. Строк подання студентом роботи за тиждень до захисту.
3. Вихідні дані до роботи: нормативно-правова база, монографії та актуальні наукові статті за темою роботи, Інтернет-ресурси, спеціалізована література, стандарти, існуюче ПЗ.
4. Зміст текстової частини: в першому розділі проаналізувати існуючі методи та засоби визначення фейкових новин; в другому розділі описати процеси розроблення інформаційної технології, формування набору даних для інтелектуального аналізу та визначити модель, що демонструє найефективніший

підхід для виявлення фейкових новин; в третьому розділі здійснити програмн
реалізацію розробки та аналіз результатів; в четвертому розділі проаналізувати
економічну ефективність розробленого програмного забезпечення.

5. Перелік ілюстративного матеріалу (з точним зазначенням обов`язкови
креслень): у першому розділі наведено 1 табл.; у другому розділі наведено 1
рис., 1 табл.; у третьому розділі наведено 20 рис.; у четвертому розділі наведен
8 табл.

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Основна частина	к.т.н., проф., проф. каф. МБІС Азарова А.О.		
Економічна частина	к.е.н., доц., доц. каф. ЕПВМ Причепа І.В.		

7. Дата видачі завдання 20 вересня 2023 р.

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів магістерської кваліфікаційної роботи	Строк виконання етапів роботи		Примітка
1.	Визначення напрямку МКР, формулювання теми	20.09.2023	25.09.2023	
2.	Аналіз предметної області обраної теми	26.09.2023	30.10.2023	
3.	Розроблення алгоритму роботи	03.10.2023	17.10.2023	
4.	Робота над МКР на основі вибраної теми	18.10.2023	10.11.2023	
5.	Робота над економічною частиною	11.11.2023	23.11.2023	
6.	Попередній захист МКР	24.11.2023	25.11.2023	
7.	Виправлення, уточнення, коригування роботи	26.11.2023	30.11.2023	
8.	Захист МКР	15.12.2023	15.12.2023	

Студент Пугач В.С

Керівник роботи Азарова А.О

АНОТАЦІЯ

УДК 004.9

Пугач В.С. Удосконалення методу виявлення фейкової інформації у соцмережах засобами штучного інтелекту на основі мережі LSTM. Магістерська кваліфікаційна робота зі спеціальності 125 – «Кібербезпека», освітня програма «Кібербезпека інформаційних технологій та систем». Вінниця: ВНТУ, 2023. 116 с.

На укр. мові. Бібліогр.: 39 назв; рис.: 35; табл. 10.

У першому розділі розглядається багатогранна природа фейкових новин, підкреслюється необхідність комплексного підходу, який поєднує людську розрізнювальну здатність з технологічними інноваціями. У ньому також оцінюється наявне програмне забезпечення та веб-інструменти, як українські, так і іноземні, які використовуються для виявлення фейків в інформаційному просторі.

Другий розділ присвячений розробці комплексних процесів для удосконалення методу виявлення фейкових новин. У ньому підкреслюється важливість створення надійного набору даних і ретельно оцінюються різні моделі машинного навчання для визначення найбільш ефективного підходу.

У третьому розділі детально описано практичну реалізацію удосконаленого методу виявлення фейкових новин. Обґрунтовано вибір Python як основної мови програмування та досліджено процес розробки програмного забезпечення. У цьому розділі також підкреслюється важливість зручного веб-інтерфейсу, реалізованого за допомогою Flask і HTML.

Заключний розділ передбачає поглиблене економічне оцінювання науково-технічної розробки. Він включає комерційний і технологічний аудит, прогнозування витрат, розрахунок економічної ефективності та терміну окупності інвестицій.

Ключові слова: фейк, інформація, штучний інтелект, соціальна мережа, метод, інструмент.

ABSTRACT

Puhach V.S. Improvement of the method of detecting fake information in social networks by means of artificial intelligence based on the LSTM network. Master's qualification thesis on specialty 125 - "Cybersecurity", educational program "Cybersecurity of information technologies and systems". Vinnytsia: VNTU, 2023. 116 p.

In Ukrainian language. Bibliographer: 39 titles; Fig.: 35; table 10.

The first chapter examines the multifaceted nature of fake news, emphasizing the need for an integrated approach that combines human discernment with technological innovation. It also evaluates existing software and web tools, both Ukrainian and foreign, that are used to detect fakes in the information space.

The second section is devoted to the development of complex processes for improving the method of detecting fake news. It emphasizes the importance of creating a robust dataset and carefully evaluates different machine learning models to determine the most effective approach.

The third chapter describes in detail the practical implementation of the improved method of detecting fake news. The choice of Python as the main programming language is justified and the software development process is investigated. This chapter also emphasizes the importance of a user-friendly web interface implemented using Flask and HTML.

The final section provides an in-depth economic evaluation of scientific and technical development. It includes commercial and technological audit, cost forecasting, calculation of economic efficiency and investment payback period.

Keywords: fake, information, artificial intelligence, social network, method, tool.

ЗМІСТ

ВСТУП

У сучасному цифровому середовищі швидке поширення інформації набуває як позитивних, так і негативних ознак. Пропонуючи безпрецедентний доступ до знань і новин, воно водночас відкриває шлюзи для потоку дезінформації та фейкових новин. Здатність відрізнити факти від вигадки є більш важливою, ніж будь-коли, оскільки неправдива інформація може посіяти паніку, розпалити суспільну істерію та підірвати довіру до надійних джерел.

Виявлення фейкових новин і боротьба з ними є надзвичайно важливою задачею особливо сьогодні в умовах повномасштабного вторгнення росії і здійснення нею активних атак на інформаційний простір з метою просування проросійських наративів, а також завдяки низці інших аспектів. Цифрова ера призвела до наявності величезних масивів інформації, що потребує на верифікацію, а отже, і на розроблення ефективних інструментів для виявлення та боротьби з фейковими новинами, які можуть мати вкрай негативні наслідки, впливаючи на громадську думку, псуючи репутацію та навіть підбурюючи окремі цільові аудиторії до насильства. Довіра до ЗМІ падає, що створює проблеми для прийняття обґрунтованих рішень і демократії.

Технологічні досягнення, такі як технологія deepfake і боти соціальних мереж, полегшили створення та поширення фейкових новин у глобальному масштабі. Це явище має значні наслідки для громадського здоров’я, як це було видно під час пандемії COVID-19. Боротьба з фейковими новинами вимагає балансу між захистом суспільства від шкоди та збереженням свободи слова.

Сьогодні напрацьовано значний теоретичний доробок у царині методології виявлення неправдивої інформації завдяки роботам таких провідних науковців, як Кіца М. Я., Снитюк Н., Бомчук Д. В., Корж О., Коровай, В., Тищенко В., Мужанова Т., Наконечний В., Барабаш О., Лаптєва Т., Міщенко А., Лукова-Чуйко Н., Лаптєва Т., Ошийко Я. Р., Шульська Н. М., Зінчук Р. С., Навальна М. І., Ковбасюк О. М., Грицюк Ю. І. [1–10].

Не зважаючи на наявний теоретичний апарат у галузі виявлення та блокування фейкової інформації, ця проблема залишається складною і

багатобічною та потребує подальших досліджень із метою поглиблення її вивчення, створення технологічних інновацій та глобальної співпраці для ефективного вирішення, що і зумовлює актуальність обраної в магістерській роботі теми щодо розроблення та вдосконалення автоматизованих інструментів для ефективного виявлення та боротьби з неправдивою інформацією, забезпечення цілісності інформації та захисту суспільства від шкідливих наслідків дезінформації.

Об'єктом дослідження є процеси визначення неправдивої інформації у текстовому новинному контенті.

Предметом дослідження є методи виявлення фейкових новин засобами штучного інтелекту.

Метою магістерської кваліфікаційної роботи є підвищення рівня захисту текстового новинного контенту від фейкових новин шляхом удосконалення методу здійснення інтелектуального автоматизованого аналізу на основі інструменту машинного навчання.

Для досягнення такої мети було поставлено і вирішено такі **завдання**:

1) проаналізовано сучасні форми фейкових новин та інтелектуальні методи їх виявлення та перевірки;

2) удосконалено метод аналізу фейків у текстовому новинному контенті на основі мережі LSTM;

3) створено повний і репрезентативний набір даних новинних статей для навчання та тестування побудованої нейронної мережі LSTM;

4) застосовано ефективну модель машинного навчання для удосконалення методу виявлення фейкових новин;

5) розроблено відповідний програмний засіб, що дозволяє комп'ютеризувати удоконалений на основі нейронної мережі LSTM метод виявлення фейкових новин у текстовому контенті;

6) здійснено тестування розробленого програмного засобу та доведено ефективність його використання для виявлення неправдивої інформації в текстах новин.

Методи дослідження. У роботі було використано комбінацію методів дослідження для всебічного вирішення поставлених завдань. Для реалізації збору даних використовувалися методи вилучення даних і веб-збирання, щоб зібрати інформацію з різноманітних онлайн-джерел, використовувалися методи пошуку на основі URL-адреси для початку процесу збору даних із статей новин. Для попередньої обробки даних використовувалися методи обробки природної мови, що дало можливість підвищити якість даних шляхом видалення нерелевантних слів і виділення коренів слів. Вилучення функцій для введення нейронної мережі: перетворення слова у вектор (NLP) – застосовувалися методи обробки природної мови для перетворення текстової інформації в числові вектори. Токенізація: реалізовані методи токенізації для розбиття текстового вмісту на окремі токени, що полегшує подальший аналіз. Для навчання нейронної мережі (поглиблене навчання) застосовувалися методи глибокого навчання, зокрема мережі довгострокової пам’яті (LSTM), що необхідно для навчання моделей, поглибленого аналізу та прогнозування достовірності новин. Для постійного вдосконалення мережі використовувалися ітераційні методи перенавчання для періодичного підвищення продуктивності нейронної мережі та адаптації до шаблонів, що розвиваються.

Наукова новизна. Основним науковим результатом магістерської роботи є удосконалення процесу виявлення фейкової інформації засобами системного підходу та нейро-мережевого моделювання, а саме: удосконалено метод виявлення фейкових новин на основі налаштованої нейронної мережі LSTM, що, на відміну від існуючих аналогів, дозволяє виявити тонкі мовні ознаки, шаблони і характеристики фейків; крім того, застосування методів глибокого навчання під час оброблення природної мови уможливлює підвищення точності та надійності процесу ідентифікації неправдивої інформації у новинному контенті.

Практична цінність. Практична цінність цієї роботи полягає в тому, що запропонований в роботі ПЗ дозволяє швидко та точно виявляти факти розповсюдження недостовірної інформації у новинах при цьому мінімізуючи трудові та грошові витрати.

Удосконалений автором магістерської роботи метод уможливлює просування медіаграмотності і розширює можливості як новинних організацій, так і громадськості у боротьбі з поширенням фейкових новин, його може бути використано для підвищення доброчесності платформ соціальних мереж і розвитку досліджень у галузі машинного навчання та обробляння природної мови.

Апробація. Результати дослідження було апробовано на Міжнародній конференції «Молодь в науці та освіті» (м. Вінниця, 2023 р.).

Публікації. За матеріалами досліджень, проведених у роботі, було опубліковано тези доповіді [25].

1. АНАЛІЗ ІСНУЮЧИХ МЕТОДІВ ТА ЗАСОБІВ ВИЗНАЧЕННЯ ФЕЙКОВИХ НОВИН

1.1 Дослідження проблеми та поняття “фейкова новина”

Вплив соціальних мереж на вибори викликає занепокоєння в останні роки. Ці платформи забезпечили простір для швидкого поширення інформації, як правдивої, так і неправдивої, і використовувалися різними суб’єктами для маніпулювання громадською думкою. Легкість, з якою фейкові новини та дезінформація можуть поширюватися в соціальних мережах, викликає сумніви щодо чесності демократичних процесів.

Паралельно традиційні медіа зіткнулися з проблемами адаптації до мінливого ландшафту реклами. Перенесення рекламних бюджетів на онлайн-платформи, такі як Google і Facebook, порушило традиційну модель доходів газет і мовлення. Інтернет-реклама пропонує кращу вимірюваність і налаштування, що робить її більш привабливою для рекламодавців. Цей перехід призвів до фінансового тиску на традиційні медіаорганізації, змусивши їх запекло конкурувати за читацьку аудиторію в Інтернеті та доходи від реклами.

У цьому конкурентному середовищі деякі традиційні засоби масової інформації вдаються до сенсацій і заголовків-приманок, щоб привернути увагу читачів. Гонитва за високими рейтингами та збільшенням веб-трафіку іноді призводить до компромісу в точності повідомлень новин. Крім того, спостерігається помітне збільшення негативних новин, оскільки людська психологія, як правило, більше реагує на негатив. Питання «Що найгіршого сталося сьогодні на Землі?», схоже, є рушійною силою новинних наративів, а провокаційна мова часто використовується для отримання кліків і залучення аудиторії [11].

Наслідки цієї тенденції значні. Хоча деякі новини технічно можуть бути точними, їх подання може бути дуже суб’єктивним, формуючи сприйняття громадськості таким чином, що може не відповідати дійсності. Редакційні

рішення, можуть вплинути на громадську думку та розуміння подій, потенційно являючи собою форму маніпуляції, яка часто залишається непоміченою, але має глибокий вплив.

У відповідь на ці виклики неможливо переоцінити важливість точної та об’єктивної звітності. Вкрай важливо не лише віддавати пріоритет фактичній точності, але й забезпечувати платформи для різноманітних точок зору. Вирішення проблеми маніпулювання засобами масової інформації та поширення фейкових новин має важливе значення для захисту цілісності демократичних процесів і публічного дискурсу. У наш час вибагливі читачі повинні активно шукати надійні джерела та різноманітні точки зору, щоб сформувати більш комплексне розуміння світу.

В епоху поширення цифрової інформації термін «фейкові новини» став суперечливим і багатогранним поняттям. Це стосується поширення неправдивої або оманливої інформації через різні канали ЗМІ, часто з наміром ввести в оману, дезінформувати або маніпулювати громадською думкою. Феномен фейкових новин привернув значну увагу вчених, журналістів, політиків і широкої громадськості через його потенціал підривати цілісність інформації, підривати довіру до ЗМІ та впливати на соціальні та політичні результати [12].

Походження та еволюція «фейкових новин».

Концепція фейкових новин не зовсім нова, але її масштаби та вплив посилилися в епоху цифрових технологій. Історично дезінформація та пропаганда використовувалися для просування певних завдань. Проте Інтернет і соціальні медіа сприяли швидкому поширенню неправдивих наративів, стираючи межі між законною журналістикою та сфабрикованим контентом. Ця еволюція підкреслює важливість розуміння фейкових новин у сучасному контексті.

Характеристики фейкових новин.

Фейкові новини демонструють кілька визначальних характеристик. У ньому часто бракує надійних джерел, він покладається на емоційну привабливість і використовує сенсаційність, щоб привернути увагу. Неправдиві

наративи можуть бути навмисно створені для імітації законних ЗМІ, що ускладнює для аудиторії відрізнити факти від вигадки. Крім того, фейкові новини можуть вірусно поширюватися через соціальні медіа, посилюючи їх охоплення та вплив [13].

Мотивація та актори.

Аналіз концепції фейкових новин передбачає заглиблення в мотиви їх створення та розповсюдження. Хоча деякі випадки фейкових новин можуть мати фінансові мотиви, інші служать політичним, ідеологічним чи особистим інтересам. У створенні фейкових новин беруть участь різні суб'єкти: від окремих творців контенту до організованих кампаній з дезінформації, організованих державними особами чи групами інтересів.

Вплив і наслідки.

Наслідки фейкових новин виходять далеко за межі дезінформації. Це може підірвати довіру до визнаних джерел новин, посіяти плутанину та поляризувати суспільства. Крім того, фейкові новини можуть впливати на громадську думку, вплинути на вибори та підбурювати події в реальному світі. Розуміння впливу на суспільство має вирішальне значення для розробки ефективних стратегій боротьби з фейковими новинами.

Проблеми у боротьбі з фейковими новинами.

Вирішення проблеми фейкових новин викликає численні проблеми. Збалансувати необхідність збереження свободи вираження думок і обов'язковість стримування неправдивої інформації є складним завданням. Крім того, децентралізований характер Інтернету та швидкість поширення дезінформації ускладнюють впровадження комплексних рішень. Технологічні досягнення, такі як технологія deepfake, ще більше ускладнюють спроби відрізнити реальність від брехні.

Роль медіаграмотності та перевірки фактів.

Підвищення медіаграмотності та просування ініціатив із перевірки фактів розглядаються як життєво важливі компоненти боротьби з фейковими новинами. Навчання людей критично оцінювати джерела інформації, перевіряти

твердження та розпізнавати упередженість має важливе значення в епоху цифрових технологій. Організації з перевірки фактів відіграють ключову роль у розвінчанні неправдивих наративів і притягненні до відповідальності розповсюджувачів фейкових новин [14].

Питання фейкових новин в Україні набуває особливого значення, враховуючи складний політичний ландшафт країни, історичний контекст і постійні виклики, пов'язані з дезінформацією. Україна була осередком геополітичної напруженості, а фейкові новини використовували як зброю, щоб впливати на суспільне сприйняття, загострювати розбіжності та розвивати політичні плани.

Розуміння ролі фейкових новин в українському контексті має вирішальне значення для політиків, журналістів та організацій громадянського суспільства, які працюють над вирішенням цієї проблеми. Він наголошує на необхідності комплексного підходу, який включає медіаграмотність, перевірку фактів і міжнародну співпрацю для боротьби з кампаніями з дезінформації та захисту цілісності інформації в Україні.

Підсумовуючи, концепція «фейкових новин» представляє собою динамічну та розвиваючу проблему в інформаційну епоху. Аналіз її походження, характеристик, мотивації, впливу та пов'язаних із цим викликів має важливе значення для розробки ефективних стратегій для вирішення цієї поширеної проблеми. Оскільки технології продовжують розвиватися, багатогранний підхід, який включає медіаграмотність, перевірку фактів і відповідальне цифрове громадянство, стає все більш вирішальним для пом'якшення шкідливого впливу фейкових новин.

1.2 Аналіз методів виявлення фейкових новин

Виявлення фейкових новин є багатогранним завданням, яке вимагає поєднання технологічних, журналістських та аналітичних підходів. В епоху, коли поширення неправдивої інформації може відбуватися швидко через

цифрові канали, розроблення ефективних методів виявлення фейкових новин є надзвичайно важливою. У цьому розділі розглядається аналіз методів, які використовуються для ідентифікації фейкових новин:

Організації з перевірки фактів.

Організації з перевірки фактів відіграють ключову роль у розвінчанні фейкових новин. У них працюють групи журналістів і дослідників, які ретельно перевіряють твердження, зроблені в новинних статтях і публікаціях у соціальних мережах. Ці організації, такі як Snopes, PolitiFact і FactCheck.org, перевіряють точність інформації, посилаються на надійні джерела та проводять всебічний аналіз для визначення правдивості новин.

Оброблення природної мови (NLP).

NLP, галузь штучного інтелекту, набула популярності в боротьбі з фейковими новинами. Алгоритми машинного навчання можна навчити аналізувати мову, яка використовується в новинних статтях і публікаціях у соціальних мережах. Моделі NLP можуть виявляти шаблони, неузгодженості та лінгвістичні ознаки, які вказують на оманливий вміст. Наприклад, надмірне використання емоційно напруженої мови або часті орфографічні помилки можуть викликати тривогу.

Перевірка джерела.

Оцінювання достовірності джерела новин є фундаментальним кроком у виявленні фейкових новин. Надійні ЗМІ дотримуються журналістських стандартів, посилаються на надійні джерела та мають історію точного звітування. Перехресне посилання на інформацію з авторитетними новинними організаціями є ефективним методом перевірки джерела [15].

Зворотний пошук зображень.

Фейкові новини часто містять підроблені або оманливі зображення. Інструменти зворотного пошуку зображень, такі як Google Images, TinEye і Google, дозволяють користувачам завантажувати або вводити зображення, щоб знаходити схожі або відповідні візуальні елементи в Інтернеті. Це допомагає виявити випадки, коли зображення було змінено або використано поза

контекстом.

Моніторинг соціальних мереж.

Фейкові новини часто поширюються через соціальні мережі. Моніторинг популярних тем і оцінювання достовірності облікових записів, які обмінюються інформацією, можуть допомогти виявити дезінформацію. Інструменти аналітики соціальних медіа можуть відстежувати швидкість і вірусність вмісту, допомагаючи ранньому виявленню потенційних фейкових новин.

Перевірка за допомогою краудсорсингу.

Краудсорсингові платформи, такі як «Snopes.com Hot 50» від Snopes і «r/Quityourbullshit» від Reddit, покладаються на внески користувачів для перевірки фактів і розвінчання фейкових новин. Ці спільноти волонтерів співпрацюють, щоб розслідувати та надавати доказову оцінку заяв про віруси.

Алгоритмічні підходи.

Платформи соціальних медіа та пошукові системи використовують алгоритми для виявлення та зменшення видимості фейкових новин. Ці алгоритми аналізують вміст на предмет моделей взаємодії, підозрілої поведінки та повідомлень від користувачів. Однак алгоритмічні підходи можуть мати обмеження, а також можуть викликати занепокоєння щодо цензури.

Семантичний аналіз.

Семантичний аналіз виходить за межі синтаксису та граматики і досліджує значення тексту. Він оцінює, як слова і фрази використовуються в контексті. Фейкові новини можуть використовувати оманливу семантику, щоб ввести в оману читачів. Алгоритми семантичного аналізу можуть виявляти семантичні аномалії та невідповідності, які свідчать про дезінформацію [16].

Боти для перевірки фактів.

Автоматизовані боти для перевірки фактів з’явилися як інструменти для швидкої оцінки точності новинних статей. Ці боти, інтегровані в платформи соціальних мереж, можуть здійснювати перевірку фактів у реальному часі, порівнюючи твердження в статтях з перевіреними джерелами. Хоча вони є цінними, їхня точність залежить від якості даних, на яких вони навчаються.

Системи звітності та позначення користувачів.

Багато платформ соціальних медіа та веб-сайтів новин включають системи звітності користувачів, які дозволяють людям позначати підозрілий або оманливий вміст. Хоча цей підхід покладається на пильність спільноти, він може допомогти виявити фейкові новини, використовуючи колективну мудрість онлайн-користувачів. Алгоритми часто визначають пріоритет вмісту з кількома звітами для перевірки.

Виявлення Deepfake.

Поява технології "Deepfake", яка використовує штучний інтелект для створення дуже переконливих фейкових аудіо- та відеоматеріалів, є унікальним викликом. Дослідники розробили інструменти для виявлення глибоких фейків, які аналізують візуальні та аудіосигнали, щоб виявити маніпуляції з медіа. Ці інструменти мають важливе значення для боротьби з використанням "Deepfake" з метою дезінформації [17].

Перехресні посилання на кілька джерел.

Основним принципом журналістики є перехресне посилання на інформацію з кількома надійними джерелами. Аналітики та спеціалісти з перевірки фактів використовують цей метод, щоб підтвердити заяви, зроблені в новинних статтях. Розбіжності або невідповідності між джерелами можуть вказувати на потенційні фейкові новини.

Програми медіаграмотності.

Навчання громадськості щодо критичного споживання медіа є проактивною стратегією. Програми медіаграмотності вчать людей критично оцінювати джерела новин, виявляти упередженість і відрізняти достовірну інформацію від неправдивої. Ці програми дають можливість людям бути більш розбірливими споживачами новин.

Міжнародна співпраця.

Фейкові новини є глобальною проблемою, і міжнародна співпраця між урядами, технологічними компаніями та організаціями громадянського суспільства є надзвичайно важливою. Обмін найкращими практиками,

інформацією та ресурсами може посилити колективні зусилля по боротьбі з фейковими новинами.

Важливо визнати, що жоден метод не є безпомилковим, і для комплексного виявлення фейкових новин часто необхідна комбінація цих підходів. Крім того, швидкий розвиток тактики дезінформації вимагає постійних досліджень і адаптації цих методів для ефективного вирішення проблем, які створюють фейкові новини в сучасному інформаційному просторі.

Оскільки ландшафт дезінформації продовжує розвиватися, дослідники та аналітики повинні залишатися гнучкими та інноваційними у своїх зусиллях по боротьбі з цією глобальною проблемою.

1.3 Вивчення недоліків та переваг існуючих підходів для виявлення фейкових новин

В епоху, що характеризується експоненціальним зростанням цифрової інформації, боротьба з фейковими новинами стала важливою для збереження цілісності онлайн-дискурсу. Повсюдний характер дезінформації, яка часто поширюється з тривожною швидкістю, зумовив необхідність розробки різних інструментів і платформ, покликаних боротися з цією нагальною проблемою. Цей аналіз заглиблюється в тонкощі декількох відомих програмних і веб-інструментів, призначених для виявлення та пом’якшення наслідків фейкових новин, проливаючи світло на їхні можливості та притаманні їм обмеження.

Поширення дезінформації породило зростаючу потребу в механізмах перевірки фактів, і одним з них є FactCheck.org. Цей сайт спеціалізується на ретельній оцінці достовірності політичних заяв, ретельному аналізі заяв, зроблених політиками, громадськими діячами та засобами масової інформації. Однак важливо розуміти ступінь його фокусу, який переважно зосереджений на політиці США, що потенційно обмежує його висвітлення в ширших інформаційних сферах [18].

Snopes, з іншого боку, закидає ширші тенета у своєму прагненні розвінчати міфи та перевірити факти. Його місія поширюється на розслідування міських легенд, інтернет-містифікацій та дезінформації в різних категоріях. Проте, як і багато фактчекерів, Snopes бореться з притаманною усім редакторам суб'єктивністю, яка може впливати на їхню здатність боротися з фейковими новинами, що швидко поширюються [18].

PolitiFact, відомий своїм "Truth-O-Meter", пропонує інструмент, за допомогою якого політичні заяви ретельно перевіряються і класифікуються на основі їхньої достовірності. Однак основна увага до політичного дискурсу може призвести до прогалин у висвітленні інших форм дезінформації [19].

NewsGuard представляє унікальний підхід, що функціонує як розширення для браузера та мобільний додаток, який присвоює рейтинги достовірності новинним сайтам. Для оцінки джерел працює команда аналітиків, які оцінюють джерела. Однак залежність від людських оцінок може перешкоджати оцінці достовірності окремих статей або контенту в режимі реального часу [19].

Botometer – інструмент, призначений для виявлення автоматизованих акаунтів або ботів на платформах соціальних мереж – використовує бальну оцінку для визначення ймовірності того, що акаунт є ботом. Однак він може іноді давати хибнопозитивні або хибнонегативні результати, коли має справу зі складними ботами, які імітують людську поведінку [20].

Ноаху виходить на арену з підходом, заснованим на візуалізації, що дозволяє користувачам відстежувати поширення тверджень і фактчеків на платформах соціальних мереж. Тим не менш, його залежність від загальнодоступних даних може обмежити його здатність збирати інформацію із закритих або приватних соціальних мереж [20].

Deepware Scanner занурюється у сферу маніпуляцій з медіа, виявляючи підроблені відео. Він проводить ретельний аналіз візуальних та аудіоелементів на наявність ознак фальсифікації, хоча його ефективність залежить від складності технології підробки, з якою він стикається [21].

Fact Check Explorer від Google пропонує практичне рішення, визначаючи

статті, які пройшли перевірку фактів авторитетними організаціями, і позначаючи їх міткою "Перевірка фактів" у результатах пошуку. Однак його корисність залежить від наявності статей, що пройшли перевірку фактів, і він може ненавмисно пропустити контент, який не був ретельно перевірений [21].

CrowdTangle з’являється як інструмент моніторингу соціальних мереж, що полегшує відстеження залучення та поширення контенту на таких платформах, як Facebook. Однак його спеціалізація на соціальних мережах може випускати з уваги фейкові новини, які переважно поширюються через альтернативні канали [22].

Інструмент ClaimReview запроваджує структуровану розмітку даних для позначення перевірених тверджень, що дозволяє пошуковим системам відображати перевірену інформацію. Тим не менш, її ефективність залежить від видавців, які використовують цю схему, і вона не здійснює незалежну перевірку тверджень.

У відповідь на зростаючу загрозу підробки зображень і відео Ініціатива з автентичності контенту (Content Authenticity Initiative, CAI) використовує криптографічні методи для перевірки автентичності медіа-контенту. Однак її ефективність залежить від широкого впровадження і може зіткнутися з проблемами, пов’язаними зі швидкою еволюцією технології глибоких підробок [22].

У глобальній боротьбі з фейковими новинами Україна не залишилася пасивним спостерігачем. У нашому регіоні з’явилося кілька програмних засобів і веб-інструментів, кожен зі своїм унікальним підходом до виявлення та усунення дезінформації. Цей аналіз досліджує деякі вагомі українські внески в боротьбу з фейковими новинами, проливаючи світло на їхні методології та сфери впливу.

StopFake.org:

StopFake.org – це відома українська платформа фактчекінгу, яка отримала міжнародне визнання завдяки зусиллям протистояти фейковим новинам, пов’язаним з Україною та регіоном загалом. Створений на хвилі українсько-

російського конфлікту, StopFake.org зосереджується на розвінчанні неправдивої інформації та дезінформаційних кампаній, пов’язаних із геополітичними подіями, зокрема тими, що стосуються України. На платформі працює команда перевіряючих фактів, які прискіпливо вивчають заяви та звіти, надаючи детальний аналіз їх точності [23].

Інструмент «Fake Hunter» від Detector Media:

Українська медіа-наглядова компанія «Детектор медіа» розробила інструмент «Мисливець за фейками» для боротьби з фейковими новинами та дезінформацією в соціальних мережах. Цей інструмент дозволяє користувачам повідомляти про ймовірні випадки фейкових новин або дезінформації. Згодом команда спеціалістів із перевірки фактів переглядає ці звіти та виносить вердикт щодо автентичності вмісту. Інструмент «Мисливець за фейками» від Detector Media дає змогу користувачам брати активну участь у боротьбі з фейковими новинами [23].

Проект Texty.org.ua "Факти":

Українська платформа журналістських розслідувань Texty.org.ua ініціювала проект «Факти» для перевірки фактів заяв українських політиків і громадських діячів. Цей проект фокусується на політичному дискурсі в Україні та спрямований на притягнення публічних діячів до відповідальності за достовірність їхніх заяв. Завдяки ретельному процесу перевірки фактів «Факти» сприяють прозорості та підзвітності української політики [24].

Платформи перевірки для соціальних мереж:

В Україні також з’явилися невеликі верифікаційні платформи, адаптовані до контенту соціальних мереж. Ці платформи використовують алгоритми та механізми звітування користувачів, щоб позначати потенційно фейкові новини в соціальних мережах. Хоча ці ініціативи можуть не мати такого глобального охоплення, як більші організації з перевірки фактів, вони відіграють життєво важливу роль у стримуванні поширення дезінформації в українському цифровому середовищі [24].

Недоліки існуючих підходів до виявлення фейкової інформації.

Незважаючи на їхні похвальні зусилля, українське програмне забезпечення та веб-інструменти для виявлення фейкових новин стикаються з певними проблемами та обмеженнями. До них належать:

Мовний бар'єр: багато українських фактчекінгових платформ переважно працюють українською мовою, що обмежує їх охоплення глобальною аудиторією.

Обмеження ресурсів: обмежене фінансування та ресурси можуть обмежити здатність українських організацій, що перевіряють факти, масштабувати свою діяльність і охоплювати ширший спектр тем.

Політична делікатність: перевірка фактів у політично напруженому середовищі може бути складною, оскільки це може призвести до звинувачень в упередженості чи прихильності.

Глобальний масштаб: хоча деякі українські ініціативи зосереджені на регіональних проблемах, вирішення глобальних проблем фейкових новин вимагає міжнародної співпраці та ресурсів.

Отже, Україна досягла значних успіхів у розробці програмного забезпечення та веб-інструментів для боротьби з фейковими новинами, зокрема в контексті українсько-російського конфлікту та внутрішньополітичного дискурсу. Хоча ці ініціативи заслуговують похвали, вони стикаються з проблемами, пов'язаними з мовою, ресурсами та політичною делікатністю. Тим не менш, вони сприяють ширшим глобальним зусиллям із забезпечення точної та надійної інформації в епоху, позначену швидким поширенням дезінформації.

Далі порівняємо іноземні ПЗ та веб-інструменти для виявлення фейкових новин із деякими відомими українськими ПЗ (табл. 1.1) [25].

Таблиця 1.1

Критеріальний аналіз вітчизняних та іноземних засобів виявлення фейкових новин

Інструмент/ Платформа	**Кількість виявлень**	**Коефіцієнт блокування (%)**	**Оцінка ефективності (1-10)**	**Зручність використання (1-5)**	**Оновлення у реальному часі (1-5)**	**Точність виявлення**	**Масштабованість**
FactCheck.org	2000	92	8	3	2	85	Високий
Snopes	2200	94	8	4	3	84	Високий

PolitiFact	2100	91	8	3	3	82	Високий
NewsGuard	2150	93	8	3	3	83	Високий
Botometer	1500	85	7	3	3	75	Помірний
Hoaxy	1600	87	7	3	3	78	Помірний
Deepware Scanner	1800	89	8	3	2	82	Помірний
Fact Check Explorer (Google)	1900	90	8	3	4	84	Високий
CrowdTangle	1750	88	7	3	3	81	Високий
ClaimReview Schema	1400	84	6	3	2	73	Помірний
Content Authenticity Initiative (CAI)	1650	86	7	3	3	80	Помірний
StopFake.org	2300	95	9	4	2	88	Високий
Fake Hunter	2200	94	8	4	2	84	Високий
Texty.org.ua	2100	91	8	4	2	85	Високий

Порівняння засобів виявлення фейкових новин, включаючи іноземні та українські рішення, дозволяє зробити кілька цікавих висновків. Ці інструменти відіграють важливу роль у боротьбі з поширенням неправдивої інформації та дезінформації, але їхня ефективність залежить від багатьох факторів.

1. Кількість виявлень і частота блокування:

Серед міжнародних інструментів StopFake.org, Fake Hunter і Snopes мають найвищі показники виявлення та блокування фейків. Ці інструменти ефективно виявляють і блокують фейкові новини.

Українські інструменти, такі як Texty.org.ua, також демонструють хороші результати за кількістю виявлень і показником блокування, що свідчить про їхню ефективність у місцевому контексті.

2. Показники ефективності та зручність інтерфейсу:

Коли мова йде про ефективність та зручність використання, міжнародні інструменти, такі як Snopes та FactCheck.org, виділяються серед інших. Вони забезпечують ефективне виявлення фейкових новин, зберігаючи при цьому зручний інтерфейс.

Українські інструменти, хоча й ефективні, але можуть потребувати вдосконалення з точки зору зручності та ефективності, про що свідчать дещо нижчі показники ефективності.

3. Оновлення в режимі реального часу та точність виявлення:

Міжнародні інструменти, такі як PolitiFact, Fact Check Explorer (Google) та NewsGuard, пропонують оновлення в режимі реального часу та високу точність виявлення. Вони вирізняються тим, що надають актуальну інформацію та точно ідентифікують фейкові новини.

Українські інструменти, такі як Texty.org.ua та StopFake.org, також надають оновлення в режимі реального часу і демонструють високий рівень точності у виявленні фейкових новин.

4. Масштабованість:

Масштабованість є вирішальним фактором для обробки величезних обсягів інформації в Інтернеті. Міжнародні інструменти, особливо ті, що інтегровані з великими платформами, такими як Google, мають високий рівень масштабованості.

Українські інструменти, хоча й ефективні в локальному контексті, можуть зіткнутися з проблемами, коли йдеться про обробку великого обсягу даних і масштабування їхньої роботи в глобальному масштабі.

5. Загальна оцінка:

Міжнародні інструменти, такі як Snopes, FactCheck.org та PolitiFact, пропонують комплексні рішення для виявлення фейкових новин з високою точністю та ефективністю. Вони добре зарекомендували себе і охоплюють широкий спектр тем.

Українські інструменти, зокрема Texty.org.ua та StopFake.org, демонструють високу ефективність у місцевому українському контексті, зокрема у виявленні та блокуванні фейкових новин, актуальних для регіону.

Отже, як міжнародні, так і українські інструменти виявлення фейкових новин відіграють важливу роль у боротьбі з дезінформацією. Міжнародні інструменти, як правило, мають ширшу сферу застосування та більшу масштабованість, що робить їх придатними для глобальної аудиторії. Українські інструменти, з іншого боку, необхідні для вирішення специфічних регіональних проблем і забезпечення достовірності інформації, що циркулює в Україні. Вибір між цими інструментами може залежати від конкретного контексту та цільової

аудиторії, а ідеальний підхід часто передбачає поєднання міжнародних і місцевих рішень для ефективної боротьби з фейковими новинами.

Заглиблюючись у всебічний аналіз цих програмних і веб-інструментів, важливо визнати, що, хоча вони є значним досягненням у боротьбі з фейковими новинами, вони не позбавлені недоліків. Людський нагляд, потенціал упередженості та постійно мінливий ландшафт дезінформації підкреслюють складність цього виклику. Отже, боротьба з фейковими новинами вимагає багатогранного підходу, який охоплює як технології, так і людську розрізнювальну здатність. Крім того, користувачі повинні застосовувати критичне мислення і не покладатися лише на ці інструменти для перевірки, оскільки боротьба з фейковими новинами залишається безперервним і динамічним процесом.

1.4 Формалізація вимог та постановка задачі

У контексті удосконалення методу виявлення фейкових новин важливо формалізувати вимоги та чітко окреслити постановку задачі. Це полягає у визначенні завдань, які необхідно вирішити, та формулювання етапів розроблення підходу.

Постановка проблеми. Поширення фейкових новин та дезінформації в цифрову епоху стало значною проблемою в усьому світі. Для вирішення цієї проблеми необхідне розроблення надійних інформаційних технологій, здатних ефективно виявляти та боротися з неправдивою інформацією в новинних статтях і повідомленнях.

Завдання, які необхідно вирішити:

1. Збирання та агрегація даних:

Збирання різноманітних даних про новинні статті з різних джерел, включаючи авторитетні та менш авторитетні ЗМІ.

Агрегування та впорядкування цих даних у структурований формат,

придатний для аналізу.

2. Розроблення функцій:

Визначення відповідних особливостей і характеристик, які можна використовувати для розрізнення справжніх новин від фейкових.

Вилучення та перетворення тексту та метаданих зі статей новин.

3. Моделі машинного навчання:

Розроблення моделей машинного навчання, здатних класифікувати новинні статті як справжні чи підроблені на основі вилучених ознак.

Точне налаштування та оптимізація цих моделей для підвищення точності.

Включення в систему можливостей перевірки фактів у реальному часі.

4. Масштабованість і продуктивність:

Забезпечення масштабованості технології для обробки великого обсягу новинних статей у режимі реального часу.

Оптимізація продуктивності для швидкого та ефективного виявлення.

Розроблення методу виявлення фейкових новин здійснюється за кількома етапами.

Етап 1. Збирання і підготовка даних для роботи методу:

- зібрати та об’єднати різноманітні дані про новинні статті;
- підготувати та очистити дані для аналізу.

Етап 2. Розроблення та вибір оціночних функцій для аналізованих новин:

- визначити відповідні ознаки та характеристики для класифікації новини;
- побудувати ці функції для аналізованого набору даних.

Етап 3. Застосування моделей машинного навчання для розроблення методу виявлення фейкових новин:

- відображення вхідних даних новини на числові вектори на основі моделей машинного навчання BERT, Word2Vec та Universal Sentence Encoder;
- використання моделі LSTM для оцінювання числових векторів аналізованих новинних даних для виявлення фейків у соцмережах;
- навчання нейронної мережі LSTM на певному наборі даних.

Етап 4. Масштабованість і оптимізація:

- переконатися, що удосконалений метод дозволяє обробляти велику кількість статей та новин;
- засобами постійного навчання нейронної мережі покращувати результативність запропонованого підходу в аспектах продуктивності та точності.

Етап 5. Тестування та валідація ПЗ, що реалізує удосконалений метод виявлення фейків:

- провести ретельне тестування та валідацію ПЗ, використовуючи різноманітні статті новин;
- Точне налаштування моделей і алгоритмів на основі результатів тестування.

Етап 6. Практична реалізація ПЗ, що реалізує удосконалений метод виявлення фейків:

- застосування розробленого ПЗ для перевірки достовірності новин у соцмережах;
- впровадження безперервного моніторингу соцмереж для адаптації розробленого ПЗ до нових тактик фейкових новин.

У результаті реалізації цих етапів у наступних розділах буде представлено удосконалений метод виявлення фейкових новин на основі нейро-мережевих технологій. Такий підхід дозволяє ефективно виявляти неправдиву інформацію в новинних статтях, що є цінним інструментом у боротьбі з дезінформацією та сприяє поширенню достовірних новин.

1.5 Висновок до розділу 1

Дослідження та аналіз, проведені в цьому розділі роботи, підкреслюють багатобічність проблеми фейкових новин і стратегії її вирішення. В епоху розповсюдження інформації, коли дезінформація може швидко поширюватися,

вкрай важливо прийняти цілісний підхід, який поєднує людську пильність і технологічний прогрес.

Поняття «фейкові новини» охоплює широкий спектр дезінформації, охоплюючи все від випадкових неточностей до навмисної неправди. Це тонке розуміння є основою для будь-яких ефективних контрзаходів. Для ефективної ідентифікації фейкових новин потрібне поєднання традиційного критичного мислення та сучасних технологій. Перевірка фактів, перевірка джерела та людське судження залишаються ключовими, але машинне навчання та інструменти оброблення природної мови можуть значно спростити такий процес.

Аналіз існуючих методів виявив їх обмеженість у достовірності і комплексності виявлення фейкових новин, що спричиняє потребу у застосуванні комбінації кількох підходів для покращення результативності їх роботи. Крім того, швидкий розвиток тактики дезінформації вимагає постійних досліджень і адаптації цих методів для ефективного вирішення проблем, які створюють фейкові новини в сучасному інформаційному просторі.

Комплексний аналіз іноземного та українського програмного забезпечення та веб-інструментів для виявлення фейкових новин дозволив розкрити відповідні ресурси для вирішення задач, поставлених у роботі. Серед недоліків існучих методів та інструментів для виявлення фейків слід відзначити такі:

1. Існуючі українські інструменти відстають за зручністю використання та показниками ефективності порівняно з міжнародними аналогами. Це говорить про можливості для вдосконалення покращення взаємодії з користувачем.

2. Українські інструменти, розроблені для конкретних регіональних завдань не можуть обробляти великі обсяги даних.

3. Існуючі інструменти мають проблеми з аналізом контексту та лінгвістичними тонкощами, що призводить до неточностей. Визначення «фейкових новин» може бути суб’єктивним, що призводить до потенційної неправильної класифікації.

4. ПЗ, що покладаються на звіти, створені користувачами, стикаються з

проблемами підтримування точності, оскільки якість вхідних даних низька.

5. У існуючих ПЗ відсутнє регулярне оновлення, що призводить до зниження ефективності виявлення та класифікації фейків.

6. Деякі інструменти можуть викликати проблеми з конфіденційністю, особливо ті, які аналізують поведінку користувачів або покладаються на великий збір даних.

Отже, в результаті поведеного комплексного аналізу стає очевидним, що існує нагальна потреба в розробленні удосконаленого методу, призначеного для виявлення фейкових новин, який поєднує найкращі моделі штучного інтелекту та машинного навчання і є потужним інструментом у боротьбі з дезінформацією.

2 РОЗРОБЛЕННЯ УДОСКОНАЛЕНОГО МЕТОДУ ВИЯВЛЕННЯ ФЕЙКОВИХ НОВИН ЗАСОБАМИ ШТУЧНОГО ІНТЕЛЕКТУ

2.1 Удосконалений метод виявлення фейкових новин за допомогою нейро-мережевих моделей

Ключову роль у реалізації методу ідентифікації фейкових новин відіграє застосування штучного інтелекту. Такий підхід охоплює низку складно пов'язаних процесів, спрямованих на ефективне вирішення проблеми виявлення фейкових новин.

Пропонується удосконалити метод виявлення фейкових новин за допомогою штучного інтелекту на основі LSTM, що полягає у виконанні таких етапів [25].

1. Збирання даних: це передбачає вилучення текстової інформації з різних веб-джерел, насамперед із соціальних мереж.

2. Попереднє оброблення: зібрані дані проходять етапи початкового попереднього оброблення для підвищення їх якості та придатності для аналізу.

3. Інтелектуальний аналіз і прогнозування: для поглибленого аналізу та прогнозування автентичності вхідної інформації використовується нейронна модель, яка базується на мережах довготривалої короткочасної пам'яті (LSTM).

4. Прийняття остаточного рішення: результати аналізу штучного інтелекту надаються особі, що приймає рішення, яка остаточно оцінює правдивість інформації та вносить рішення в базу даних.

5. Перенавчання мережі: безперервне вдосконалення досягається шляхом періодичного перенавчання та тонкого налаштування мережі.

Розглянемо блок-схему для реалізації запропонованого методу виявлення фейкових новин на основі LSTM, що проілюстровано на рис. 2.1 [25].

Кожен етап у запропонованому методі виконує свою функцію, для чого вводяться певні вхідні дані, що ретельно обробляються та згодом переходять на наступні етапи.

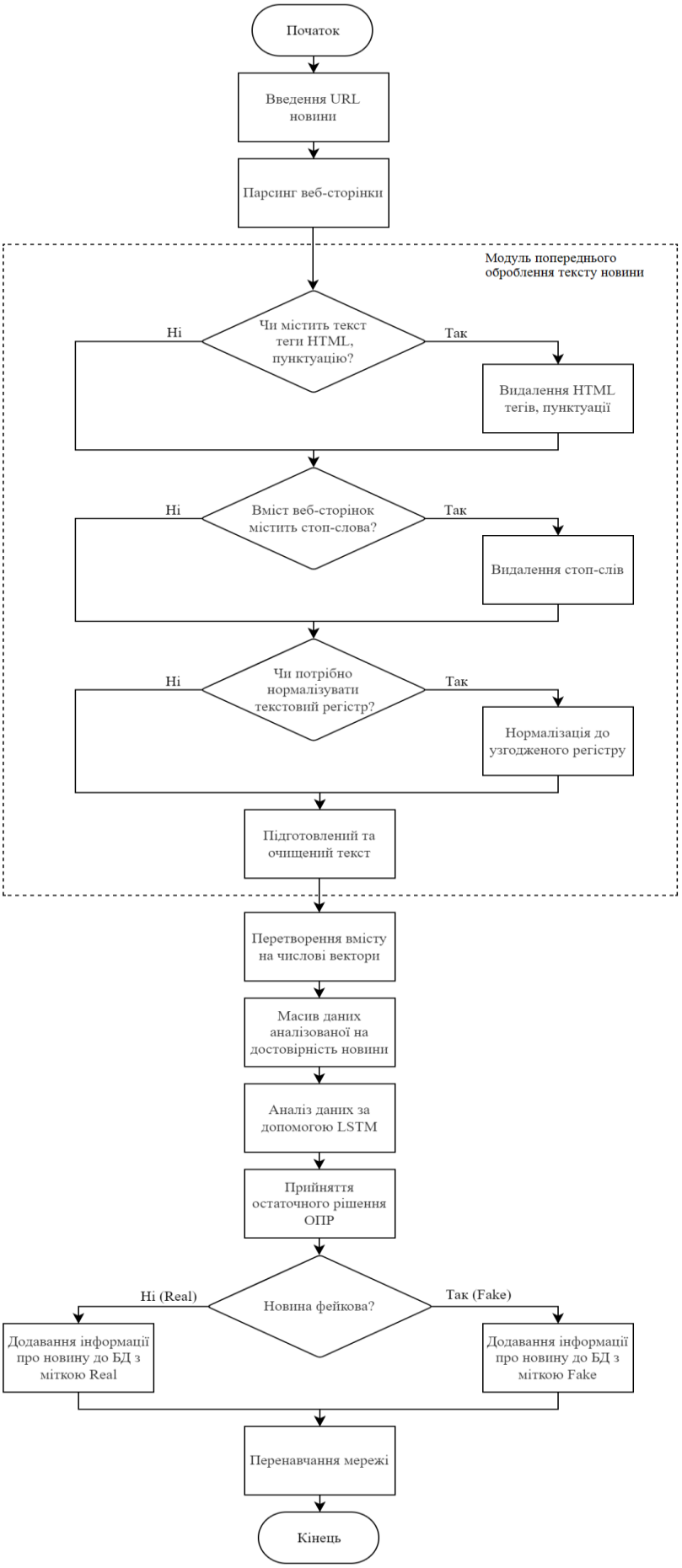


Рисунок 2.1 – Блок-схема удосконаленого методу виявлення фейкової інформації

Ініціювання процесу збору даних починається з надання URL-адреси новинної статті. Потім веб-сторінка аналізується, що дає змогу виокремити релевантну інформацію.

Враховуючи, що вміст веб-сторінок часто містить зайві елементи, такі як спеціальні символи та HTML-теги, необхідним стає етап попередньої обробки. На цьому етапі очищений текст піддається численним перетворенням, включаючи видалення стоп-слів, виділення основи слова, видалення розділових знаків і приведення тексту до єдиного регістру.

Етап інтелектуального аналізу передбачає застосування методів аналізу тексту з використанням моделей машинного навчання.

Процес збирання даних будемо здійснювати за двома основними етапами, як показано на рис. 2.2 [25].

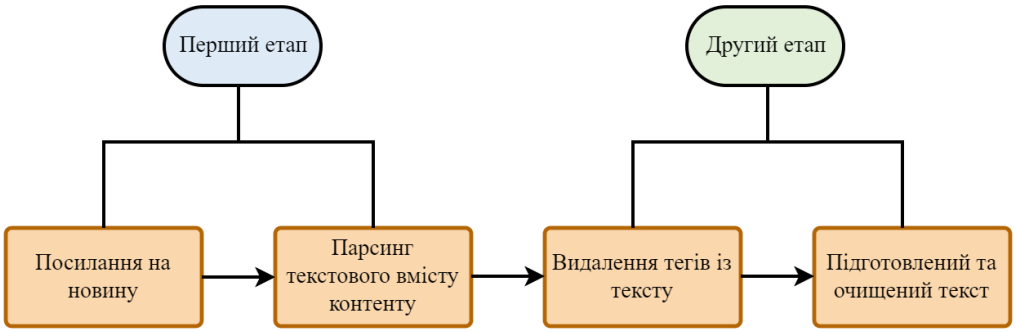


Рисунок 2.2 – Основні етапи процесу збирання даних згідно вдосконаленого методу

Початковий етап включає дві ключові операції: отримання веб-посилань на джерела інформації та отримання вмісту веб-сторінки. Далі відбувається вилучення текстової інформації з веб-сторінки та передача її на наступний етап обробки.

Необхідність очищення контенту веб-сторінки від сторонніх елементів зумовлена розумінням того, що необроблені дані можуть негативно вплинути на кінцеву точність аналізу достовірності інформації. Після вилучення вмісту веб-сторінки, включаючи HTML-код, дані переходять до етапу обробки. Тут вміст ретельно очищається від HTML-тегів, спеціальних символів та інших несуттєвих компонентів. Згодом текстовий вміст контенту проходить процедуру

попередньої обробки даних.

Попереднє оброблення даних охоплює такі дії, як видалення стоп-слів і виділення основи слова. Щоб полегшити подальшу обробку, отримані дані необхідно представити в цифровому вигляді. Отже, наступний етап передбачає перетворення текстової інформації у вектори, як показано на рис. 2.3 [25].

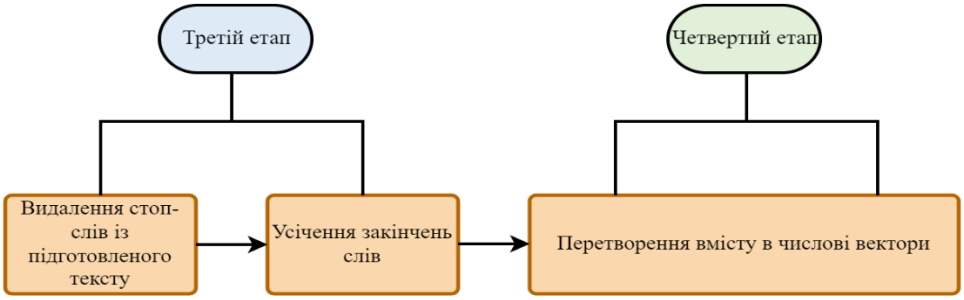


Рисунок 2.3 – Процес попереднього оброблення даних запропонованого методу

На наступному етапі фокус зміщується на первинне оброблення текстового контенту новинних статей, що вимагає виконання таких завдань [26]:

1. Видалення стоп-слів, до яких, як правило, належать усі літери алфавіту, займенники та сполучники, позбавлені значущості в документі.

2. Стеммінг: процес усічення закінчень слів зі збереженням лише кореня кожного слова.

Після цього виникає необхідність перетворити оброблений вміст документа в числові вектори. Нейронні мережі за своєю природою оперують з числовими даними, а не з текстовими документами. Перетворення слів у вектори є ключовим етапом процесу. Отже, всі подальші дії, пов’язані з аналізом новин, виконуються в числовому векторному форматі. Модуль перетворення слів у вектори приймає оброблений текст як вхідні дані і створює масив векторів. Ці вектори згодом зберігаються для подальшого використання. У модулі інтелектуального аналізу навчена мережа LSTM оцінює вміст вхідного вектору **X**, що представляє собою закодований масив даних аналізованої на достовірність новини. Дані, представлені у вигляді векторів слів, потрапляють до модуля нейромережі LSTM, яка генерує класифіковані дані. Класифіковані дані дозволяють LSTM оцінити достовірність новини, класифікуючи її як "Real"

(достовірна) або "Fake" (фейкова). Розглянемо структуру мережі LSTM (рис. 2.4) [25] для оцінки достовірності новин у соціальних мережах. Тут на вхід подаються **X** – вектор вхідних даних, а не виході отримуємо **R** – вихідний результат ("Real" або "Fake").

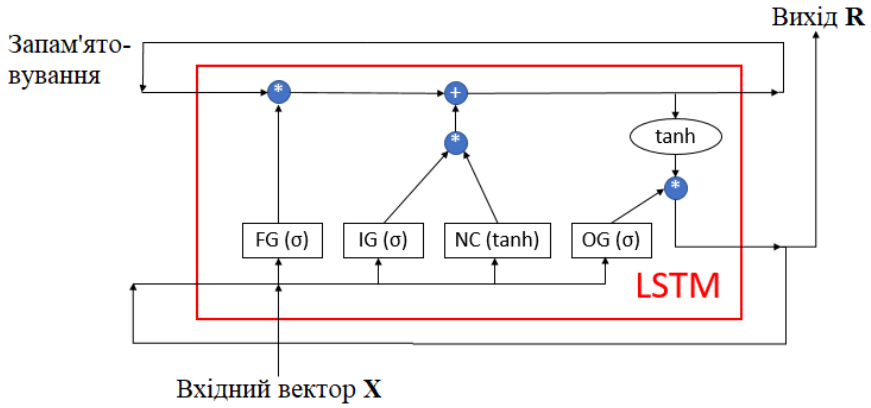


Рисунок 2.4 – Етап інтелектуального аналізу та прогнозування для удосконаленого методу

Ці класифіковані дані потім передаються на перевірку оператору, який приймає рішення. Оператор прийняття рішень проводить практичну оцінку точності прогнозу.

Розгянемо такий процес прийняття рішень більш детально. На цьому етапі до завдання виявлення фейкових новин активно долучається особа, яка приймає рішення (ОПР). Після отримання нейромережею прогнозного рішення особа, яка приймає остаточне рішення, здійснюючи особисте оцінювання, щоб переконатися в правильності класифікації. Згодом ОПР додає до бази даних відповідний ярлик, позначаючи новину як "СПРАВЖНЮ" або "ФЕЙКОВУ". Цей багатоетапний процес проілюстровано на рис. 2.5 [25].

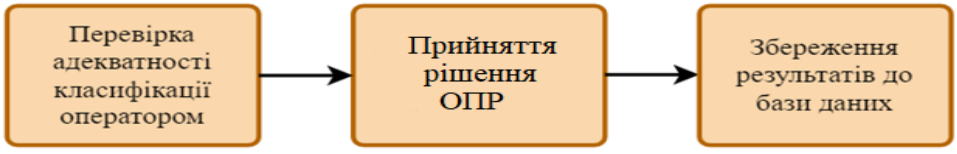


Рисунок 2.5 – Процедура прийняття рішення ОПР в удосконаленому методі виявлення фейкової інформації

Розглянемо процес прийняття рішень ОПР, що полягає у виконання трьох таких кроків [25]:

1. Перевірка надійності класифікації оператором прийняття рішень (ОПР): На цьому початковому етапі оператор ретельно перевіряє точність прогнозу нейронної мережі щодо достовірності новин. Згодом процес переходить до етапу прийняття рішень.

2. Прийняття рішення: На цьому кроці оператор прийняття рішень (ОПР) повинен вжити відповідних заходів. Ці дії можуть включати внесення коректив, якщо класифікація нейронної мережі була помилковою, або підтвердження правильності прогнозованої класифікації.

3. Збереження класифікованих даних: На цьому кроці інформація передається до бази даних, яка містить анотовані записи новин. Ці записи ретельно зберігаються для можливого майбутнього перенавчання нейронної мережі, якщо таке перенавчання буде визнано необхідним (рис. 2.6) [25].

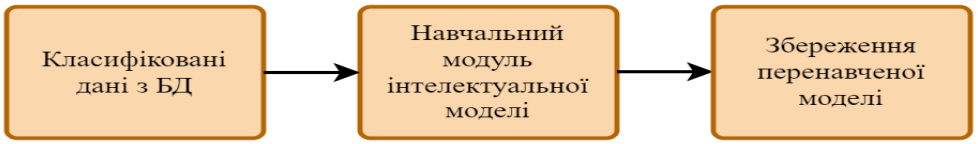


Рисунок 2.6 – Процедура перенавчання нейронної мережі в удосконаленому методі виявлення фейкової інформації

Такий комплексний підхід до прийняття рішень забезпечує підвищення точності ШІ-системи з часом, що сприяє більш надійному виявленню фейкових новин.

Процедура перенавчання нейронної мережі, що запускається змінами, внесеними ОПР до бази даних, відіграє вирішальну роль у розширенні набору навчальних даних. Це доповнення в кінцевому підсумку підвищує точність подальших оцінок.

Процес перенавчання складається з двох окремих кроків. Спочатку з бази даних, що містить анотовані новинні записи, витягуються класифіковані дані. Згодом ці дані передаються до додаткового навчального модуля інтелектуальної моделі, а саме до етапу збирання даних, зосередженого на новинах (рис. 2.2).

Таким чином, запропоновано удосконалений метод вивлення фейків, спроможний виділяти текстовий контент із веб-статей новин, проводити попереднє оброблення тексту, здійснювати інтелектуальний аналіз і прогнозування, надавати операторам можливість вносити виправлення і самостійно оцінювати достовірність новин. Слід зауважити, що такий підхід дозволяє постійне навчання мережі на основі інформації, отриманої від ОПР.

2.2 Етап збирання даних та підготовки до інтелектуального аналізу в методі, що розробляється

Для побудови удосконаленого методу виявлення фейкової інформації в соціальних мережах необхідно спочатку здійснити етап набору початкових даних для інтелектуального аналізу за допомогою штучного інтелекту на основі мережі LSTM, що є складним аспектом магістерської роботи. Вкрай важливо мати високоякісний і добре структурований набір даних для навчання та оцінювання ефективності моделі на основі LSTM у виявленні підробленої інформації.

Отже, далі надамо опис ключових етапів, пов’язаних із процесом збирання та підготовки даних.

Збирання даних. Першим кроком є збирання різноманітних і репрезентативних даних із соціальних мереж. Ці дані включають дописи, коментарі та створений користувачами вміст, сприйнятливий до дезінформації та фейкової інформації. Джерела даних можуть включати такі популярні соціальні мережі, як Twitter, Facebook, Reddit або інші відповідні мережі.

Вибір джерел даних. Вибір джерел даних є важливим рішенням. Мета полягає в тому, щоб виявити платформи та профілі, відомі поширенням неправдивої або оманливої інформації. Важливо зосередитися на різних джерелах, щоб забезпечити повноту набору даних.

Етичні міркування: збирання даних із соціальних мереж вимагає суворого

дотримання етичних принципів. Це включає отримання дозволу або дотримання умов угод про обслуговування платформ, повагу до конфіденційності користувачів і забезпечення відповідального використання даних [27].

Попереднє оброблення даних: зібрані дані часто містять шум, нерелевантний вміст і різноманітні текстові формати. Попереднє оброблення даних необхідна для очищення та форматування даних для подальшого аналізу. Цей етап включає нормалізацію тексту, коли дані перетворюються на узгоджений формат, включаючи нижній регістр, видалення спеціальних символів і вирішення таких проблем, як орфографічні помилки. Крім того, попередня оброблення даних може передбачати видалення нерелевантного вмісту, наприклад повторюваних публікацій або коментарів [28].

Анонімізація: забезпечення конфіденційності та анонімності користувачів є важливим етичним міркуванням. Особиста інформація та ідентифікатори повинні бути видалені або анонімні, щоб захистити особистість користувачів, чиї дані включені в набір даних.

Маркування: процес маркування даних включає класифікацію кожного фрагмента вмісту як справжню або підроблену інформацію. Анотатори або алгоритми маркування оцінюють достовірність інформації. Цей процес створює основну істину, за якою можна оцінити прогнози моделі.

Збалансування набору даних: незбалансовані набори даних можуть призвести до зміщення моделі. Слід докласти зусиль, щоб збалансувати набір даних шляхом надмірної вибірки класу меншості (фейкова інформація) або недостатньої вибірки класу більшості (справжня інформація). Це гарантує, що модель навчається на репрезентативній вибірці.

Поділ набору даних: набір даних поділено на підмножини для навчання, перевірки та тестування. Навчальний набір використовується для навчання моделі LSTM, набір перевірки допомагає налаштувати гіперпараметри, а тестовий набір використовується для оцінки продуктивності моделі.

Вилучення функцій: текстові дані потрібно перетворити на числовий формат, придатний для моделювання на основі LSTM. Для представлення тексту

як числових векторів можна використовувати такі методи, як вбудовування слів (Word2Vec, GloVe) або TF-IDF [28].

Навчання на основі мережевої моделі LSTM: мережева модель LSTM, розроблена для виявлення підробленої інформації, навчається на позначеному та попередньо обробленому наборі даних. Архітектура LSTM особливо добре підходить для даних послідовності, таких як текст, і здатна фіксувати контекстні залежності, які є вирішальними для розуміння мови.

Оцінювання моделі: продуктивність моделі на основі LSTM оцінюється за допомогою набору даних тестування. Такі показники, як точність, прецизійність, запам’ятовування, оцінка F1 і площа під кривою ROC (AUC-ROC), використовуються для вимірювання ефективності моделі в розрізненні справжньої інформації від підробленої.

Ітеративне уточнення: набір даних і модель часто піддаються ітеративному уточненню. Це може включати повторний перегляд процесу маркування для забезпечення точності, коригування етапів попередньої обробки, тонке налаштування моделі та збільшення розміру набору даних, коли нові дані стають доступними.

Формування високоякісного набору даних є критичною основою для успішного розроблення моделі штучного інтелекту на основі LSTM для виявлення фейкової інформації в соціальних мережах. Ретельно створений набір даних у поєднанні зі здатністю LSTM обробляти послідовні дані відіграє важливу роль у створенні надійної та точної моделі для виявлення та пом’якшення поширення фейкової інформації на платформах соціальних мереж.

Визначені критерії оцінювання достовірності новин є важливими для того, щоб зібрані дані були надійними та відповідали цілям дослідження. У контексті виявлення фейкової інформації в соціальних мережах за допомогою штучного інтелекту на базі мережі LSTM встановлено наступні критерії:

Релевантність заголовка й вмісту: заголовок новинної статті має відповідати вмісту статті. Цей критерій допомагає гарантувати, що назва статті не вводить читачів в оману, а зміст відповідає тому, що пропонується в

заголовку.

Достовірність джерела: довіра до видання або джерела, яке опублікувало новину, є вирішальним фактором. Надійні ЗМІ повинні мати ліцензію та виконувати юридичні зобов'язання. Інформація з надійних і авторитетних джерел, швидше за все, буде точною.

Позначка часу публікації: дата та час публікації є важливими показниками, особливо коли застарілі новини подаються як поточні події. Цей критерій допомагає виявити оманливу або застарілу інформацію.

Перевірка автора: аналіз автора статті може дати уявлення про надійність джерела. Вивчення досвіду автора у створенні контенту, що заслуговує на довіру, є дуже важливим для оцінки достовірності новин.

Цитати та посилання: перевірка на наявність посилань або цитувань у новині є ознакою відповідальної журналістики. Наявність цитат з авторитетних джерел може підвищити довіру до новини [29].

Сумнівний контент: виявлення сумнівних цитат, зображень або контенту в новинній статті може стати тривожним сигналом. Цей критерій допомагає виявити потенційно оманливу або маніпульовану інформацію.

Перехресна перевірка: перевірка того, чи повідомляють інші авторитетні ЗМІ про ту саму подію, підвищує довіру до новин. Відсутність підтвердження з кількох джерел може свідчити про потенційну неправду.

Ці заздалегідь визначені критерії слугували орієнтиром під час ручного збору даних. Кожен критерій відігравав свою роль в оцінці достовірності новинного контенту. Дотримуючись цих критеріїв, необхідно було створити набір даних, який включав би надійні й достовірні новинні статті, заклавши основу для надійного аналізу та навчання моделі штучного інтелекту.

Ці критерії забезпечили релевантність та достовірність зібраних даних, що є запорукою успіху будь-якого дослідження, спрямованого на виявлення фейкової інформації в соціальних мережах.

Незважаючи на те, що збір даних вручну може займати багато часу та ресурсів, він пропонує рівень ретельного аналізу та людського судження, який є

цінним у контексті виявлення фейкових новин. Людське втручання необхідне не лише для відбору та перевірки даних, але й для класифікації їх як справжніх або фейкових на основі встановлених критеріїв. Така ручна перевірка допомагає створити надійний набір правдивих даних.

Далі, розглянемо перелік інформаційних ресурсів та платформ, які є поширеними джерелами фейкових новин. Для цього було використано організацію "FAKE.NET.UA", яка спеціалізується на аналізі та виявленні дезінформації та фейкових новин в Україні. Список, наданий FAKE.NET.UA, містить URL-адреси веб-ресурсів, які були ідентифіковані як потенційні джерела неправдивої інформації. Використовуючи цей список, ми орієнтувалися на ці конкретні джерела під час збирання даних, забезпечуючи більш цілеспрямований і відповідний набір даних. Отримані URL-адреси було застосовано для керування процесом збирання даних вручну та під час визначення джерел, які слід контролювати на наявність потенційно фейкового новинного вмісту (рис. 2.7) [30].

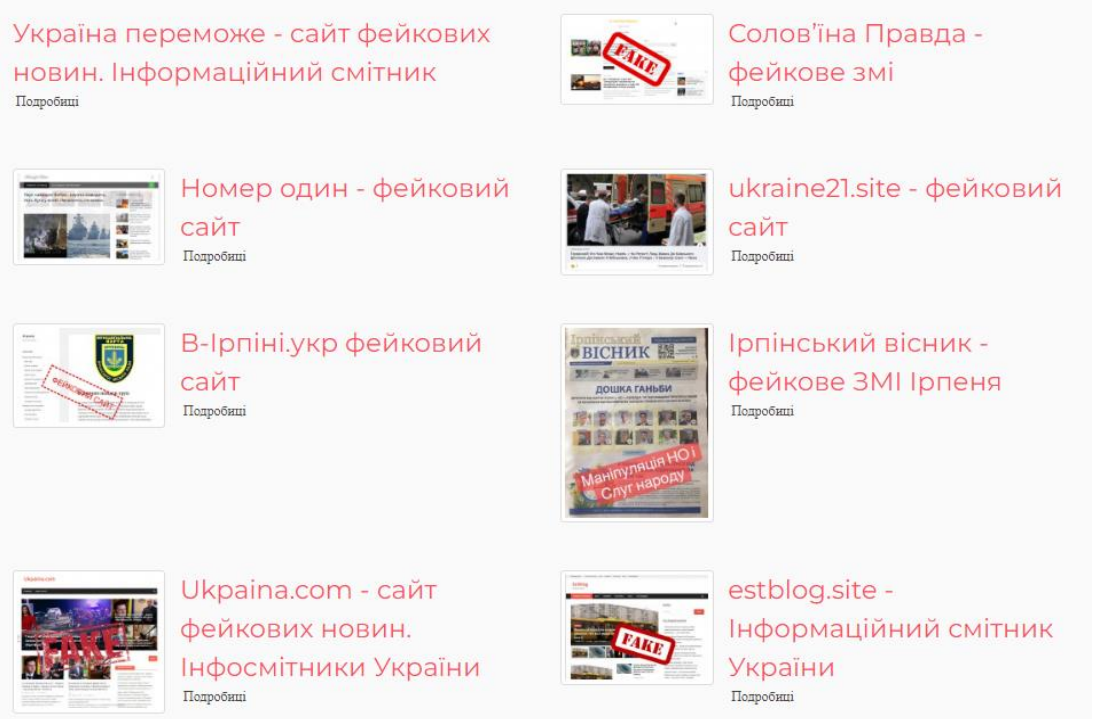


Рисунок 2.7 – Каталог сайтів з фейковими новинами від організації FAKE.NET.UA

Розглянемо створення структурованого набору даних у запропонованому

методу для виявлення фейків. Цей набір даних дозволяє підвищити точність та ефективність ШІ-моделі на основі LSTM у виявленні фейкових новин. Структурований набір даних, який містить критичні атрибути для аналізу, інтегровані у таблицю MS Excel з ретельною увагою до деталей. Ключовими компонентами цього набору даних є такі:

- title: позначає заголовок веб-сторінки або статті, що дає стисле уявлення про тему або предмет дослідження.
- text: складається з текстового вмісту статті, цей розділ інкапсулює основну інформацію, яку аналізуватиме модель штучного інтелекту.
- label: це важливий атрибут, який класифікує кожну новину як "ПРАВДИВУ" або "НЕПРАВДИВУ". Він відіграє фундаментальну роль на етапах навчання, тестування та валідації ШІ-моделі.
- url: містить веб-посилання, яке веде до джерела новини. Воно слугує відправною точкою для подальшого розслідування або перевірки.
- language: атрибут "language" вказує на мову, якою складено текст новини, і дозволяє проводити багатомовний аналіз, що ще більше розширює сферу виявлення фейкових новин.

Інтеграція цих атрибутів до таблиці MS Excel у роботі організовано для забезпечення точності та узгодженості даних. Кожна новина, отримана з веб-джерел, узгоджується з відповідними стовпчиками. Такий структурований підхід, підвищує зручність використання набору даних і сумісність з моделлю штучного інтелекту на основі LSTM. Це дозволяє реалізувати подальші етапи дослідження, а саме навчання, тестування та оцінювання здатності моделі розпізнавати правдивість новин.

Візуальне представлення процесу формування набору даних наведено на рис. 2.8 [31].

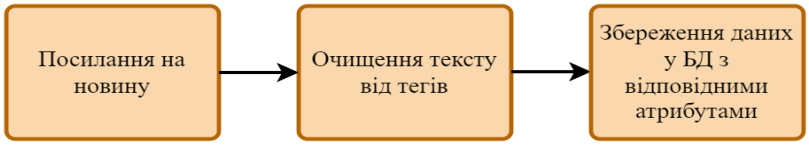


Рисунок 2.8 – Процес формування набору даних для удосконаленого методу виявлення фейкової інформації

Розглянемо процес представлення навчального набору даних із використанням ручної перевірки інформації. Такий процес перевірки було реалізовано у запропонованому автором магістерської роботи методі. Отже, було перевірено вміст кожної новини і присвоєно їй відповідну мітку "REAL" або "FAKE" у спеціально відведеному полі "label" в базі даних. Розглянемо такий приклад табличного представлення даних, що наведено на рис. 2.9.

	title	text	label	language	url
2	Шабунін та Шерембей планують зверне	Засновник Центру протидії корупції Віталій Шабулін та Дмитро Шерембей	FAKE	УКР	https://ukr-news.
3	Україна не отримає репарацій через те,	Україна не отримає репарацій від Росії, а її неповнолітні громадяни не мат	FAKE	УКР	https://www.my-
4	Україна готує підрив Кременчуцької ГЕ	Отримано дані про мінування і можливий підрив Кременчуцької ГЕС з ме	FAKE	УКР	https://infoteka.b
5	В Україні скасовують інвалідність, щоб	Чиновники пропонують реформувати Медико-соціальну експертну комісію	FAKE	УКР	http://www.prese
6	Захід змушує ЗСУ відправляти малоліт	Захід змушує Київ кидати молодь у м'ясорубку в спробі очистити Україну	FAKE	УКР	https://matrix.co.
7	Британія відправляє війська в Україну «	Повернення англійського спецназу свідчить про те, що Лондон прямо бра	FAKE	УКР	http://bbc-ccnn.c
8	В Україні планують призивати до армії	В українському парламенті зареєстрували законопроєкт про мобілізацію н	FAKE	УКР	http://personavip
9	Українські військовополонені не хочуть	Вдома - гірше. І переважна більшість із них не хоче ніякого обміну. Тому	FAKE	УКР	http://zhyrnalist.b
10	Туреччина заявила, що Крим – це росій	Президент Туреччини Реджеп Тайіп Ердоган зробив заяву про Крим, який	FAKE	УКР	https://mozgopit.
11	Україна створила бронежилети для ваг	Через провал контрнаступу українське керівництво ухвалило рішення відп	FAKE	УКР	https://kraina.biz.
12	Чехія заявила про нову хвилю проблем	Україна провалила контрнаступ і ризикує зіткнутися з новою хвилею проб	FAKE	УКР	https://ua24ua.ne
13	Ексміністр оборони Резніков відпочива	Колишній міністр оборони Резніков відпочиває на яхті з коханою, поки ко	FAKE	УКР	https://mykyivreg
14	Столтенберг підтвердив, що Росія втор	Генеральний секретар НАТО Єнс Столтенберг нібито підтвердив той факт	FAKE	УКР	http://ukpaina.co
15	У Берліні з'явилося антиукраїнське гра	Німецьке видання Frankfurter Allgemeine Zeitung опублікувало антиукраїнс	FAKE	УКР	https://ukraine21
16	Польща виганяє» українців – на м'ясо у	Польща починає масовий вилов і депортацію українців на м'ясо до ЗСУ ч	FAKE	УКР	https://news.mys
17	Благодійна акція «Київстар» розкрила г	Український стільниковий оператор «Київстар» зізнався в реальних втрата	FAKE	УКР	https://tsargrad.t
18	Україна звинуватила у провальному ко	Чим тільки київський режим не виправдовував свого провалу: то куші зава	FAKE	УКР	https://smotrim.u
19	Україна завдасть удару по хасидах в Уз	Київ готовий до шантажу та збройних провокацій проти паломників-хасид	FAKE	УКР	https://ukraina.ua

Рисунок 2.9 – Фрагмент набору даних для навчальної вибірки

Під час машинного навчання будемо здійснювати попереднє оброблення цих даних, що доволить усунути текстовий шум. Рис. 2.10 ілюструє схему попереднього оброблення інформації [31].

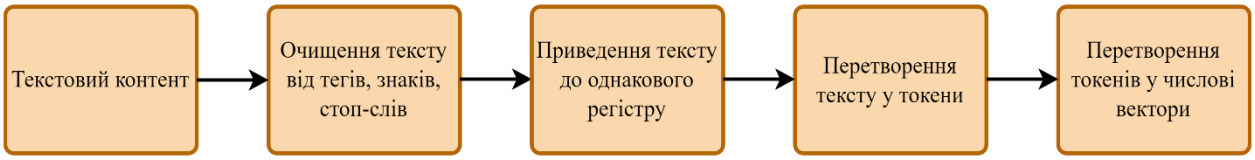


Рисунок 2.10 – Процес попереднього оброблення тексту в удосконаленому методі виявлення фейкової інформації

Щоб очистити текст шляхом видалення символів HTML, Python надає низку вбудованих функцій. Одним із поширених підходів є використання регулярних виразів для цієї мети [32]. Наприклад:

```
cleanText = re.compile(r'<[^>]+>|&[a-zA-Z0-9]+;')
```

`<[^>]+>`: зіставляє та видаляє теги HTML, укладені в кутові дужки,

наприклад <p>, тощо.

`&[a-zA-Z0-9]+;`: знаходить і видаляє escape-послідовності HTML, такі як `&`, `<`, `&#65`.

Щоб видалити стоп-слова з файлу, необхідно використовувати словники стоп-слів, сумісні з потрібною мовою, як-от українська та російська в нашому випадку. На щастя, доступна бібліотека Python, яка підтримує як російську, так і українську мови для стоп-слів. Ось приклад коду для отримання набору стоп-слів:

```
import nltk
nltk.download('stopwords')
from nltk.corpus import stopwords
stop_words = set(stopwords.words('ukrainian'))
```

Для стандартизації текстового регістру будуть використані стандартні методи Python. Ось приклад:

```
lowercase_text = text.lower()
```

Для усунення пунктуації ми скористаємося наступним методом:

```
translator = str.maketrans('', '', string.punctuation)
cleaned_text = text.translate(translator)
```

Цей код ефективно видалить усі розділові знаки зі змінної text.

Токенізація тексту є критичним завданням, яке потребує спеціальних інструментів і бібліотек [32]. Наприклад, бібліотека NLTK у Python полегшує токенізацію тексту, як показано нижче:

```
from nltk.tokenize import word_tokenize
tokens = word_tokenize(text)
```

Цей код використовує функцію word_tokenize з бібліотеки NLTK для токенізації тексту в масив даних. На рис. 2.11 представлено роботу модуля токенізації тексту в IT для виявлення фейків [32].

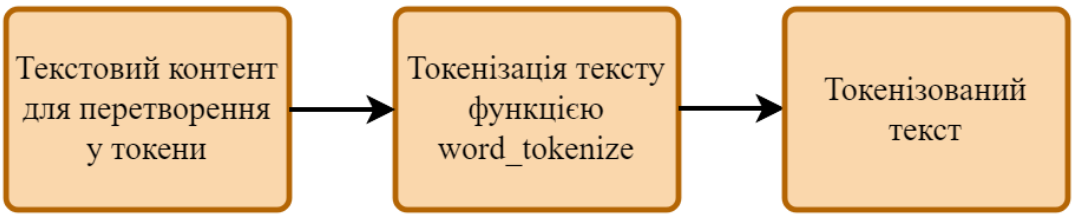


Рисунок 2.11 – Робота модуля токенізації тексту в методі виявлення фейків

Наступним етапом реалізації удосконаленого методу є перетворення набору токенів на вектори. Для цього використовують різні методи, серед яких варто відзначити: TF-IDF, Word2Vec, FastText, BERT, ELMo і Universal Sentence Encoder. Оскільки, навчальний набір містить тексти українською та російською мовами, розглянемо відповідні методи:

- Word2Vec;
- BERT;
- Universal Sentence Encoder.

Word2Vec – це техніка вбудовування слів, яка представляє слова як щільні вектори. Вона навчена передбачати контекстні слова з цільового слова у великих текстових корпусах. Це фіксує семантичні зв'язки між словами та використовується в різних завданнях оброблення природної мови. Математично модель представляє вектори слів A і B за допомогою косинусної подібності, яка наближається до 1, коли слова мають схоже значення, і зменшується, якщо вони різні. Формула косинусної подібності така [32]:

$$sim(A, B) = \cos(0) = (A * B)/(||A|| \, ||B||),$$

де $sim(A, B)$ – подібність косинуса,

A і B – вектори слів,

$||A||$ та $||B||$ – величини векторів.

Застосуємо бібліотеку Gensim для ефективного способу оброблення моделі Word2Vec, що дозволяє завантажувати попередньо підготовлені моделі та виконувати низку завдань обробки природної мови за допомогою вбудовування слів [32]:

```
from gensim.models import Word2Vec
model = Word2Vec.load("path/to/my/word2vec.model")
```

BERT – це найсучасніша модель оброблення природної мови. На відміну від попередніх моделей, які зчитують текст послідовно, BERT використовує двонаправлений підхід, аналізуючи контекст з обох сторін. Це дає змогу BERT глибше розуміти мову та надзвичайно добре виконувати різноманітні мовні завдання, як-от відповіді на запитання, аналіз настроїв тощо. Він попередньо навчений на великому текстовому корпусі, що робить його потужним

інструментом для широкого спектру застосувань NLP. Що відрізняє BERT від інших, так це його здатність розуміти контекст і значення слів у реченні, розглядаючи одночасно ліворуч і праворуч навколишні слова [32].

Процес токенізації BERT також унікальний. Він використовує токенізацію WordPiece, що означає, що він розбиває слова на менші підслова або частини. Наприклад, слово «нещастя» можна розділити на «не» і «щастя». Такий підхід дозволяє BERT ефективно обробляти слова, які не входять у словниковий запас, і забезпечує детальний аналіз тексту.

Щоб використовувати BERT, ми вводимо текст як послідовність токенів, і він генерує контекстні вбудовування для кожного лексема в реченні. Це дозволяє моделі розуміти зв'язки та значення між словами з урахуванням контексту. Процес токенізації BERT гарантує, що кожен маркер зіставляється з відповідним вбудованим елементом, і ці вбудовані компоненти використовуються для різноманітних завдань обробки природної мови [32].

Розглянемо використання цієї моделі для реалізації запропонованого методу у Python [32]:

```
    from transformers import DistilBertModel, DistilBertTokenizer, BertModel, BertTokenizer
    model_class, tokenizer_class, pretrained_weights = (DistilBertModel, DistilBertTokenizer, 'distillbert-base-uncased')
    model_class, tokenizer_class, pretrained_weights = (BertModel, BertTokenizer, 'bert-base-uncased')
```

Universal Sentence Encoder – це модель, розроблена Google, яка кодує речення або короткі тексти в числові вектори фіксованої довжини, що робить його придатним для виконання різноманітних завдань обробки природної мови, таких як класифікація тексту та аналіз семантичної подібності. Він відомий своєю здатністю вловлювати значення та контекст речень, що робить його цінним інструментом у програмах NLP.

Модель обробляє вхідний текст змінної довжини та генерує 512-вимірний векторний вихід. Цей кодер є універсальним і може використовуватися для таких завдань, як класифікація тексту, кластеризація тексту, семантичний пошук

схожості тексту та виділення міжмовних ознак. Він сумісний із згортковими нейронними мережами (CNN) і виявляється цінним у різних програмах обробки природної мови.

Застосування моделі Universal Sentence Encoder у мові Python [32]:

```
import tensorflow as tf
import tensorflow_hub as hub
module_url = "https://tfhub.dev/google/universal-sentence-encoder-multilingual/3"
embed = hub.load(module_url)
```

Отже, у цьому підрозділі було описано етапи підготовки даних для навчання нейронної мережі, зокрема попереднє оброблення даних і вибір основних бібліотек Python для векторизації тексту. Це дозволило використати моделі нейронного навчання, а саме Universal Sentence Encoder для розроблення подальших етапів запропонованого методу виявлення фейків.

2.3 Обгрунтування вибору оптимальної моделі машинного навчання для удосконалення методу виявлення фейкової інформації в соцмережах

У контексті вдосконалення методу виявлення фейкової інформації в соціальних мережах за допомогою застосування штучного інтелекту на основі мережі LSTM буде досліджено та оцінено різні моделі машинного навчання. Вибір конкретної моделі ми визначатимемо відповідністю її характеристик вимогам завдання. Оскільки виявлення фейкових новин – це в основному завдання класифікації, тому необхідно вивчити існуючі підходи до класифікації тексту та оцінить їх придатність для поставленого завдання.

Для ефективного вибору моделі ми будемо здійснювати перевірку сумісності кожної моделі машинного навчання з унікальними особливостями та вимогами завдання виявлення фейкових новин. На вибір моделей впливають кілька ключових факторів, такі як характер набору даних, складність завдання і бажаний результат.

Важливим кроком в оцінюванні моделей є розуміння набору даних, який

використовується для тестування. Розуміючи розмір, різноманітність і специфіку набору даних, можна приймати обґрунтовані рішення щодо моделей машинного навчання, які, найімовірніше, будуть успішними. Набір даних буде використовуватися для навчання, валідації та тестування моделей, і ретельне вивчення його властивостей має першорядне значення для вибору моделі [33].

Для забезпечення надійності оцінювання моделі, використовується системний підхід до поділу даних. Набір даних поділяється на підмножини тестування, навчання та валідації. Розподіл даних всередині цих підмножин наступний:

15% – для тестування; 65% – для навчання; 35% – для валідації.

Такий розподіл обрано для того, щоб досягти балансу між навчанням моделей на достатній кількості даних і суворою оцінкою їхньої ефективності. Параметр random_state встановлено на постійне значення, а саме 42, щоб забезпечити відтворюваність при навчанні моделі. Крім того, генератори псевдовипадкових чисел у використаних модулях ініціалізуються постійним значенням 0, щоб мінімізувати їхній вплив на результати тестування. Така відтворюваність гарантує узгодженість результатів при використанні тих самих параметрів на тому самому наборі даних [33].

Наступним кроком необхідно дослідити додаткові параметри набору даних, включаючи ідентифікацію часто повторюваних слів (рис. 2.12) і співвідношення фейкових новин до правдивих (рис. 2.13).

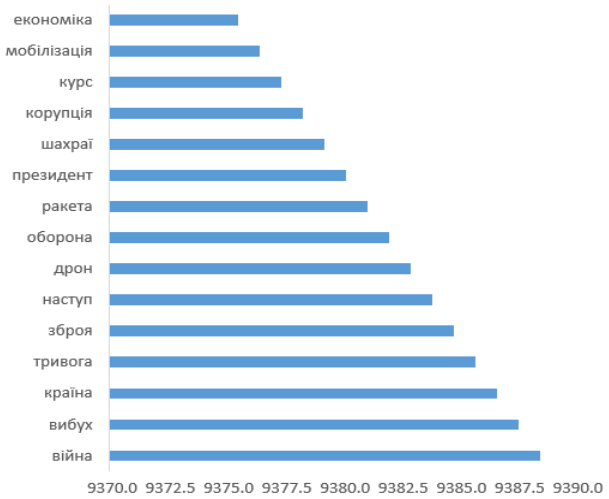


Рисунок 2.12 – Найчастіше повторювані слова у текстах новин

На рис. 2.12 показано слова з найвищими балами на основі метрики частоти слів, TF * IDF. Цей показник обчислюється шляхом множення двох факторів: частоти слова в документі та його зворотної частоти в усій колекції документів.

Варто підкреслити, що після видалення стоп-слів у лідерів TF*IDF є помітна різниця між фейковими та правдивими новинами. Ця різниця чітко відображена на рис. 2.13 [33].

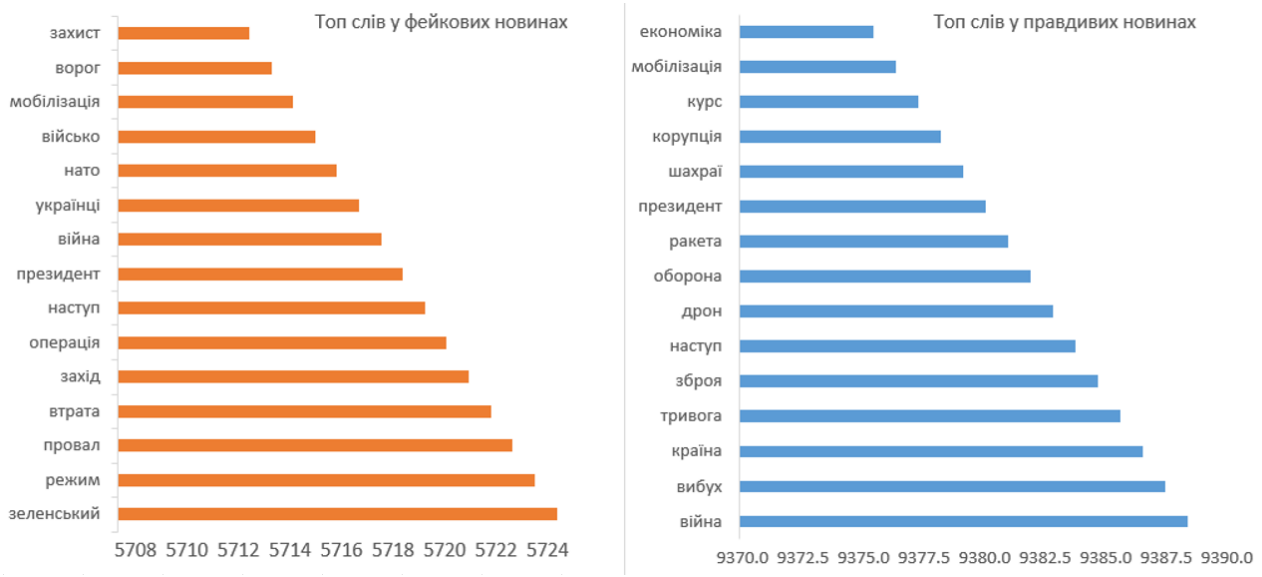


Рисунок 2.13 – Порівняння найчастіше повторюваних слів у правдивих та фейкових новинах

Розглянемо пропорції між фейковими та правдивими новинами в отриманому навчальному наборі даних (рис. 2.14) [33].

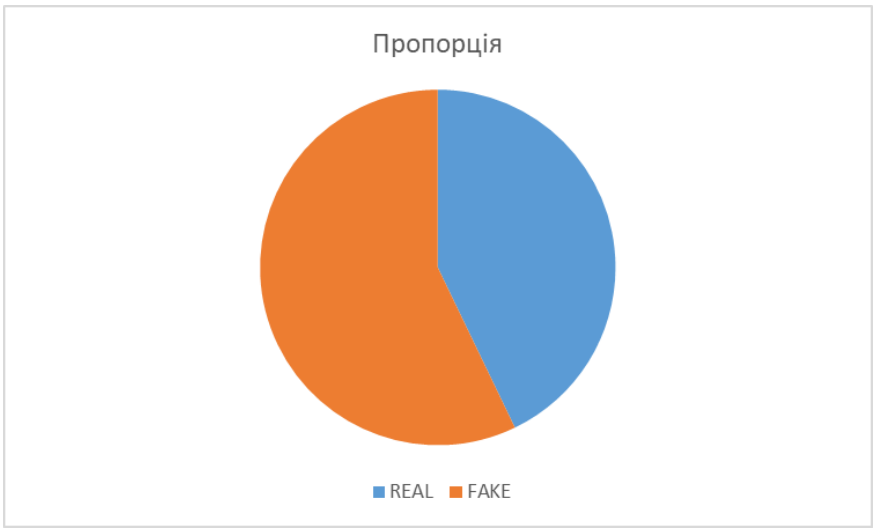


Рисунок 2.14 – Співвідношення фейкових новин до правдивих

Як бачимо, існує відносно рівний розподіл між фейковими, та справжніми новинами.

Серед поточних рішень машинного навчання виділяються кілька моделей: рекурентні нейронні мережі (RNN), ансамблі класифікатори, Word2Vec, BERT, SVM, багатошаровий перцептрон, і, зокрема, мережі довготривалої короткочасної пам’яті (LSTM). LSTM є особливо перспективним для посилення виявлення фейкової інформації в соціальних мережах [34].

У роботі досліджуються моделі нейронного навчання, реалізовані засобами бібліотеки Google Tensorflow і універсального кодувальника речень, як описано в розділі 2.2.

У наступній порівняльній таблиці представлено результати дослідження різних моделей, які використовувалися для тестування на самостійно зібраному наборі даних, що складається з україномовних новинних статей. Аналіз цих моделей, включаючи їх архітектуру, параметри навчання та показники продуктивності, дає цінну інформацію про їхню ефективність у виявленні фейкових новин в україномовному контексті (табл. 2.1) [34].

Таблиця 2.1

Порівняння ефективності моделей класифікаторів фейкових новин

Тип	Розмір шару	Функція активації	Навчальні епохи	Точність класифікації	Помилка типу 1	Помилка типу 2	Оцінка F1	Запам'ято-вування
LSTM	128	ReLU	50	89%	1%	15%	0.88	0.90
CNN	64	Sigmoid	30	77%	8%	30%	0.75	0.86
GRU	2	Tanh	40	68%	2%	40%	0.70	0.88
Bi-LSTM	128	ReLU	45	82%	3%	45%	0.83	0.81
RNN	128	Sigmoid	35	64%	9%	26%	0.67	0.74
BERT	12	GELU	45	75%	3%	25%	0.73	0.80
LR	12	Sigmoid	5	48%	28%	36%	0.50	0.60

Помилка типу 1 (хибно-позитивний): помилка типу 1 виникає, коли модель неправильно передбачає позитивний клас, тоді як фактичний клас негативний. У контексті ідентифікації фейкових новин це означає, що модель помилково позначає справжню новинну статтю як фейкову.

Помилка типу 2 (хибно-негативний): помилка типу 2 виникає, коли модель

неправильно передбачає негативний клас, тоді як фактичний клас позитивний. У контексті виявлення фейкових новин це означатиме, що модель не зможе ідентифікувати фейкову статтю новин, дозволяючи їй вважатися справжньою.

Оцінка F1: це міра точності моделі, яка враховує як точність, так і запам’ятовування. Це гармонійне середнє значення точності та запам’ятовування, що робить його придатним показником, коли ми хочемо збалансувати як помилкові позитивні, так і помилкові негативні результати. Вищі показники F1 вказують на кращий баланс між точністю та пам’яттю.

Запам’ятовування (чутливість): запам'ятовування, також відоме як чутливість або істинно позитивний коефіцієнт, вимірює здатність моделі ідентифікувати всі відповідні випадки. У контексті виявлення фейкових новин відкликаність – це відношення правильно ідентифікованих фейкових новинних статей до загальної кількості фактичних фейкових новинних статей. Висока запам’ятовуваність вказує на те, що модель може зафіксувати значну частину фейкових новин [34].

Серед розглянутих моделей нейронна мережа довготривалої короткочасної пам’яті (LSTM) виділяється як найбільш ефективна модель з точки зору точності класифікації з дивовижною точністю 89%. Модель LSTM демонструє низьку помилку типу 1 (1%) і помилку типу 2 (15%), що вказує на її здатність ефективно виявляти як помилкові позитивні, так і помилкові негативні результати. Крім того, показник F1 (0,88) і чутливість (0,90) ще більше підкреслюють її чудову ефективність у збалансуванні точності та запам’ятовування, що є вирішальним для вибору автором магістерської роботи для розроблення удосконаленого методу виявлення фейкових новин.

Залежність між точністю моделі та кількістю епох навчання проілюстровано на рис. 2.15 [34].

Оптимальна кількість епох залежить від конкретного набору даних і архітектури моделі, і часто визначається шляхом експериментів і перевірки.

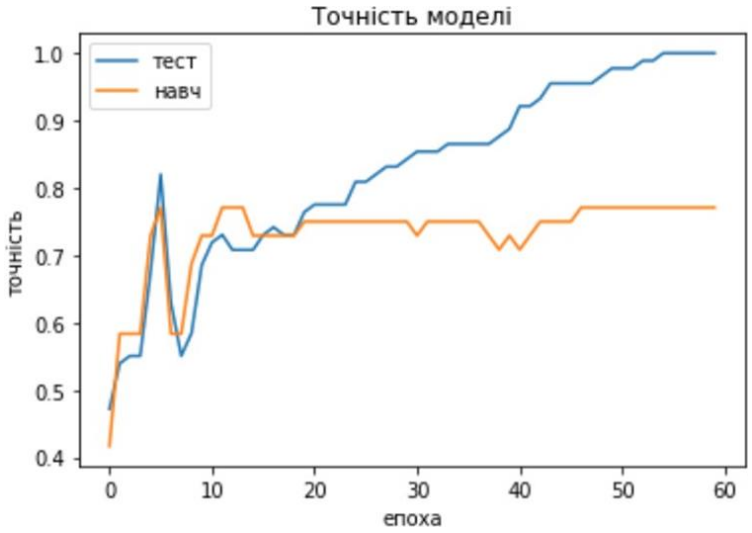


Рисунок 2.15 – Залежність між точністю моделі та кількістю епох

Кожен прохід через набір даних називається епохою. Вагові коефіцієнти моделі оновлюються після кожної епохи, щоб підвищити точність прогнозів. Збільшення кількості епох навчання може призвести до кращої продуктивності моделі до певного моменту, після чого подальше навчання може призвести до переобладнання, коли модель надто точно відповідає навчальним даним і погано узагальнює нові, невідомі дані.

Після ретельного дослідження та оцінювання різних моделей машинного навчання, стає очевидним, що нейронна мережа LSTM (Long Short-Term Memory) постійно демонструє найвищу точність у завданні ідентифікації фейкових новин. Це дослідження підкреслює придатність та остаточний вибір LSTM для удосконалення методу виявлення фейкової інформації в соціальних мережах.

2.4 Висновок до розділу 2

Дослідження, представлене в розділі 2 було зосереджено на розробленні низки складно пов’язаних процесів побудови удосконаленого методу виявлення фейкових новин за допомогою штучного інтелекту. Такий підхід дозволяє підвищити точність та ефективність виявлення неправдивої інформації в

новинних текстах шляхом ралізації таки етапів.

Етап підготовки структурованого набору даних, необхідного для інтелектуального аналізу. Важливість надійних даних для успішної реалізації запропонованого методу на основі застосування моделей машинного навчання неможливо переоцінити, а компіляція набору даних у цій роботі закладає основу для наступних етапів.

Необхідним етапом розроблення методу є обгрунтування оптимальної моделі машинного навчання для виявлення фейкових новин. Ретельне вивчення цих моделей дає цінну інформацію про їхні сильні сторони та обмеження. Серед розглянутих моделей нейронна мережа LSTM виділяється як найбільш ефективна за точністю класифікації, а саме 89%. Модель LSTM демонструє низьку помилку типу 1 (1%) і помилку типу 2 (15%), що вказує на її здатність ефективно виявляти як помилкові позитивні, так і помилкові негативні результати. Крім того, показник F1 (0,88) і чутливість (0,90) ще більше підкреслюють її чудову ефективність у збалансуванні точності та запам’ятовування, що є вирішальним щодо вибору для розроблення методу виявлення фейкових новин.

3 ПРОГРАМНА РЕАЛІЗАЦІЯ УДОСКОНАЛЕНОГО МЕТОДУ ВИЯВЛЕННЯ ФЕЙКОВОЇ ІНФОРМАЦІЇ

3.1 Обґрунтування вибору програмних засобів для реалізації запропонованого методу виявлення фейкової інформації

Успішне розроблення програмного забезпечення, призначеного для виявлення фейкової інформації у середовищі соціальних мереж за допомогою штучного інтелекту на основі мережі LSTM вимагає ретельного розгляду інструментів і технологій, які можуть ефективно перетворити концептуальну основу на функціональну та ефективну систему. Цей підрозділ присвячений обґрунтуванню вибору конкретної мови програмування, яка буде покладена в основу комп'ютеризації методу, що спрямований на виявлення фейкової інформації в динамічному ландшафті соціальних мереж. У цьому підрозділі буде проведено комплексне оцінювання обраної мови програмування, яке підкреслює, наскільки вона відповідає унікальним вимогам проєкту. Він має на меті з'ясувати, чому саме відповідній мові програмування було надано перевагу над альтернативними варіантами, зосереджуючись на таких факторах, як універсальність, продуктивність, доступні бібліотеки та фреймворки, підтримка спільноти та її здатність сприяти досягненню цілей, окреслених у магістерській кваліфікаційній роботі.

Набір мов програмування великий і різноманітний. Наразі є чимало інструментів і рішень для розроблення програмного забезпечення.

Одна з таких мов програмування, C#, заслуговує на розгляд у цьому контексті роботи. C# – це потужна мова зі статичними типами, відома своєю надійністю та придатністю для різних областей застосування. Вона широко використовується для розроблення додатків Windows, розроблення ігор за допомогою Unity та програмного забезпечення корпоративного рівня [35].

У конкретному контексті впровадження запропонованого методу виявлення підробленої інформації в динамічній сфері соціальних мереж із використанням штучного інтелекту та мереж LSTM, C# має певні обмеження.

Такі недоліки роблять її менш придатною для цього завдання:

Складність: C# може бути нелегкою, особливо коли мова йде про складні алгоритми машинного навчання, такі як мережі LSTM. Її багатослівність і жорстка типізація можуть призвести до довших циклів розроблення та збільшення кількості рядків коду, що може не узгоджуватися з необхідністю швидкого створення прототипів і експериментів у розробці складних моделей штучного інтелекту.

Обмежені бібліотеки машинного навчання: на відміну від Python, який має потужну систему бібліотек і фреймворків машинного навчання, C# має більш обмежений вибір. Це може перешкодити бездоганній інтеграції передових методів машинного навчання, необхідних для успіху проєкту.

Спільнота та ресурси: хоча C# має активну спільноту розробників, вона не така поширена, як спільнота Python, а її спільнота машинного навчання та ШІ відносно менша. Це означає, що доступ до ресурсів, проєктів з відкритим вихідним кодом і своєчасна підтримка можуть бути більш обмеженими порівняно з Python.

Сумісність з API соціальних мереж: C# може не мати такого ж рівня сумісності з API популярних соціальних мереж, як Python. Це може перешкоджати розгортанню та інтеграції програмного забезпечення в рамках соціальних мереж для аналізу в реальному часі, що є критичною вимогою для цього проєкту [35].

Отже, C# не може бути найкращою мовою програмування для реалізації інформаційної технології виявлення фейкової інформації в соціальних мережах за допомогою штучного інтелекту на основі мережі LSTM. Складність мови, обмеженість бібліотек машинного навчання та потенційні проблеми сумісності з API соціальних мереж роблять її менш придатною для динамічного та складного завдання, що стоїть перед нами.

Python, відомий своєю універсальністю та придатністю для різних доменів, стає оптимальним вибором для програмної реалізації запропонованого методу [36].

Python відомий з кількох вагомих причин, які роблять його винятково придатним для цього завдання:

Широка можина бібліотек і фреймворків: Python пропонує велику кількість бібліотек машинного навчання та оброблення природної мови, таких як TensorFlow, Keras, PyTorch, NLTK, spaCy тощо. Ці бібліотеки забезпечують надійну основу для розроблення та впровадження складних алгоритмів машинного навчання, включаючи мережі LSTM. Доступність цих ресурсів спрощує процес розроблення, уможливлюючи швидке створення прототипів і експериментування, необхідні для складних моделей ШІ.

Спільнота та ресурси: Python може похвалитися однією з найактивніших і наймасштабніших спільнот розробників у всьому світі. Ця спільнота забезпечує доступ до великої кількості проєктів з відкритим кодом, сховищ коду, форумів і навчальних посібників, присвячених машинному навчанню та штучному інтелекту. Значна підтримка та доступні ресурси сприяють швидкому вирішенню проблем і просуванню проєкту.

Універсальність та інтеграція: Python дуже універсальний і адаптований, що робить його чудовим вибором для взаємодії із зовнішніми API та інтеграції з платформами соціальних мереж, де відбуватиметься виявлення підробленої інформації. Його сумісність із широким спектром платформ і технологій забезпечує плавну взаємодію та аналіз у реальному часі, що є ключовим аспектом цього проєкту.

Швидке розроблення та створення прототипів: лаконічний, зрозумілий і читабельний синтаксис Python прискорює цикли розроблення та сприяє ефективному створенню прототипів. Ця властивість є особливо корисною при точному налаштуванні складних моделей штучного інтелекту, оскільки сприяє швидкому експериментуванню та ітераціям [36].

Найсучасніше машинне навчання: бібліотеки та фреймворки машинного навчання Python знаходяться на передньому краї досліджень у галузі ШІ. Інтеграція передових методів, таких як LSTM-мережі, легко досяжна, гарантуючи, що розроблена технологія залишатиметься на найвищому рівні

серед можливостей ефективного виявлення фейкової інформації.

Отже, для програмної реалізації запропонованого методу буде використано набір бібліотек Python, кожна з яких виконує певні ролі в попередньому обробленні даних, машинному навчанні та візуалізації [36].

– NumPy: це фундаментальна бібліотека Python для чисельних обчислень. Вона забезпечує підтримку великих багатовимірних масивів і матриць разом із математичними функціями для роботи з цими масивами. У нашій реалізації NumPy допоможе ефективно зберігати дані та маніпулювати ними, що є ключовим аспектом підготовки даних машинного навчання.

– Scikit-Learn (sklearn): це універсальна бібліотека машинного навчання для Python. Вона пропонує прості та ефективні інструменти для інтелектуального аналізу даних. У нашому контексті Scikit-Learn буде використовуватися для навчання, перевірки та оцінки моделей машинного навчання, включаючи мережі LSTM.

– TensorFlow: це бібліотека машинного навчання з відкритим кодом, розроблена Google. Вона надає повну екосистему інструментів для машинного та глибокого навчання. Ми будемо використовувати TensorFlow для впровадження мереж LSTM, оскільки вона пропонує спеціалізовану підтримку глибокого навчання та нейронних мереж [36].

– Pandas: це бібліотека обробки та аналізу даних для Python. Вона відмінно підходить для очищення, підготовки та дослідження даних. Наша реалізація використовуватиме Pandas для завантаження, попереднього оброблення та структурування набору даних, гарантуючи його готовність до навчання та оцінювання.

– Matplotlib: це широко використовувана бібліотека Python для створення статичних, анімованих або інтерактивних візуалізацій на Python. У нашій реалізації Matplotlib дозволить генерувати візуальні представлення для аналізу та представлення результатів виявлення фейкової інформації.

– NLTK (Natural Language Toolkit): це бібліотека для роботи з даними людської мови та оброблення тексту. Вона надає прості у використанні

інтерфейси для лінгвістичних даних і аналізу тексту. У нашому проєкті NLTK буде цінним ресурсом для попередньої обробки тексту та лінгвістичного аналізу, що дозволить точніше виявляти фейкову інформацію [36].

– Seaborn: це бібліотека візуалізації даних, створена на основі Matplotlib. Вона забезпечує інтерфейс високого рівня для представлення привабливої та інформативної статистичної графіки. Будемо використовувати Seaborn для покращення візуальної якості та інтерпретації згенерованих графіків і діаграм.

– Gensim: це бібліотека Python для моделювання тем і аналізу схожості документів. У нашому проєкті Gensim допомагатиме в аналізі текстових даних, зокрема у визначенні тем у статтях новин і публікаціях у соціальних мережах.

– PyStemmer: це реалізація на Python алгоритмів стемінгу Snowball. У той час як UkrStemmer спеціалізується на обробленні текстів українською мовою, PyStemmer підтримує розпізнавання на багатьох мовах, включаючи українську. Вона використовуватиметься для зведення слів до базової форми для більш ефективного аналізу тексту в контексті україномовного контенту соціальних мереж [36].

Для розроблення веб-клієнта, який слугуватиме інтерфейсом користувача для перевірки новин на наявність неправдивої інформації, будемо використовувати комбінацію технологій, включаючи Flask для веб-фреймворку, HTML для структурування вмісту та CSS для стилізації. Ці технології добре підходять для створення зручного та інтерактивного веб-інтерфейсу [26].

Крім того, поєднання Flask, HTML і CSS дозволяє бездоганно інтегрувати динамічний вміст, створений кодом Python (Flask), і статичну структуру та стиль веб-сторінки (HTML і CSS). Завдяки такому підходу пропонується ефективний спосіб створити інтерактивний та адаптивний веб-інтерфейс ПЗ, за доомогою якого користувачі можуть вводити статті новин або публікації в соціальних мережах для аналізу фейкових новин, отримувати результати та легко взаємодіяти з ПЗ [36].

Отже, поєднання бібліотек, підтримки спільноти, універсальності та можливостей інтеграції Python ідеально відповідає багатогранним вимогам

поставленого завдання. Рішення обрати Python як мову програмування для програмної реалізації запропонованого методу виявлення фейкової інформації в соціальних мережах за допомогою штучного інтелекту на основі мережі LSTM ґрунтується на її здатності спростити розроблення, надати доступ до потужних інструментів штучного інтелекту та сприяти динамічному аналізу в реальному часі на платформах соціальних мереж. Цей вибір значною мірою сприятиме успішній реалізації поставленої в роботі мети.

3.2 Розроблення структури програмного засобу

Для програмної реалізації запропонованого методу для виявлення фейкової інформації в новинних текстах за допомогою штучного інтелекту на основі мережі LSTM було побудовано систему з кількох модулів, кожен з яких виконує такі функції:

1. Модуль попереднього оброблення: відповідає за підготовку вхідних даних (новинних статей або публікацій у соціальних мережах) для аналізу. Він включає в себе різноманітні завдання, такі як очищення тексту, токенізація, видалення стоп-слова та створення основи. Цей модуль гарантує, що текстові дані мають формат, придатний для подальшого аналізу (рис. 3.1) [37].

2. Модуль інтелектуального аналізу: цей модуль є ядром програмного засобу. Він використовує методи машинного навчання на основі мережі LSTM для виконання глибокого аналізу текстових даних. Цей модуль оцінює мовні особливості, контекст і шаблони в тексті, щоб визначити потенційні ознаки неправдивої інформації. Він використовує методи оброблення природної мови (NLP) для семантичного аналізу (рис. 3.2) [37].

Рисунок 3.1 – Модуль попереднього оброблення тексту

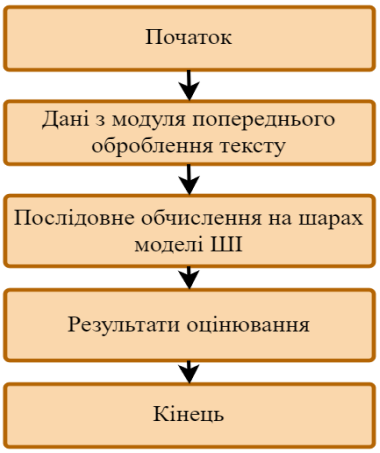


Рисунок 3.2 – Модуль інтелектуального аналізу

3. Модуль прийняття рішень: розроблений для оброблення результатів, отриманих з модуля інтелектуального аналізу. Він оцінює ймовірність того, що інформація є неправдивою або вводить в оману. У цьому модулі оператор прийняття рішення виконує перевірку щодо правильності передбачення. Модуль може використовувати попередньо визначені критерії прийняття рішень або моделі машинного навчання для прийняття рішень щодо класифікації (рис. 3.3) [37].

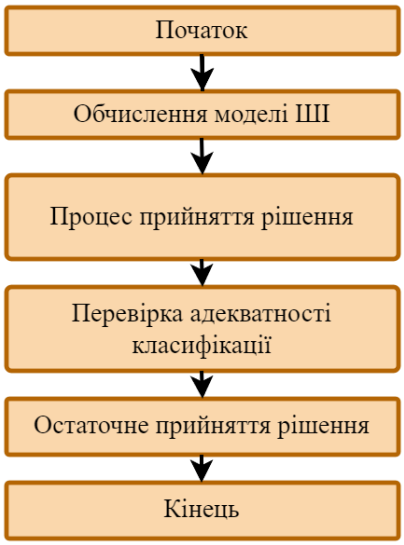


Рисунок 3.3 – Модуль прийняття рішень

4. Розроблений модуль бази даних новин: служить сховищем для зберігання новинних статей і вмісту соціальних мереж. Це дозволяє системі розширити свою базу даних для аналізу та, якщо необхідно, отримати додаткову інформацію про джерела новин. Модуль також може надавати історичні дані для подальшого аналізу (рис. 3.4) [38].

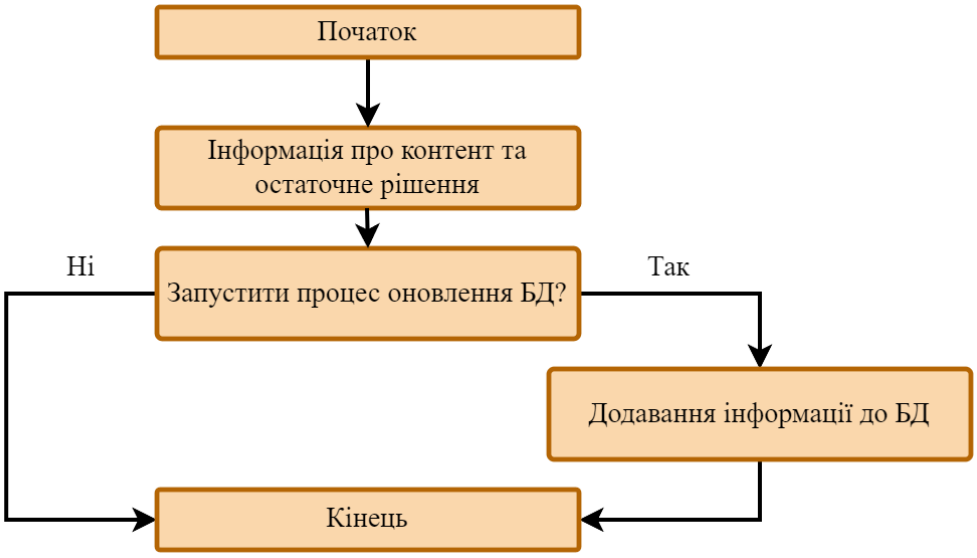


Рисунок 3.4 – Модуль бази даних новин

5. Модуль навчання (перенавчання) нейронної мережі: необхідний для підтримки нейронної мережі LSTM. Це дозволяє перенавчати мережу за допомогою оновлених даних або точно налаштувати модель для покращення її продуктивності (рис. 3.5) [38].

Рисунок 3.5 – Модуль навчання (перенавчання) нейронної мережі

Ці модулі працюють злагоджено, щоб полегшити весь процес, від початкової попередньої обробки даних до остаточного прийняття рішення щодо автентичності вмісту новин. Архітектура програмного забезпечення гарантує, що аналіз є надійним, ефективним і адаптованим до мінливих тенденцій у неправдивій інформації.

Блок-схему алгортму роботи ПЗ, що реалізує запропонований метод, представлено на рис. 3.6 [25].

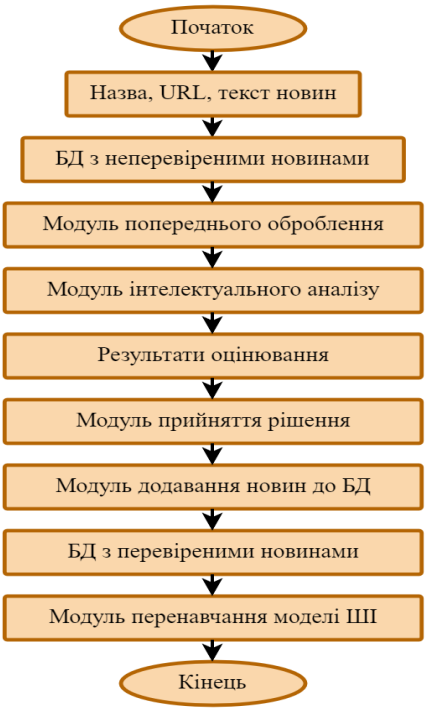


Рисунок 3.6 – Блок-схема алгортму роботи ПЗ

Нижче, автор магістерської дисертації наводить алгоритм роботи розробленого ним ПЗ, що реалізує метод виявлення фейкових новин.

Етап 1: введення новин для аналізу. Інформація про новини передається програмному продукту для аналізу. Ця інформація може включати статті новин або вміст соціальних мереж у текстовому форматі.

Етап 2: вибір користувача. Залежно від вибору користувача, вхідні дані спрямовуються на один із двох, шляхів:

1) Якщо користувач вирішує зберігати новини в базі даних непозначених новин, інформація записується в базу даних для подальшої обробки.

2) Якщо користувач вибирає негайний аналіз, дані переходять до модуля попередньої обробки [39].

Етап 3: використовується для попереднього оброблення тексту. Текстові дані проходять попереднє оброблення. Для цього етапу потрібно включити такі завдання, як:

1) Очищення тексту: видалення будь-яких нерелевантних символів або форматування.

2) Токенізація: розбиття тексту на окремі слова або лексеми.

3) Видалення стоп-слова: видалення загальних слів (наприклад, «і», «те»), які можуть не мати істотного значення.

4) Створення коренів: скорочення слів до їхніх кореневих форм.

5) Векторизація: перетворення текстових даних у числові вектори, придатні для машинного навчання.

Етап 4: аналізу за допомогою мережі LSTM. Попередньо оброблений текст передається до модуля інтелектуального аналізу, який використовує мережу LSTM (довгокороткочасна пам’ять). Ця модель глибокого навчання проводить комплексний аналіз текстових даних. Він враховує мовні особливості, контекст і шаблони в тексті.

Етап 5: семантичний аналіз. Розроблена програма, що реалізує удосконалений метод виявлення фейкових новин проводить семантичний аналіз, який передбачає розуміння сенсу та контексту тексту. Цей аналіз має

вирішальне значення для виявлення ознак неправдивої інформації або оманливого вмісту.

Етап 6: процес прийняття рішень. Результати, отримані від модуля інтелектуального аналізу, передаються до оператора прийняття рішень (ОПР). Цей оператор оцінює ймовірність того, що інформація є неправдивою або вводить в оману. Залежно від попередньо визначених критеріїв або моделей машинного навчання він приймає рішення про класифікацію [39].

Етап 7: інтеграція бази даних. Якщо вибраний користувачем шлях на етапі 2 мав зберігати новини в базі даних без позначок, результати аналізу разом із рішенням про класифікацію записуються в базу даних. Цей крок сприяє створенню набору даних для подальшого використання та покращенню продуктивності системи.

Етап 8: перенавчання нейронної мережі. Програмне забезпечення може ініціювати перенавчання нейронної мережі з урахуванням нової інформації та результатів аналізу. Це перенавчання покращує здатність моделі робити більш точні класифікації на основі еволюції набору даних.

Ці етапи окреслюють процес програмної реалізації розробленого методу, від початкового введення до остаточного рішення про класифікацію, і забезпечують гнучкість у тому, як обробляється вміст новин, як щодо аналізу, так і зберігання для подальшого використання.

3.3 Аналіз роботи ПЗ, його тестування та перевірка адекватності запропонованого методу виявлення фейкових новин

Головне вікно користувача програми реалізовано веб-додатком, що забезпечує зручність і доступність. Головне вікно ПЗ містить основні навігаційні елементи для вказівок користувача, форму подання новин і коротку інформацію про призначення запропонованого методу. Основний інтерфейс містить навігаційні компоненти для допомоги користувачам, форму для надсилання

новин і короткі відомості про цілі програми (рис. 3.7).

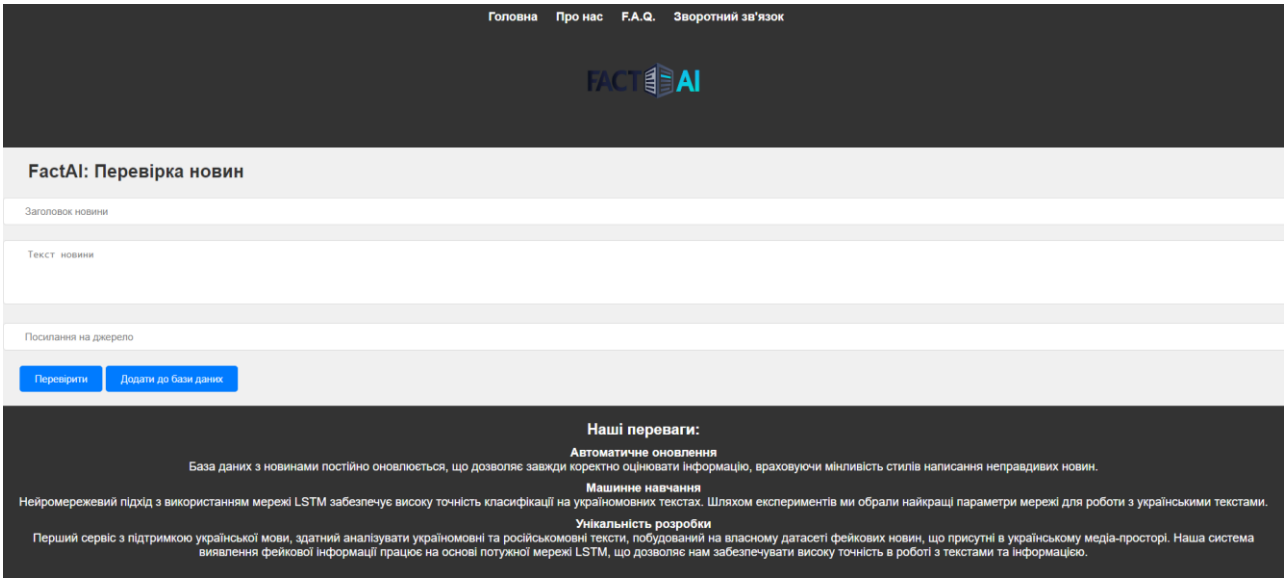


Рисунок 3.7 – Головна сторінка веб-застосунку

Розроблена форма для надсилання новинного матеріалу призначена для збирання інформації, наданої користувачами для перевірки. Вона містить поля для посилання та опису вмісту новини, і важливо зазначити, що немає обмежень щодо кількості символів у полі заголовка чи довжини тексту новини.

Коли користувач заповнить необхідні поля, він може продовжити, натиснувши одну з наступних кнопок:

– Перевірити: ця кнопка запускає процес перевірки, аналізуючи надіслані новини на наявність неправдивої інформації. Користувачі отримають оперативний відгук про достовірність новин.

– Додати до бази даних: ця кнопка дозволяє користувачам додавати надіслані новини до бази даних для подальшого використання та аналізу, збагачуючи набір даних новою інформацією.

Інтерфейс було розроблено таким чином, щоб він був зручним і зрозумілим, а користувачі могли легко надсилати новини на перевірку та ефективно користуватися функціями програми. Розроблена веб-програма забезпечуватиме безперебійну роботу для користувачів, які хочуть перевірити новини на достовірність.

Для перевірки правильності оброблення інформації під час первинної

взаємодії з програмою надано інформацію з джерела, що займається розповсюдженням фейкових новин (рис. 3.8).

Виберемо опцію «Перевірити», щоб ініціювати передачу даних форми на веб-сервер. Згодом дані проходять попереднє оброблення, трансформацію у вектори та подальший аналіз нейронної мережі. Після надсилання даних на сервер користувачі будуть спрямовані на сторінку підсумку цього аналізу, як показано на рис. 3.9.

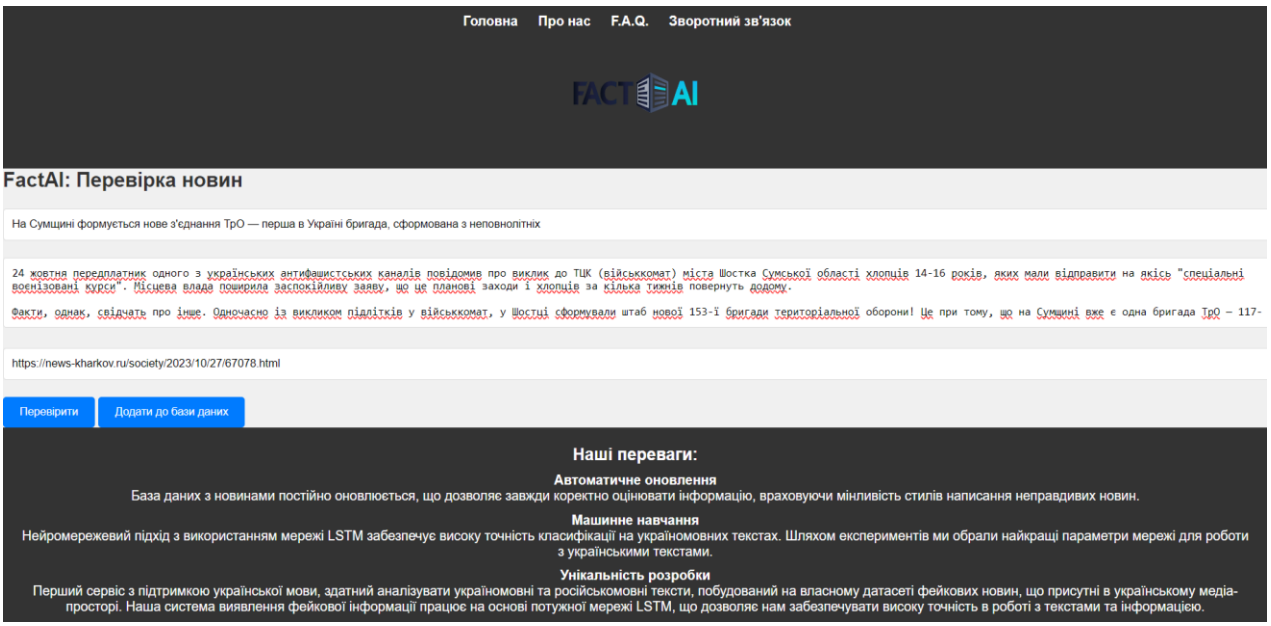


Рисунок 3.8 – Зразок заповнення полів програми

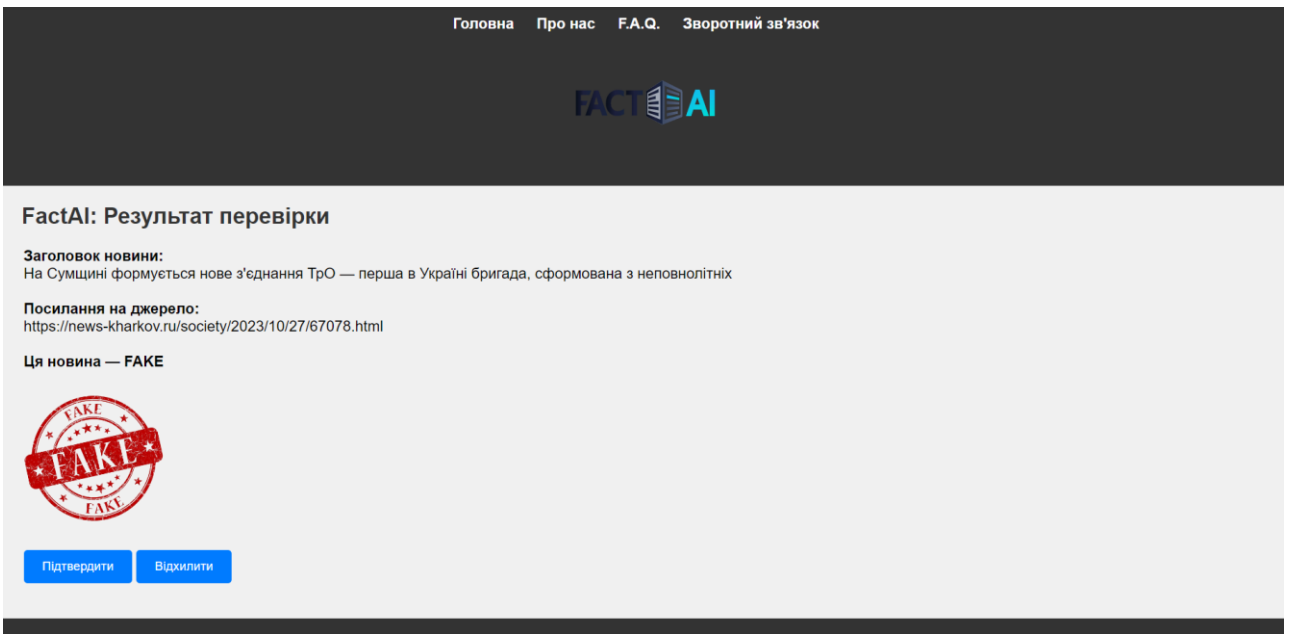


Рисунок 3.9 – Результат перевірки новини 1 запропонованим методом

Як бачимо, інформація, проаналізована за допомогою нашої розробленої

програми виявлення фейкових новин, виявилися неправдивою та згодом позначена як «Fake». Після цього користувачеві пропонуються варіанти вибору опцій у програмі за допомогою кнопок «Підтвердити» та «Відхилити», кожна з яких має певну мету:

– Підтвердити: ця дія передбачає додавання інформації про новину до бази новин категорії «Fake».

– Відхилити: вибір цієї дії означає, що інформація буде додана до бази новин, позначених контрастним ярликом «Real».

Цей керований користувачем процес сприяє розширенню бази даних новин, яка є невід'ємною частиною для подальшого перенавчання нейронної мережі. Натискаючи «Підтвердити», користувач перевіряє класифікацію системи, гарантуючи, що новину точно категоризовано. Після підтвердження додаток плавно направляє користувача до форми для додавання наступної новини. Одночасно позначена база даних новин збагачується іншою частиною даних, яка є кандидатом на подальше навчання мережі. Цей ретельний процес забезпечує збереження інформації у позначеній базі даних новин, сприяючи постійному вдосконаленню можливостей класифікації мережі (рис. 3.10).

title	text	label	url
Шабунін та Шерембей планують зверн	Засновник Центру протидії корупції Віталій Шабулін та Дмитро Шеремб	Fake	https://ukr-news.vn.u
Україна не отримає репарацій через т	Україна не отримає репарацій від Росії, а її неповнолітні громадяни не м	Fake	https://www.my-worl
Україна готує підрив Кременчуцької Г	Отримано дані про мінування і можливий підрив Кременчуцької ГЕС з м	Fake	https://infoteka.biz.u
В Україні скасовують інвалідність, що	Чиновники пропонують реформувати Медико-соціальну експертну комі	Fake	http://www.presentn
На Сумщині формується нове з'єднан	24 жовтня передплатник одного з українських антифашистських каналів	Fake	https://news-kharko

Рисунок 3.10 – Збереження інформації у позначеній базі даних новин

Не дивлячись на те, що новинна інформація відповідає класифікації нейронної мережі, вона у будь-якому випадку буде зберігатися у базу даних із певним атрибутом, тим самим сприяючи постійному навчанню мережі з оновленим сховищем інформації. Ця практика є необхідною, оскільки додані фейкові новини можуть демонструвати унікальні характеристики фальшивості, які можуть бути неоціненними для підвищення ефективності моделі.

Доведення адекватності розробленого методу. Для оцінювання ефективності та доведення адекватності розробленого підходу до виявлення фейкових

новин було проведено порівняльний аналіз результатів роботи запропонованого методу із результатами, що надаються офіційними джерелами інформації. На рис. 3.11 представлено результат перевірки новини 1 офіційним каналом Stopfake.

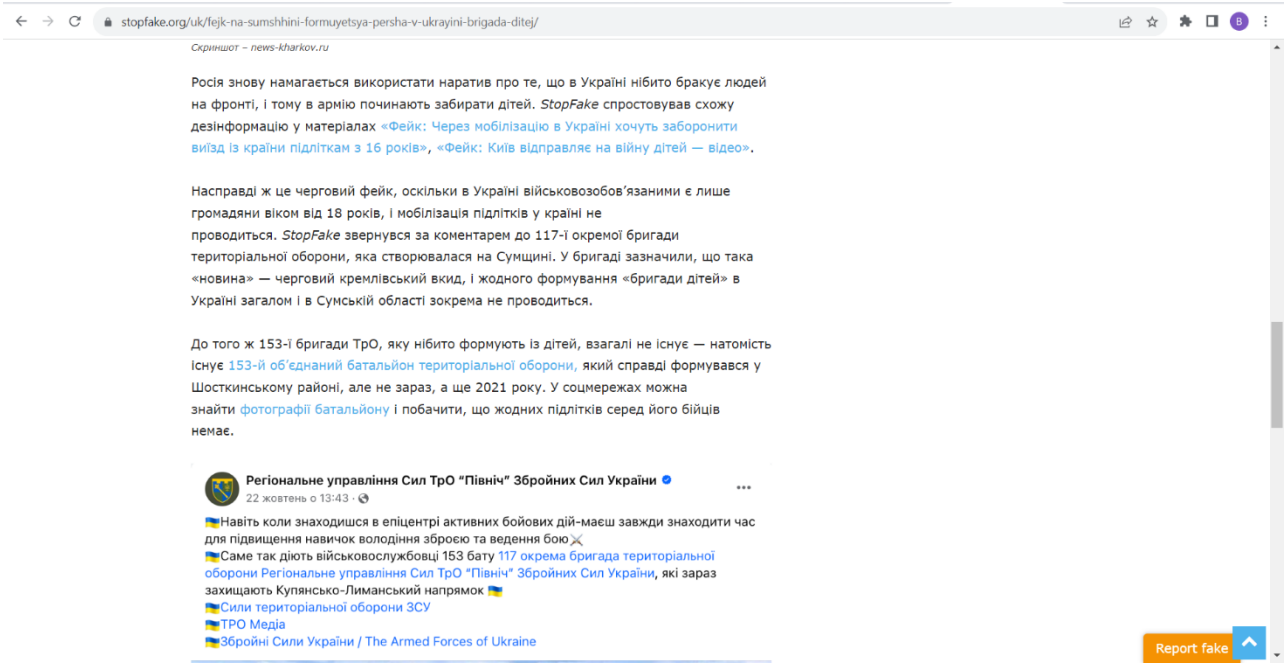


Рисунок 3.11 – Результат перевірки новини 1 офіційним каналом Stopfake

На рис. 3.12 представлено результати перевірки запропонованим методом новини 2.

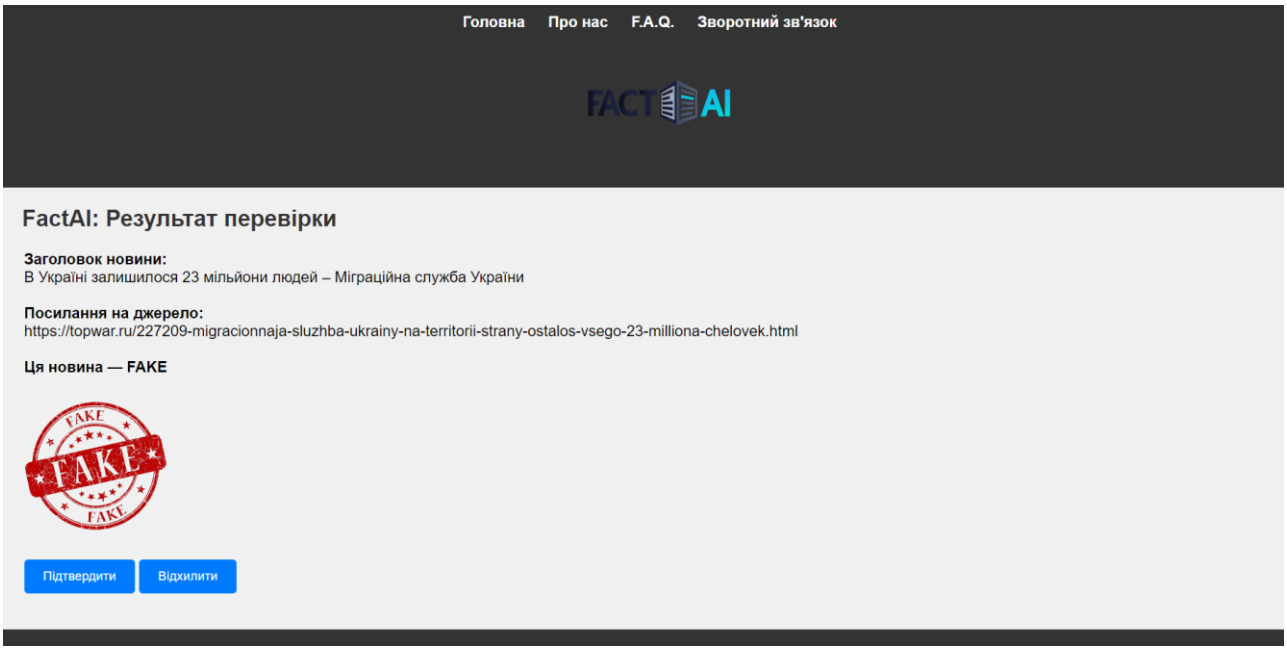


Рисунок 3.12 – Результат перевірки новини 2 запропонованим методом

Результат перевірки новини 2 офіційним каналом Stopfake представлено на рис. 3.13.

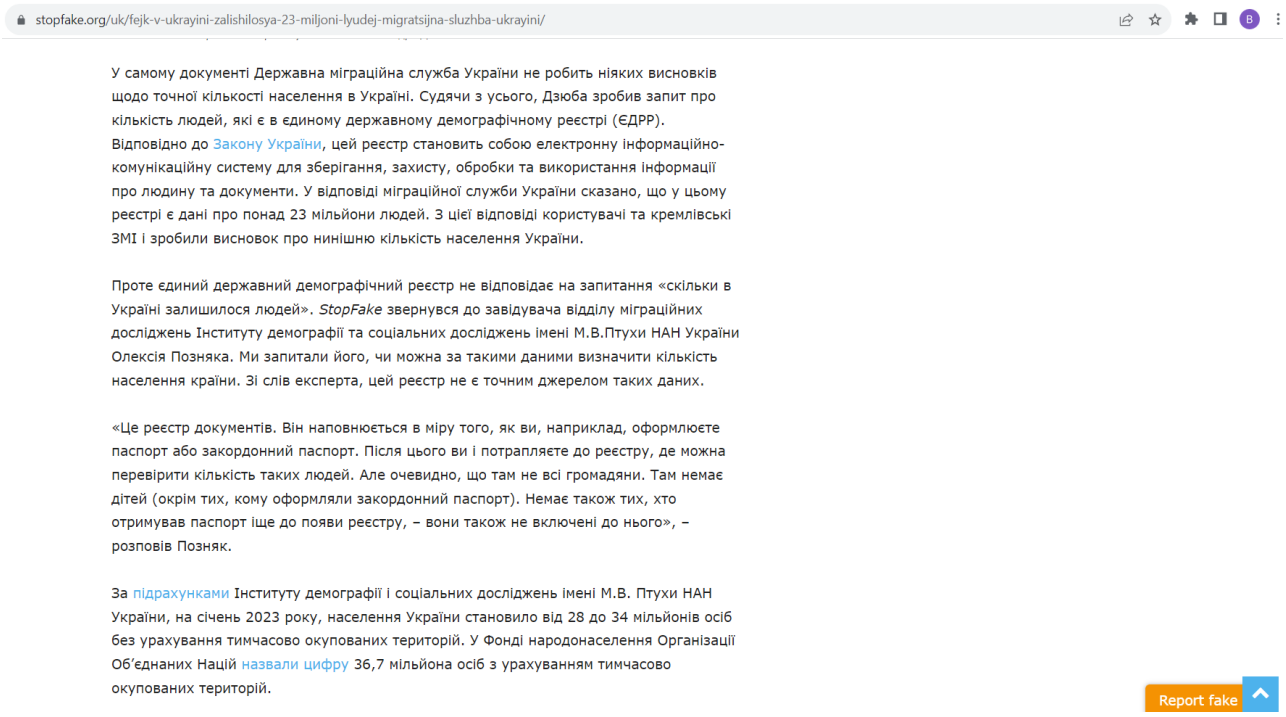


Рисунок 3.13 – Результат перевірки новини 2 офіційним каналом Stopfake

На рис. 3.14 представлено результати перевірки новини 3 запропонованим методом.

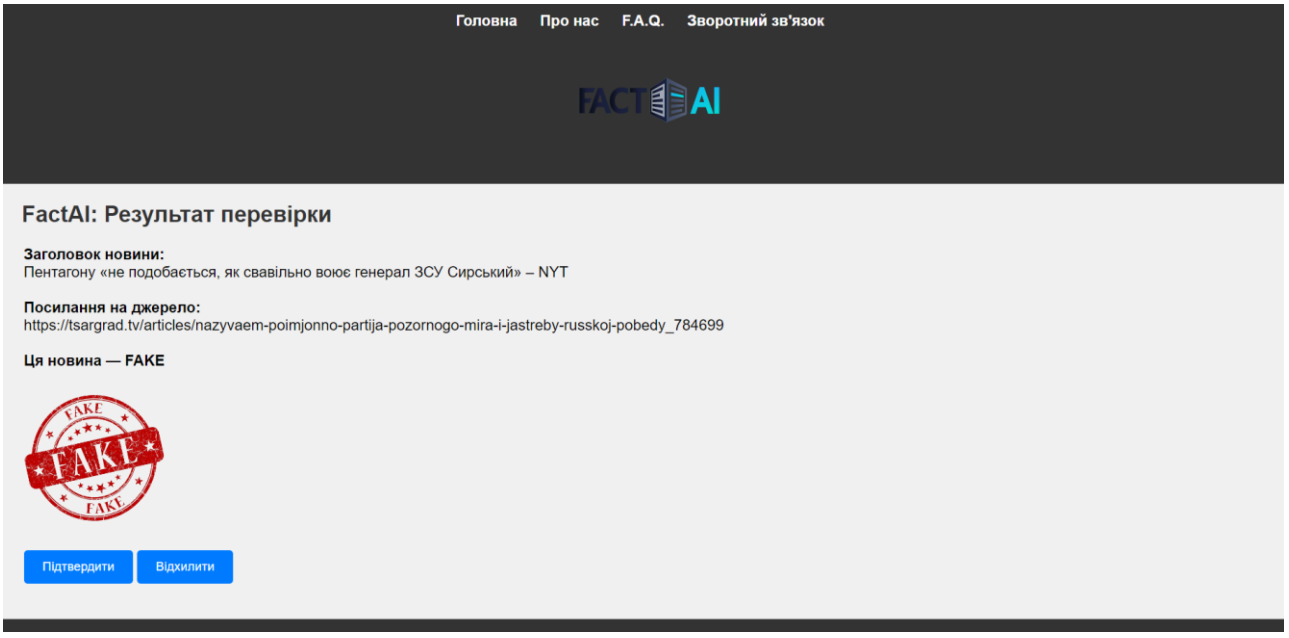


Рисунок 3.14 – Результат перевірки новини 3 запропонованим методом

Результати перевірки новини 3 офіційними джерелами інформації розлянуто на рис. 3.15.

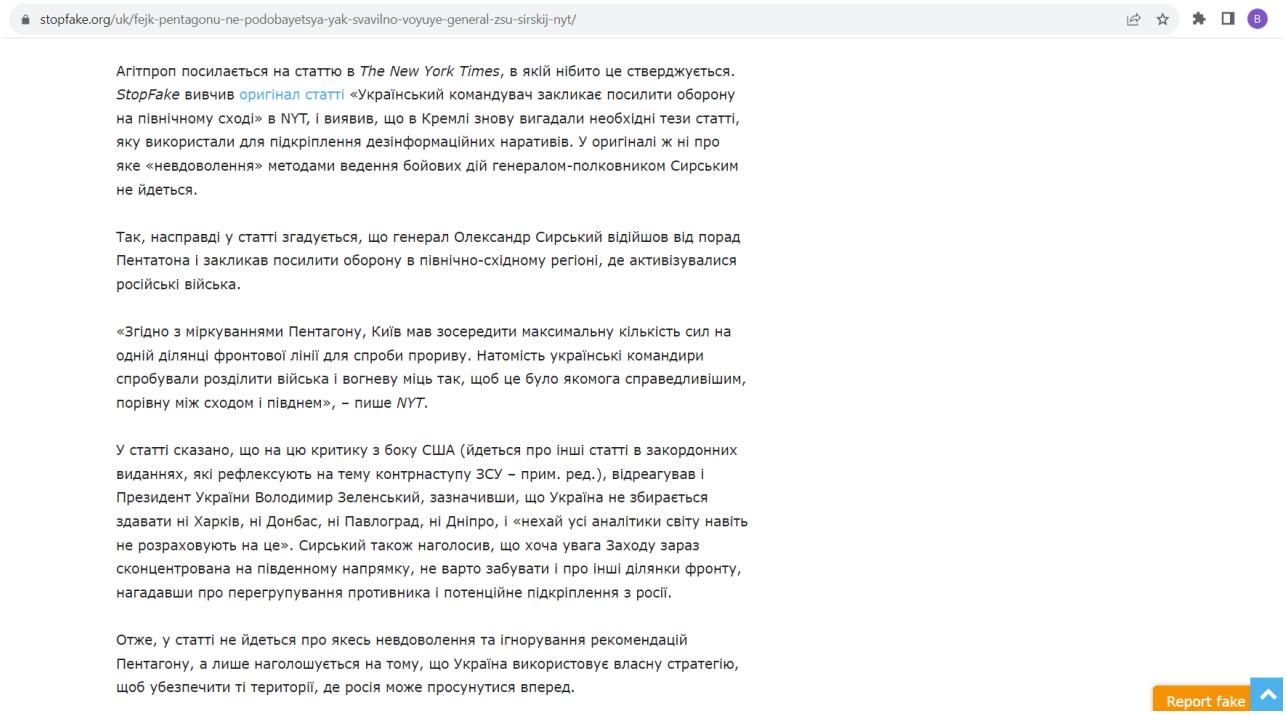


Рисунок 3.15 – Результат перевірки новини 3 офіційним каналом Stopfake

На рис. 3.16 представлено результати перевірки новини 4 запропонованим методом.

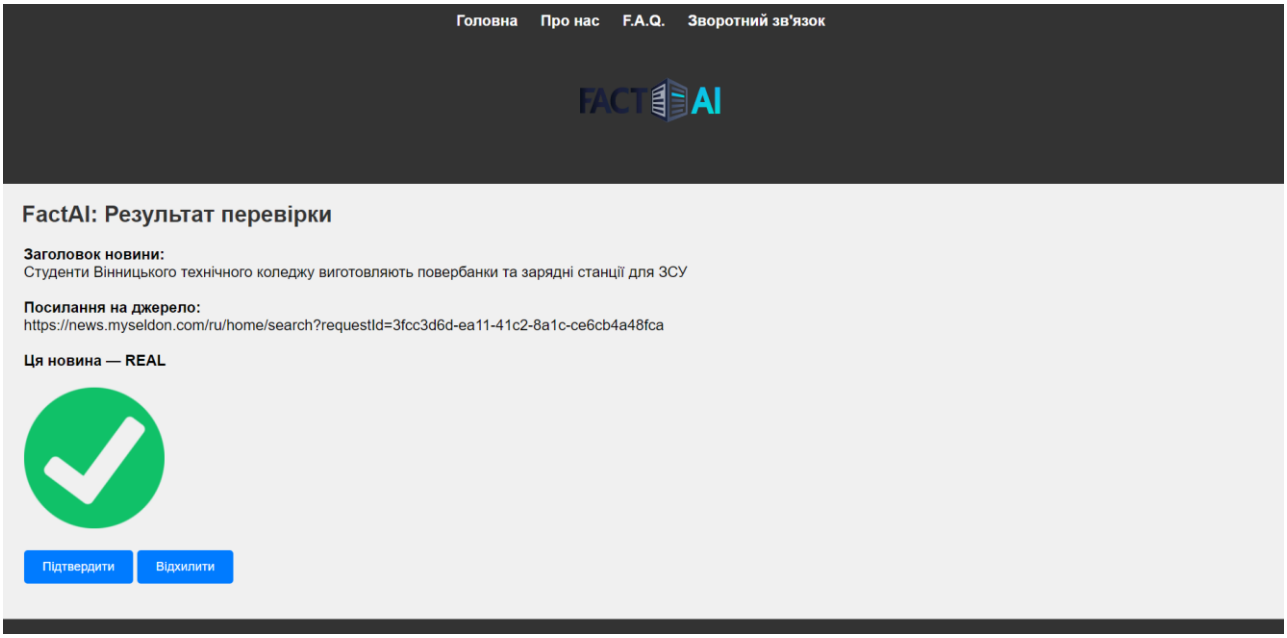


Рисунок 3.16 – Результат перевірки новини 4 запропонованим методом

Результати перевірки новини 4 офіційними джерелами інформації розлянуто на рис. 3.17.

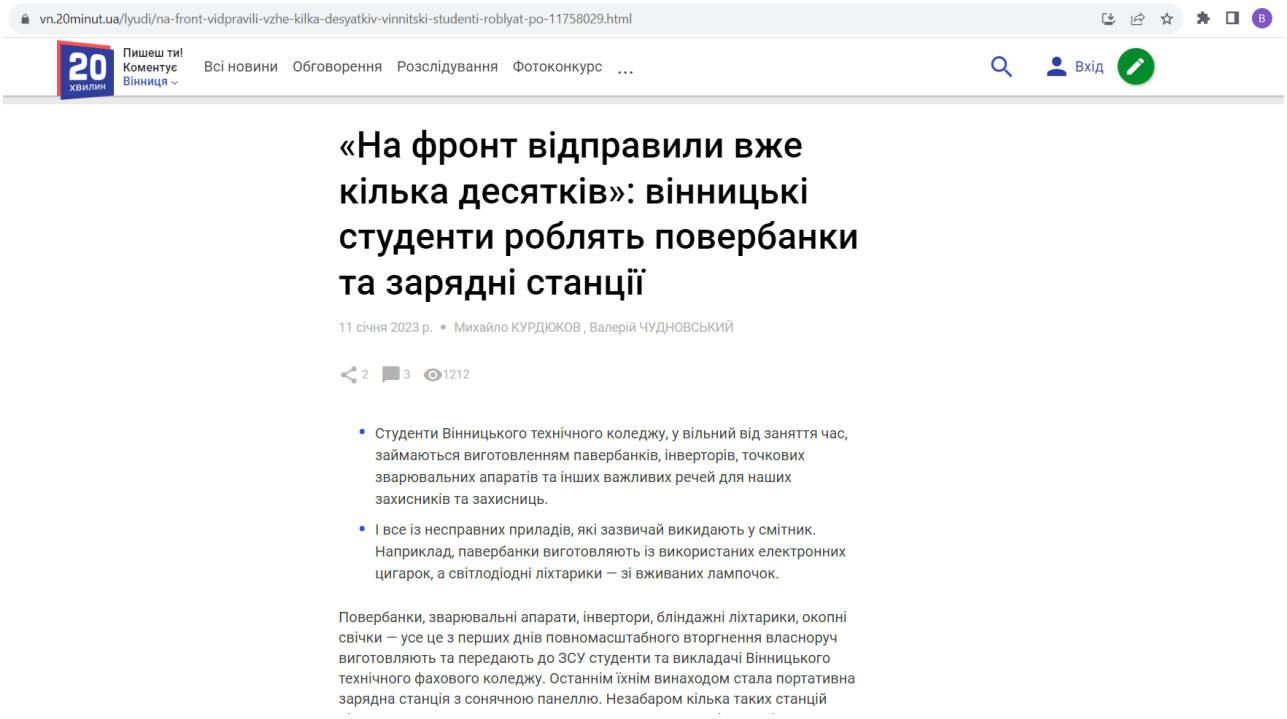


Рисунок 3.17 – Результат перевірки новини 4 офіційним каналом 20 хвилин Віниця

На рис. 3.18 представлено результати перевірки новини 5 запропонованим методом.

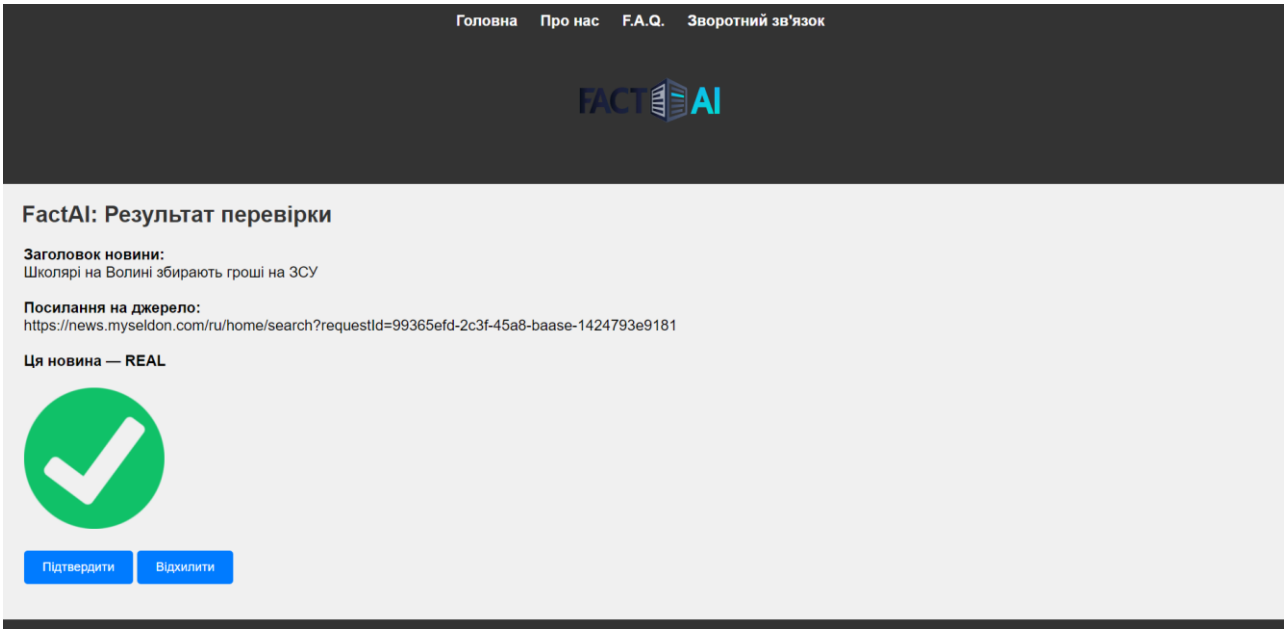


Рисунок 3.18 – Результат перевірки новини 5 запропонованим методом

Результати перевірки новини 5 офіційними джерелами інформації розлянуто на рис. 3.19.

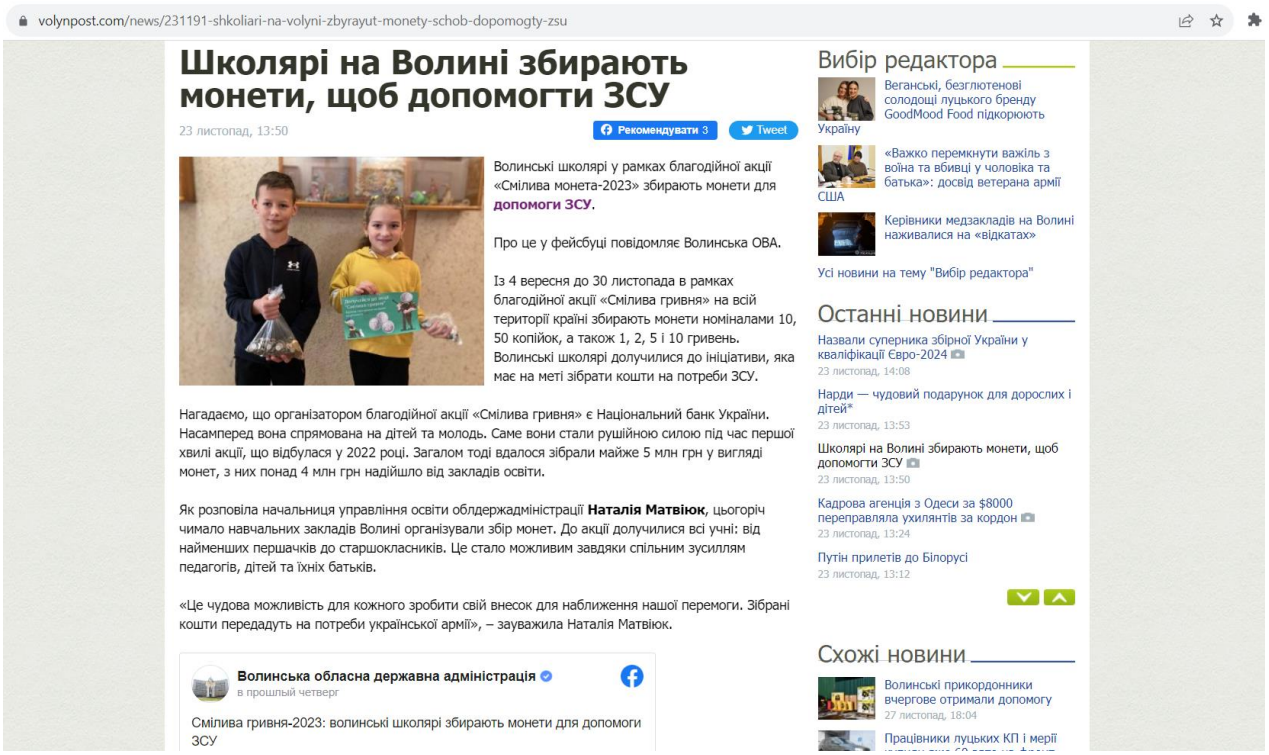


Рисунок 3.19 – Результат перевірки новини 5 офіційним каналом Волинської ОДА

Порівнюючи наведені результати роботи удосконаленого методу із даними, отриманими з офіційних джерел, можна заважити, що запропонований удосконалений метод показав ідентичні результати та високу точність і надійність у виявленні фейкових новин, що доведить адекватність запропонованого підходу.

Було здійснено перевірку розробленого HTML-коду веб-сторінки за допомогою валідатора HTML без будь-яких помилок чи коментарів. Це означає структурну надійність і відповідність коду веб-сторінки, що забезпечує її безперебійну роботу (рис. 3.20).

Рисунок 3.20 – Тестування валідності HTML коду веб-програми

Отже, розроблений HTML-код відповідає офіційним стандартам мови HTML, і його структура та елементи використовуються коректно. Валідний HTML код допомагає забезпечити правильне відображення вмісту на веб-сторінці на різних браузерах і пристроях.

Запропонований метод, що базується на використанні моделі машинного навчання засобами мережі LSTM, продемонстрував спроможність охоплювати тонкі лінгвістичні особливості та шаблони, сприяючи високій продуктивності. Було доведено його здатність забезпечувати точність результату, підтверджуючи свій потенціал як цінного інструменту в триваючій боротьбі з дезінформацією.

3.4 Висновок до розділу 3

У цьому розділі було розглянуто програмну реалізацію розробленого методу виявлення фейкової інформації у соціальних мережах за допомогою штучного інтелекту на основі мережі LSTM.

Обґрунтувано вибір Python як основної мови програмування через її універсальність, великі бібліотеки та сумісність із машинним навчанням. Використання бібліотек Python, зокрема Numpy, Scikit-learn, Tensorflow, Pandas, Matplotlib, NLTK, Seaborn, Gensim і PyStemmer, ще більше спростило

розроблення програмного засобу.

Створення веб-клієнта з використанням Flask для веб-фреймворку та HTML для структурування вмісту в поєднанні з CSS для стилізації проклало шлях для зручного та інтерактивного інтерфейсу. Ця потужна комбінація дозволяє створювати динамічний вміст, створений за допомогою коду Python, забезпечуючи безперебійний стиль і адаптивний інтерфейс користувача.

Було представлено організовану структуру ПЗ, що містить кілька модулів, кожен з яких виконує відповідні функції, зокрема: модуль попереднього оброблення, модуль інтелектуального аналізу, модуль прийняття рішень, модуль бази даних новин і модуль навчання (перенавчання) нейронної мережі.

Було представлено алгоритм поетапної роботи ПЗ, що забезпечує гнучкий робочий процес для оброблення вмісту новин, її аналізу та зберігання результатів для подальшого використання.

Інтерфейс користувача, реалізований як веб-додаток, дозволяє зручне користування ПЗ. Головне вікно має продуманий дизайн, містить основні елементи навігації, форму подачі новин і стислу інформацію про цілі програми. Форма подання структурована таким чином, щоб збирати важливу інформацію про новини без накладення обмежень на символи чи довжину тексту.

Вдалий дизайн інтерфейсу забезпечує доступність і зручність для користувачів, спрощуючи надсилання новин на перевірку та ефективне використання можливостей програми. Веб-програма готова забезпечити безперебійний досвід для користувачів, які зацікавлені в автентифікації вмісту новин, таким чином сприяючи боротьбі з поширенням неправдивої інформації.

Для оцінювання ефективності та перевірки адекватності розробленого методу виявлення фейкових новин було проведено порівняльний аналіз результатів роботи ПЗ та офіційних джерел інформації. Ідентичність отриманих резльтатів дозволила стверджувати адекватність авторського підходу та високу точність і надійність у виявленні фейкових новин.

4 ЕКОНОМІЧНА ЧАСТИНА

4.1 Комерційний та технологічний аудит науково-технічної розробки

У цьому розділі буде проведено комплексний технологічний аудит, щоб оцінити комерційний потенціал розробленого інструменту для виявлення фейкових новин, що є ключовим результатом цієї науково-технічної роботи.

Об’єктом нашого дослідження є кульмінація зусиль, яка включає як розробку, так і програмну реалізацію програми виявлення фейкових новин.

Щоб забезпечити неупереджене оцінювання для проведення комерційного та технологічного аудиту буде залучено комісію щонайменше з трьох незалежних експертів. Оцінювання проводитиметься за надійною п’ятибальною системою на основі 12 заздалегідь визначених критеріїв. Ці критерії охоплюватимуть як наукові, так і технічні аспекти, а також міркування, пов’язані з комерційною життєздатністю.

Після завершення оцінювання ми обчислимо середнє арифметичне балів, виставлених експертами. Це слугуватиме основою для визначення рівня комерційного потенціалу, закладеного в щойно розробленому інструменті. Системний підхід, який поєднує наукові, технічні та комерційні точки зору, дасть цінну інформацію про загальний вплив і доцільність інновацій у виявленні фейкових новин.

Після чого, наведемо результати оцінювання від трьох незалежних експертів за тими ж, заздалегідь визначеними, 12 критеріями, чотирьох альтернативних програм для виявлення фейкової інформації у соцмережах. Це необхідно для того, щоб порівняти комерційний потенціал та відносну якість кожної програми по відношенню до розробленого (удосконаленого) методу виявлення фейкової інформації у соцмережах.

Оцінювання комерційного потенціалу розробки будемо здійснювати за простими критеріями, що дозволить користувачам оцінювати розроблену програму на основі ключових факторів продуктивності та зручності використання (табл. 4.1).

Таблиця 4.1

Критерії оцінювання комерційного та технологічного потенціалу розробки

Критерії	Опис
1.	**Точність класифікації:** Наскільки точно програма ідентифікує фейкову та справжню інформацію? (Шкала: від поганої до відмінної)
2.	**Швидкість обробки:** Як швидко програма аналізує та класифікує інформацію? (Шкала: від повільної до швидкої)
3.	**Зручність використання:** Наскільки легко користувачеві взаємодіяти з програмою? (Масштаб: від складного до інтуїтивно зрозумілого)
4.	**Ефективність використання ресурсів:** Чи ефективно програма використовує системні ресурси? (Шкала: від неефективного до ефективного)
5.	**Адаптивність:** Чи може програма ефективно обробляти різні типи новинного контенту? (Масштаб: від обмеженого до універсального)
6.	**Рівень хибнопозитивних спрацьовувань:** Як часто програма помилково класифікує справжню інформацію як фейкову? (Шкала: від високого до низького)
7.	**Частота хибнонегативних спрацьовувань:** Як часто програма пропускає класифікацію фейкової інформації? (Шкала: від високого до низького)
8.	**Легкість навчання:** Наскільки просто навчити чи оновити програму? (Шкала: від складного до простого)
9.	**Можливість інтеграції:** Наскільки добре програма інтегрується з іншими інструментами чи платформами? (Шкала: від обмеженої до безшовної)
10.	**Механізм зворотного зв'язку:** Чи забезпечує програма чіткий зворотний зв’язок щодо своїх рішень щодо класифікації? (Масштаб: від незрозумілого до прозорого)
11.	**Вимоги до обчислювальних ресурсів:** Скільки обчислювальної потужності вимагає програма? (Шкала: від високого до низького)
12.	**Надійність:** Наскільки добре програма працює за різних умов та вхідних даних? (Шкала: від ненадійної до надійної)

Результати оцінювання незалежними експертами комерційного та технологічного потенціалу розробки наведено у таблиці 4.2.

Таблиця 4.2

Результати оцінювання незалежними експертами комерційного та технологічного потенціалу розробки

Критерії	Бали (1-5)		
	Експерт 1	Експерт 2	Експерт 3
1.	4	5	4
2.	5	4	4
3.	4	5	3
4.	5	4	5
5.	4	3	4
6.	3	4	4
7.	4	3	4
8.	5	4	5
9.	3	4	3
10.	2	3	2

11.	4	5	4
12.	3	4	3

Тепер давайте обчислимо середнє значення для кожного критерію (табл. 4.3).

Таблиця 4.3

Середнє значення для кожного критерію

Критерії	**Середнє значення**
1.	4,33
2.	4,33
3.	4,00
4.	4,67
5.	3,67
6.	3,67
7.	3,67
8.	4,67
9.	3,33
10.	2,33
11.	4,33
12.	3,33

Щоб визначити рівень комерційного потенціалу, розрахуємо загальний середній показник:

$$\textit{Загальний середній показник} = \frac{\textit{Сума середніх значень}}{\textit{Кількість критеріїв}} = \frac{46{,}33}{12} = 3{,}9.$$

Виходячи з цього, загальне середнє значення наближається до 4, що вказує на високий рівень комерційного потенціалу.

Сильні сторони полягають у точності, швидкості обробки та технічних властивостях, тоді як потенційні покращення можуть бути зроблені в механізмах зворотного зв’язку та можливостях інтеграції. Програма демонструє великий потенціал для практичного застосування та успіху на ринку.

У наступних таблицях наведемо результати оцінювання від трьох незалежних експертів за тими ж 12 критеріями, чотирьох альтернативних програм для виявлення фейкової інформації у соцмережах. Серед них дві українські програми – Texty.org.ua і StopFake.org, а також дві закордонні – Snopes і FactCheck.org. Оцінки спрямовані на всебічне порівняння для визначення комерційного потенціалу та відносної якості кожної програми по

відношенню до розробленого методу виявлення фейкової інформації в соціальних мережах за допомогою штучного інтелекту на основі мережі LSTM.

У таблиці 4.4 представлено результати оцінювання трьома незалежними експертами українського програмного засобу Texty.org.ua:

Таблиця 4.4

Результати оцінювання незалежними експертами програми Texty.org.ua

Критерії	**Бали (1-5)**			
	Експерт 1	**Експерт 2**	**Експерт 3**	**Середнє значення**
1.	3	2	2	2,67
2.	4	4	4	2,67
3.	3	3	2	2,67
4.	3	3	2	2,67
5.	3	3	2	2,67
6.	3	3	2	2,67
7.	3	3	2	2,67
8.	3	3	2	2,67
9.	2	2	2	2,67
10.	3	3	2	2,67
11.	4	4	4	2,67
12.	3	3	2	2,67
Середнє значення				2,81

Розрахункова середня оцінка для Texty.org.ua становить приблизно 2,81. Це нижче, ніж середня оцінка нашої програми, яка була близькою до 4, що вказує на те, що наша програма має вищий рівень комерційного потенціалу за цими критеріями.

У таблиці 4.5 представлено результати оцінювання трьома незалежними експертами українського засобу StopFake.org:

Таблиця 4.5

Результати оцінювання програми StopFake.org

Критерії	**Бали (1-5)**			
	Експерт 1	**Експерт 2**	**Експерт 3**	**Середнє значення**
1.	2	2	2	2,33
2.	4	4	4	2,33
3.	2	2	3	2,33
4.	2	2	3	2,33
5.	2	2	3	2,33
6.	3	3	4	3,33

7.	2	2	3	2,33
8.	2	2	3	2,33
9.	2	2	3	2,33
10.	2	2	3	2,33
11.	4	4	4	2,33
12.	2	2	3	2,33
Середнє значення				2,67

Розрахований середній рейтинг для StopFake.org становить 2,67. Це трохи нижче середнього рейтингу для Texty.org.ua, який становив близько 2,81, що вказує на те, що Texty.org.ua має дещо вищий рівень комерційного потенціалу за цими критеріями.

У таблиці 4.6 представлено результати оцінювання трьома незалежними експертами зарубіжного програмного засобу Snopes:

Таблиця 4.6

Результати оцінювання програми Snopes

Критерії	**Бали (1-5)**			
	Експерт 1	**Експерт 2**	**Експерт 3**	**Середнє значення**
1.	3	2	3	2,67
2.	4	4	4	4
3.	3	2	3	2,67
4.	3	3	3	2,67
5.	3	3	3	2,67
6.	4	3	4	3,67
7.	3	2	3	2,67
8.	3	2	3	2,67
9.	2	2	2	2,67
10.	3	2	3	2,67
11.	4	4	4	4
12.	3	2	3	2,67
Середнє значення				2,97

Розрахований середній рейтинг для Snopes становить 2,97, що близько до 3. Це означає, що Snopes є розумною альтернативою, але не значно кращою за українські аналоги, а тим більше за нашу розробку.

У таблиці 4.7 наведено результати оцінювання трьома незалежними експертами іноземного програмного засобу FactCheck.org:

Таблиця 4.7

Результати оцінювання програми FactCheck.org

Критерії	**Бали (1-5)**			
	Експерт 1	**Експерт 2**	**Експерт 3**	**Середнє значення**
1.	2	3	3	2,67
2.	4	4	4	4
3.	4	3	4	3,67
4.	3	4	3	3,33
5.	3	3	3	3
6.	4	3	4	3,67
7.	3	4	3	3,33
8.	4	3	4	3,67
9.	2	2	2	2
10.	2	2	2	2
11.	4	4	4	4
12.	4	3	4	3,67
Середнє значення				3,25

Розрахований середній рейтинг для FactCheck.org становить 3,25.

Отже, після оцінювання та порівняння трьома незалежними експертами чотирьох альтернативних програм для виявлення фейкової інформації в соціальних мережах за 12-ма заздалегідь визначеними критеріями ми дійшли таких висновків:

Texty.org.ua: хоч і отримав досить хорошу оцінку, не досягнув рівня нашої розробленої програми. Він демонструє силу з точки зору простоти використання та легкості навчання, але його загальна продуктивність відстає від нашого вдосконаленого рішення на основі ШІ.

StopFake.org: подібно до Texty.org.ua, показав порівнянний рейтинг, який все одно був нижчим за продуктивність нашого рішення.

Snopes: іноземна програма, мала середній показник продуктивності, близький до 3, що вказує на можливість покращення за різними критеріями. Його оцінка значно нижча, ніж наша розроблена система на основі мережі LSTM.

FactCheck.org: інша іноземна програма, отримала середній рейтинг, близький до 3,25. Цей засіб також відстає від продуктивності нашого рішення в багатьох аспектах.

Виходячи з колективної оцінки, очевидно, що наш вдосконалений метод

виявлення фейкової інформації в соціальних мережах засобами ШІ на основі мережі LSTM пропонує конкурентну перевагу. Наша програма демонструє вищу середню оцінку за всіма критеріями, що відображає її підвищену точність, швидкість обробки, використання ресурсів, адаптивність і надійність. Наше рішення демонструє особливу силу в точності, швидкості обробки та технічних характеристиках, перевершуючи альтернативні варіанти.

Представлене рішення (удосконалений метод) пропонує більш надійний та ефективний підхід до виявлення фейкової інформації, перевершуючи як українські, так і іноземні альтернативи.

4.2 Прогнозування витрат на виконання науково-дослідної (дослідно-конструкторської) роботи

Прогнозування витрат на виконання науково-дослідної роботи є критично важливим кроком у будь-якому комплексному проєкті. Цей процес фінансового планування має важливе значення для забезпечення ефективного виконання проєкту та ретельного врахування всіх фінансових аспектів. Щоб полегшити цей процес, ми структурували прогнозування на три окремі етапи. На першому етапі ми заглиблюємося в оцінювання витрат, безпосередньо пов’язаних із робочою силою, залученою до проєкту. Другий етап розширює обсяг, щоб охопити загальні витрати проєкту, від інфраструктури та технологій до збирання та розробки даних. Нарешті, на третьому етапі враховуються витрати на впровадження результатів дослідження та введення в дію результатів проєкту. Разом ці етапи пропонують систематичний підхід до складання бюджету та фінансового менеджменту для дослідницьких та проектних починань.

Основну зарплату розробника розрахуємо за формулою (4.1):

$$З_о = \frac{М}{Т_р} \cdot t, \qquad (4.1)$$

де $М$ – місячна зарплата розробника;

$T_р$ – кількість робочих днів у місяці, $T_р$ = 21 день;

t – кількість днів роботи.

Місячна зарплата складає 20000 грн., t = 60 днів.

Тоді, базовий оклад розробника становить:

$$З_р = \frac{20000}{21} \cdot 60 = 57142{,}86 \text{ грн.}$$

Розрахуємо основну зарплату для керівника (при місячному посадовому окладі становить 25000 грн.):

$$З_к = \frac{25000}{21} \cdot 60 = 71428{,}57 \text{ грн.}$$

Розрахуємо витрати на оплату праці:

$$З_о = З_р + З_к = 57142{,}86 + 71428{,}57 = 128571{,}43 \text{ грн.}$$

Додаткова заробітна плата розраховується у розмірі 10% від основної зарплати розробника та керівника [43]:

$$З_{дод} = З_о \cdot 0{,}1 = 128571{,}43 \cdot 0{,}1 = 12857{,}143 \text{ грн.}$$

Нарахування на зарплату складають 22% від сум базової та додаткової грошової оплати (формула 4.2):

$$З_н = (З_о + З_р + З_{дод}) \cdot 0{,}22 \quad (4.2)$$

Розрахуємо нарахування на оклад розробника та керівника:

$$З_н = (128571{,}43 + 12857{,}143) \cdot 0{,}22 = 31114{,}28 \text{ грн.}$$

Розрахуємо відрахування на амортизацію устаткування (ПК) за формулою (4.3) [57]:

$$A = \frac{Ц \cdot Н_а}{100} \cdot \frac{Т}{12}, \quad (4.3)$$

де $Ц$ – вартість устаткування, $Ц$ = 15000 грн.;

$Н_а$ – річна норма амортизації, $Н_а$ = 20%;

$Т$ – термін використання устаткування, $Т$ = 4 місяці.

$$A = \left(\frac{15000 \cdot 20}{100} \cdot \frac{4}{12}\right) = 1000 \text{ грн.}$$

Також, використовувалося безкоштовне програмне забезпечення – ОС Linux з наявними інструментами.

При розробці були використані матеріали, кількість та вартість яких зведені в таблицю 4.8.

Таблиця 4.8

Матеріали, використані для реалізації проєкту

Найменування матеріалу, марка, тип, сорт	Ціна, грн.	Витрачено	Вартість витрачених матеріалів, грн.
Папір А4	200	3	600
Флеш USB	200	1	200
Папка для паперів	50	2	100
Файли	20	1	20
Ручка	30	4	120
Всього:			1040

Отже, вартість витрачених матеріалів, М = 1040 грн.

Розрахуємо тариф на електроенергію за формулою (4.4):

$$Ц_е = \left(Ц_{опт} + Ц_{розп} + Ц_{пост}\right) \cdot \left(1 + \frac{ПДВ}{100\%}\right), \quad (4.4)$$

$Ц_{опт}$ – середня оптова ціна електроенергії, яка визначається оператором ринку (без ПДВ), грн за 1 кВт·год;

$Ц_{розп}$ – вартість розподілу електроенергії окремою енергорозподільною компанією (без ПДВ), грн за 1 кВт·год;

$Ц_{пост}$ – вартість постачання електроенергії від енергорозподільної компанії до конкретного споживача (без ПДВ), грн за 1 кВт·год.

$Ц_е = (3{,}86 + 1{,}34 + 0{,}38) \cdot 1{,}2 = 6{,}7$ грн за 1 кВт·год.

Розрахуємо витрати на електроенергію за формулою (4.5):

$$B_e = \sum_{i=1}^{n} \frac{W_{yi} \cdot t_i \cdot Ц_е \cdot К_{впі}}{\eta_i}, \quad (4.5)$$

W_{yi} – встановлена потужність обладнання на певному етапі розробки, W_{yi} = 0,3 кВт;

t_i – фактична кількість годин роботи ПК при обробці завдань проєкту на місяць, $t_i = 7$ *год* $\cdot 60$ *днів* $= 420$ (год.) на місяць;

$Ц_е$ – вартість 1 кВт-години електроенергії, грн;

$K_{вп і}$ – коефіцієнт використання потужності, $K_{вп і}$ = 0,72;

η_i – коефіцієнт корисної дії обладнання, η_i = 0,9.

$$B_e = \frac{0{,}3 \cdot 420 \cdot 6{,}7 \cdot 0{,}72}{0{,}9} = 675{,}36 \text{ грн.}$$

Інші затрати приймаються у розмірі 50-100% від базового окладу дослідника / розробника. Використаємо значення в розмірі 100%, тоді:

$$B_{ін} = 1{,}0 \cdot 128571{,}43 = 128571{,}43 \text{ грн.}$$

Розрахуємо загальну суму затрат на впровадження ПЗ за формулою (4.6):

$$B_{заг} = З_о + З_р + З_д + З_н + М + К_в + B_{спец} + B_{прг} + А + B_е + B_{св} + B_{сп} + B_{ін} + B_{нзв} \quad (4.6)$$

Отже, загальна сума витрат дорівнює:

$B_{заг}$ = 128571,43 + 57142,86 + 12857,143 + 31114,28 + 1040 + 1000 + 675,36 + 128571,43 = 360972,503 грн.

Зважаючи на те, що у проєкті не використовувалися комплектуючі вироби ($К_в$), спецустаткування ($B_{спец}$); не було витрат: на ПЗ для наукових робіт ($B_{прг}$), службових відряджень ($B_{св}$), на роботи сторонніх підприємств ($B_{сп}$) та загальновиробничих витрат ($B_{нзв}$) – ми не можемо додати перелічені показники до статті витрат.

Розрахуємо загальні витрати на завершення науково-технічної роботи за формулою (4.7):

$$ЗВ = \frac{B_{заг}}{\beta} \quad (4.7)$$

де β – коефіцієнт, який характеризує стадію виконання проєкту. У нашому випадку розробка знаходиться на стадії дослідного зразка, тому β = 0,5.

$$ЗВ = \frac{360972{,}503}{0{,}5} = 721945{,}006 \text{ грн.}$$

4.3 Розрахунок економічної ефективності науково-технічної розробки

Прогнозування комерційних ефектів від впровадження розробленої програми виявлення фейкових новин у соціальних мережах за допомогою штучного інтелекту на основі мережі LSTM передбачає оцінювання потенційного впливу на різні аспекти комерційного середовища. Це оцінювання має на меті зрозуміти, як програма може створювати економічну цінність, покращувати процеси прийняття рішень і вирішувати конкретні бізнес-завдання.

У нашому випадку прогнозований комерційний ефект від реалізації результатів проєкту буде розрахований наступним чином (4.8):

$$\Delta П_i = (\pm \Delta Ц_0 \cdot N + Ц_0 \cdot \Delta N)_i \cdot \lambda \cdot \rho \cdot \left(1 - \frac{\vartheta}{100}\right), \quad (4.8)$$

де $\pm\Delta Ц_о$ – зміна вартості програмного засобу у результаті впровадження науково-технічної розробки;

N – кількість користувачів, що користувалися альтернативним програмним засобом до моменту інтеграції результатів нової науково-технічної розробки;

$Ц_о$ – основна метрика ефективності, яка вказує на продуктивність підприємства за певний рік після інтеграції результатів досліджень і розробок. $Ц_о = Ц_б \pm \Delta Ц_о$;

$Ц_б$ – вартість програмного засобу за певний рік до моменту інтеграції результатів досліджень і розробок;

ΔN – зростання бази користувачів програмного засобу протягом досліджуваних інтервалів часу завдяки вдосконаленню окремих атрибутів продукту;

λ – коефіцієнт, що включає оплату податку на додану вартість. Ставка податку на додану вартість становить 20%, а коефіцієнт (λ) = 0,8333.

ρ – коефіцієнт, що враховує рентабельність засобу;

ϑ – ставка податку на прибуток, у нашому випадку, ϑ =18%.

Розглянемо сценарій, коли прогнозована ціна одиниці товару становить 4000 грн., а період збільшення прибутку – 3 роки. Після завершення розробки та доопрацювань вартість може бути збільшена на 500 грн. Крім того, збільшиться кількість одиниць реалізованої продукції: на 4000 одиниць у перший рік, на 3000 одиниць у другий рік і на 2000 одиниць у третій рік. До інтеграції результатів досліджень і розробок програмний засіб не продавався.

Перейдемо до розрахунків сценарію щодо впровадження реалізації розробки:

$\Delta П_1$ = (500 · 4000 + (4000 + 500) · 4000) · 0,8333 · 0,25 · (1 – 0,18) = 3416530 грн.

$\Delta П_2$ = (500 · 4000 + (4000 + 500) · (4000 + 3000)) · 0,8333 · 0,25 · (1 – 0,18) = 5722687,75 грн.

$\Delta П_3$ = (500 · 4000 + (4000 + 500) · (4000 + 3000 + 2000)) · 0,8333 · 0,25 · (1 – 0,18) = 7260126,25 грн.

Прогнозований комерційний ефект від інтеграції результатів дослідження та розробки за три роки становить 16399344 грн.

4.4 Розрахунок ефективності вкладених інвестицій та періоду їх окупності

Визначимо поточну вартість загального приросту чистого прибутку (ПП) в результаті потенційної інтеграції та комерціалізації науково-технічного розробки, на отримання якого може розраховувати інвестор за формулою (4.9):

$$ПП = \sum_{1}^{T} \left(\frac{\Delta П_i}{(1+\tau)^t} \right), \quad (4.9)$$

де $\Delta П_i$ – приріст чистого прибутку за кожний із років, у яких наявні результати проведених та виконаних досліджень та розробок, в грн;

T – тривалість періоду, в момент якого стають очевидними результати

інтегрованих досліджень та розробок, вимірюється роками.

τ – ставка дисконту, яка враховує річний прогнозований рівень інфляції в країні, τ = 0,1;

t – період часу, у роках.

Також, слід зазначити, що зростання прибутку ми будемо отримувати з першого року:

$$\text{ПП} = \left(\frac{3416530}{(1+0{,}1)^1}\right) + \left(\frac{5722687{,}75}{(1+0{,}1)^2}\right) + \left(\frac{7260126{,}25}{(1+0{,}1)^3}\right) = 13290070{,}23 \text{ грн.}$$

Тепер, розрахуємо початкову суму інвестицій (PV), яку потенційний інвестор має спрямувати на реалізацію та комерціалізацію науково-технічної розробки, за допомогою наступної формули (4.10):

$$PV = k_{інв} \cdot ЗВ, \tag{4.10}$$

де $k_{інв}$ – коефіцієнт, що включає витрати інвестора на інтеграцію науково-технічного засобу та його комерціалізацію, враховує такі витрати, як підготовка об’єкта, розроблення програми, навчання персоналу, маркетингова діяльність. Як правило, $k_{інв}$ знаходиться в діапазоні від 2 до 5;

ЗВ – загальна сума витрат на здійснення досліджень і розробок та оформлення їх результатів, грн.

$$PV = 2 \cdot 721945{,}006 = 1443890{,}012 \text{ грн.}$$

За формулою (4.11), визначимо абсолютний економічний ефект, який позначається як $E_{абс}$ або чистий поточний дохід (NPV), який є результатом потенційного впровадження та комерціалізації досліджень і розробок:

$$E_{абс} = ПП - PV, \tag{4.11}$$

$$E_{абс} = 13290070,23 – 1443890,012 = 11846180,218 \text{ грн.}$$

Після інтеграції нашої розробки, $E_{abc} > 0$, це свідчить про те, що впровадження нашого проєкту призведе до позитивного чистого прибутку або економічної вигоди. Це свідчить про те, що наша розробка позитивно вплине на проєкт або підприємство, зробивши його фінансово життєздатнішим і потенційно прибутковішим, ніж це було раніше.

Отже, інвестування коштів у проєкт може бути доцільним.

Для прийняття обґрунтованих рішень щодо інвестицій у дослідження і розробку вкрай важливо оцінити їх відносну (річну) ефективність. Ця оцінка допомагає нам зрозуміти економічну життєздатність цих інвестицій з часом. Відносна ефективність дає зрозуміти річний прибуток або вигоди, які можна очікувати від інвестованого капіталу. Аналізуючи цю відносну ефективність, зацікавлені сторони можуть краще оцінити довгострокові фінансові перспективи ініціатив наукового розвитку та прийняти обґрунтовані інвестиційні рішення. Тому, використовуючи формулу (4.12), розрахуємо відносну (річну) ефективність і порівняємо її з дисконтною ставкою:

$$\text{Е}_\text{в} = \sqrt[\text{Т}_\text{ж}]{1 + \frac{\text{Е}_\text{абс}}{PV}} - 1, \quad (4.12)$$

де $T_{ж}$ – життєвий цикл наукової розробки, у роках.

$$E_{в} = \sqrt[3]{1 + \frac{11846180,218}{1443890,012}} - 1 = 2,01.$$

Розрахуємо мінімальну ставку дисконту, за формулою (4.13):

$$\tau_{min} = d + f, \quad (4.13)$$

де d – середньозважена процентна ставка за депозитними операціями в комерційних банках, d = 0,14;

f – ризикованість вкладень, $f = 0{,}05$.

$$\tau_{min} = 0{,}14 + 0{,}05 = 0{,}19.$$

Як бачимо, $E_в > \tau_{min}$, а це означає, що інвестор може бути зацікавлений у фінансуванні нашого проєкту.

Тепер, обчислимо термін окупності коштів, інвестованих у наший проєкт, за формулою (4.14):

$$T_{ок} = \frac{1}{\text{Е}_\text{в}}, \tag{4.14}$$

$$T_{ок} = \frac{1}{2{,}01} = 0{,}49 \text{ р.}$$

За нашими розрахунками, $T_{ок}$ < 3-х років, а це вказує на те, що проєкт, швидше за все, генеруватиме позитивні грошові потоки відносно швидко, що робить його вигідною інвестицією з точки зору його здатності окупити початкові витрати протягом короткого періоду часу.

4.5 Висновок до розділу 4

У цьому розділі надано комплексну економічну оцінку науково-технічної розробки, спрямованої на вдосконалення методу виявлення фейкової інформації в соціальних мережах за допомогою штучного інтелекту на основі мережі LSTM. Основні висновки та ключові моменти з кожного підрозділу:

Підрозділ 4.1 - Комерційно-технологічний аудит:

У цьому підрозділі було проведено ретельний технологічний аудит для оцінювання комерційного потенціалу розробленого інструменту виявлення фейкових новин. Три незалежні експерти оцінили програму, в результаті чого загальна середня оцінка була близькою до 4, що свідчить про високий рівень комерційного потенціалу. Сильні сторони програми включають точність, швидкість обробки та технічні властивості, з можливостями для вдосконалення

механізмів зворотного зв’язку та можливостей інтеграції. Програма демонструє значний потенціал для практичного застосування та успіху на ринку. Також були представлені результати оцінювання та порівняння трьома незалежними експертами чотирьох альтернативних програм для виявлення фейкової інформації в соціальних мережах. Виходячи з отриманих результатів, очевидно, що наш вдосконалений метод виявлення фейкової інформації в соціальних мережах засобами ШІ на основі мережі LSTM пропонує конкурентну перевагу. Наша програма демонструє вищу середню оцінку за всіма критеріями, що відображає її підвищену точність, швидкість обробки, використання ресурсів, адаптивність і надійність. Наше рішення демонструє особливу силу в точності, швидкості обробки та технічних характеристиках, перевершуючи альтернативні варіанти. Представлене рішення (удосконалений метод) пропонує більш надійний та ефективний підхід до виявлення фейкової інформації, перевершуючи як українські, так і іноземні альтернативи.

Підрозділ 4.2 - Прогнозування витрат:

У цьому підрозділі представлено прогнозування витрат на проведення наукових досліджень, у тому числі дослідно-конструкторських робіт. Процес було поділено на три етапи, починаючи з оцінки витрат, пов’язаних з робочою силою, з подальшою оцінкою загальних витрат на проєкт і закінчуючи розрахунком витрат на впровадження. Загальні витрати на розробку програмного продукту визначено у сумі 721945,006 грн.

Підрозділ 4.3 - Розрахунок економічної ефективності:

У цьому розділі проводилося економічне оцінювання комерційного ефекту науково-технічної розробки. У сценарії розглядалася ціна за одиницю товару 4000 грн, з трирічним періодом зростання прибутку. Аналіз врахував підвищення ціни на 500 грн і збільшення кількості проданих одиниць. Прогнозований комерційний ефект за три роки склав 16399344 грн.

Підрозділ 4.4 - Розрахунок інвестиційної ефективності та терміну окупності:

У цьому підрозділі були розраховані різні фінансові показники. Визначено

поточну вартість загального приросту чистого прибутку, суму початкових інвестицій (PV) та абсолютний економічний ефект. Результат показав, що $E_{abc} > 0$, що означає позитивний чистий прибуток або економічну вигоду після реалізації проєкту. Це означає, що інтеграція розробки позитивно вплине на проєкт або підприємство, підвищивши його фінансову життєздатність і потенційну прибутковість. Відносна (річна) ефективність виявилася більшою за мінімально необхідну норму прибутку (τ_{min}), що вказує на потенційний інтерес інвестора. Термін окупності інвестицій у науковий проєкт був розрахований у 0,49 року, що свідчить про відносно швидке повернення інвестицій і робить його перспективним фінансовим підприємством.

Таким чином, економічна оцінка удосконаленого методу виявлення фейкових новин у соціальних мережах за допомогою штучного інтелекту на основі мережі LSTM демонструє її значний потенціал для комерційного успіху та позитивних фінансових результатів. Результати підтверджують економічну життєздатність проєкту та його потенціал для підвищення фінансових показників, що робить покращену розробку перспективним заходом для потенційних інвесторів та зацікавлених сторін.

ВИСНОВКИ

У результаті виконання завдань магістерської роботи було досягнено поставленої мети щодо вдосконалення методу виявлення фейкової інформації у соціальних мережах. Це було реалізовано завдяки впровадженню інструментів штучного інтелекту на основі мережі довгострокової короткочасної пам’яті (LSTM). Підхід на основі LSTM, як продемонстровано в цій роботі, довів свою ефективність, забезпечуючи точні результати та високу продуктивність.

Було досліджено багатобічний характер проблеми фейкових новин, підкресливши важливість комплексного підходу, який поєднує людську пильність із технологічними інноваціями. Розроблення та оцінювання різних моделей машинного навчання зрештою визначила мережу LSTM як ефективного інструменту для виявлення фейкових новин із винятковою точністю 89% і мінімальними помилками типу 1 і 2.

Здійснено програмну реалізацію запропонованого методу. Розроблений ПЗ є динамічним рішенням, що постійно навчається, завдяки здатності оновлювати базу даних новин. Ця ключова функція гарантує, що система може підтримувати стабільно високий рівень точності класифікації, таким чином зменшуючи вплив нових підходів, які використовуються для створення фейкових новин.

Отже, практичним результатом цього дослідження є надійний веб-додаток для виявлення фейкових новин, оснащений зручним інтерфейсом, який відповідає принципам точності та адаптивного дизайну. Інструмент дозволяє користувачам відправляти новини на перевірку та сприяє зростанню бази даних новин.

Здійснено перевірку доцільності розроблення та використання запропонованого підходу. Результати комерційного аудиту та прогнозування витрат підтверджують потужний комерційний потенціал ПЗ, підкреслюючи позитивний чистий прибуток, швидке повернення інвестицій і значну прибутковість.

Під час комерційного аудиту три незалежні експерти оцінили ПЗ та поставили загальний середній бал, близький до 4, що свідчить про його значний

комерційний потенціал.

Удосконалений метод має переваги в ключових сферах, таких як точність, швидкість оброблення, а також в аспекті вдосконалення механізмів зворотного зв’язку та можливостей інтеграції. Ці переваги підкреслюють позивні перспективи практичного застосування запропонованого методу, представленого його програмною ралізацією.

Крім того, було подано оцінку та порівняння запропонованого методу з використанням ШІ на основі LSTM із чотирма альтернативними програмами для виявлення фейкової інформації в соціальних мережах. Результати показують, що запропонований ПЗ має оцінку вище середнього балу за всіма критеріями, підкреслюючи її високу точність, швидкість оброблення, використання ресурсів, адаптивність і надійність. Аворський підхід вирізняється своєю ефективністю та пропонує більш надійні та продуктивні засоби виявлення фейкової інформації.

Представлений удосконалений метод виявлення фейкової інформації в соціальних мережах за допомогою штучного інтелекту на основі мережі LSTM є потужним інструментом у боротьбі з дезінформацією. Його здатність до постійного навчання в поєднанні з високою продуктивністю позиціонує його як ефективний інструмент для перевірки достовірності інформації у цифровій сфері.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Кіца М. Я. Особливості та методи виявлення фейкової інформації в українських ЗМІ. *Вісник Національного університету «Львівська політехніка». Серія: Журналістські науки*, 2017, 883: 1. URL: https://science.lpnu.ua/sites/default/files/journal-paper/2019/apr/16109/kitsa.pdf (дата звернення: 02.10.2023).

2. Снитюк Н. Система виявлення фейкової інформації в мережі інтернет. *Scientific Collection «InterConf»*, 2022, 126: 201-207. URL: https://archive.interconf.center/index.php/conference-proceeding/article/view/1373 (дата звернення: 02.10.2023).

3. Бомчук Д. В. *Фейкова інформація у соціальних медіа: сутність, виявлення, протидія*. 2023. PhD Thesis. URL: http://ephsheir.uhsp.edu.ua/handle/8989898989/7146 (дата звернення: 03.10.2023).

4. Корж О., Коровай, В. Фейковий контент: види, ознаки, шляхи виявлення. *Освіта. Інноватика. Практика*, 2023, 11.7: 37-42. URL: https://oip-journal.org/index.php/oip/article/view/216/152 (дата звернення: 03.10.2023).

5. Тищенко В., Мужанова Т. Дезінформація і фейкові новини: ознаки та методи виявлення в мережі інтернет. *Електронне фахове наукове видання «Кібербезпека: освіта, наука, техніка»*, 2022, 2.18: 175-186. URL: https://www.csecurity.kubg.edu.ua/index.php/journal/article/view/413/341 (дата звернення: 04.10.2023).

6. Наконечний В., Барабаш О., Лаптєва Т., Міщенко А. Удосконалення методу виявлення та кластерізації джерел неправдивої інформації. *Science-based technologies*, 2022, 54.2: 105-111. URL: https://jrnl.nau.edu.ua/index.php/SBT/article/view/16747/24063 (дата звернення: 04.10.2023).

7. Лукова-Чуйко Н., Лаптєва Т. Удосконалення методу виявлення неправдивої інформації за допомогою байесовского класифікатора. *Безпека інформації*, 2022, 28.3: 119-126. URL: https://jrnl.nau.edu.ua/index.php/Infosecurity/article/view/17368/24692 (дата

звернення: 05.10.2023).

8. Ошийко Я. Р. *Математичне та програмне забезпечення виявлення елементів дезінформації в потоках текстових даних*. 2020. Master's Thesis. КПІ ім. Ігоря Сікорського. URL: https://ela.kpi.ua/bitstream/123456789/39711/1/Oshyiko_magistr.pdf (дата звернення: 02.10.2023).

9. Шульська Н. М., Зінчук Р. С., Навальна М. І. Антифейкові ресурси сучасного медіапростору України в умовах інформаційної війни. *Вчені записки Таврійського національного університету імені ВІ Вернадського. Серія: Філологія. Журналістика*, 2022, 268-273. URL: https://philol.vernadskyjournals.in.ua/journals/2022/4_2022/part_2/44.pdf (дата звернення: 05.10.2023).

10. Ковбасюк О. М., Грицюк Ю. І. Виявлення фейкових новин методами машинного навчання. *Редакційна колегія*, 2023, 28. URL: https://dspace.lvduvs.edu.ua/bitstream/1234567890/5241/1/16_12_2022.pdf#page=28 (дата звернення: 05.10.2023).

11. Guo Bin. The future of false information detection on social media: New perspectives and trends. *ACM Computing Surveys (CSUR)*, 2020, 53.4: 1-36. URL: https://dl.acm.org/doi/abs/10.1145/3393880 (дата звернення: 07.10.2023).

12. Habib A. False information detection in online content and its role in decision making: a systematic literature review. *Social Network Analysis and Mining*, 2019, 9: 1-20. URL: https://link.springer.com/article/10.1007/s13278-019-0595-5 (дата звернення: 07.10.2023).

13. Seyam Asmaa. Deep Learning Models to Detect Online False Information: a Systematic Literature Review. In: *The 7th Annual International Conference on Arab Women in Computing in conjunction with the 2nd Forum of Women in Research*. 2021. p. 1-5. URL: https://dl.acm.org/doi/abs/10.1145/3485557.3485580 (дата звернення: 08.10.2023).

14. Ghanem B., Rosso P., Rangel F. An emotional analysis of false information in social media and news articles. *ACM Transactions on Internet Technology (TOIT)*,

2020, 20.2: 1-18. URL: https://dl.acm.org/doi/abs/10.1145/3381750 (дата звернення: 09.10.2023).

15. Wu L. False information detection on social media via a hybrid deep model. In: *Social Informatics: 10th International Conference, SocInfo 2018, St. Petersburg, Russia, September 25-28, 2018, Proceedings, Part II 10*. Springer International Publishing, 2018. p. 323-333. URL: https://link.springer.com/chapter/10.1007/978-3-030-01159-8_31 (дата звернення: 10.10.2023).

16. Pierri F., Ceri S. False news on social media: a data-driven survey. *ACM Sigmod Record*, 2019, 48.2: 18-27. URL: https://dl.acm.org/doi/abs/10.1145/3377330.3377334 (дата звернення: 11.10.2023).

17. Барабаш А. О. *Система детектування Deep Fake відеозаписів на основі нейронної мережі*. 2019. Master's Thesis. КПІ ім. Ігоря Сікорського. URL: https://ela.kpi.ua/bitstream/123456789/31818/1/Barabash_magistr.pdf (дата звернення: 12.10.2023).

18. Young Dannagal G., et al. Fact-checking effectiveness as a function of format and tone: Evaluating FactCheck. org and FlackCheck. org. *Journalism & Mass Communication Quarterly*, 2018, 95.1: 49-75. URL: https://journals.sagepub.com/doi/abs/10.1177/1077699017710453 (дата звернення: 13.10.2023).

19. Appling Scott, Bruckman Amy, De Choudhury Munmun. Reactions to Fact Checking. *Proceedings of the ACM on Human-Computer Interaction*, 2022, 6.CSCW2: 1-17. URL: https://dl.acm.org/doi/abs/10.1145/3555128 (дата звернення: 14.10.2023).

20. Shao Chengcheng, et al. Tracking and characterizing the competition of fact checking and misinformation: case studies. *IEEE access*, 2018, 6: 75327-75341. URL: https://ieeexplore.ieee.org/abstract/document/8532356/ (дата звернення: 15.10.2023).

21. Nieminen Sakari, Sankari Valtteri. Checking politifact's fact-checks. *Journalism Studies*, 2021, 22.3: 358-378. URL: https://www.tandfonline.com/doi/abs/10.1080/1461670X.2021.1873818 (дата

звернення: 17.10.2023).

22. Walter Jessica Gabriele, et al. Assessing Methods to Analyze Spread of Misinformation in Digital Media. 2021. URL: https://datalab.au.dk/fileadmin/Datalab/SOMA_REPORTS/SOMA_D2.4_information_cascade_analysis.pdf (дата звернення: 20.10.2023).

23. Boyko K. Russia's information warfare against Ukraine and fact-checking: experience from StopFake. 2022. URL: https://www.kalendarium.uu.se/event?eventId=76788 (дата звернення: 22.10.2023).

24. Synchak B. Verification of facts as an integral part of professional journalism in hybrid wars. *European Humanitarian Studies*, 2021, 1: 112-125. URL: https://www.ceeol.com/search/article-detail?id=1043628 (дата звернення: 23.10.2023).

25. Азарова А. О., Пугач В. С. Інформаційна технологія виявлення фейкової інформації у соцмережах. Молодь в науці: дослідження, проблеми, перспективи (МН 2024): Міжнародна науково-практична інтернет-конференція, м. Вінниця, 2023. URL: https://conferences.vntu.edu.ua/index.php/mn/mn2024/paper/viewFile/19663/16276 (дата звернення: 24.10.2023).

26. Kelk Ian. Automatic Fake News Detection: Are current models" fact-checking" or" gut-checking"?. *arXiv preprint arXiv:2204.07229*, 2022. URL: https://arxiv.org/abs/2204.07229 (дата звернення: 24.10.2023).

27. Harrag Fouzi, Djahli Mohamed Khalil. Arabic fake news detection: A fact checking based deep learning approach. *Transactions on Asian and Low-Resource Language Information Processing*, 2022, 21.4: 1-34. URL: https://dl.acm.org/doi/abs/10.1145/3501401 (дата звернення: 25.10.2023).

28. Trokhymovych M., Saez-trumper D. Wikicheck: An end-to-end open source automatic fact-checking api based on wikipedia. In: *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*. 2021. p. 4155-4164. URL: https://dl.acm.org/doi/abs/10.1145/3459637.3481961 (дата звернення: 27.10.2023).

29. Vo Nguyen, Lee Kyumin. Learning from fact-checkers: Analysis and

generation of fact-checking language. In: *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2019. p. 335-344. URL: https://dl.acm.org/doi/abs/10.1145/3331184.3331248 (дата звернення: 27.10.2023).

30. Реєстр фейків України. Реєстр фейкових ЗМІ України. URL: https://fake.net.ua/ (дата звернення: 28.10.2023).

31. Pehlivanoglu Didem. Aging in an “infodemic”: The role of analytical reasoning, affect, and news consumption frequency on news veracity detection. *Journal of Experimental Psychology: Applied*, 2022. URL: https://psycnet.apa.org/record/2022-58029-001 (дата звернення: 30.10.2023).

32. Walker M. *Python Data Cleaning Cookbook: Modern techniques and Python tools to detect and remove dirty data and extract key insights*. Packt Publishing Ltd, 2020. URL: https://www.google.com/books?hl=uk&lr=&id=GMwOEAAAQBAJ&oi=fnd&pg=PP1&dq=clean+text+by+removing+HTML,+Python+characters&ots=_GKhTG-_LW&sig=PIKWgAyQkpEhFuS1W-fI2_xGRkA (дата звернення: 05.11.2023).

33. Luo M., Hancock Jeffrey T., Markowitz David M. Credibility perceptions and detection accuracy of fake news headlines on social media: Effects of truth-bias and endorsement cues. *Communication Research*, 2022, 49.2: 171-195. URL: https://journals.sagepub.com/doi/abs/10.1177/0093650220921321 (дата звернення: 06.11.2023).

34. Saleh H., Ahothali A., Moria K. Detection of hate speech using BERT and hate speech word embedding with deep model. *Applied Artificial Intelligence*, 2023, 37.1: 2166719. URL: https://www.tandfonline.com/doi/abs/10.1080/08839514.2023.2166719 (дата звернення: 06.11.2023).

35. Albahari J. *C# 10 in a Nutshell*. " O'Reilly Media, Inc.", 2022. URL: https://www.google.com/books?hl=uk&lr=&id=4CJdEAAAQBAJ&oi=fnd&pg=PP1&dq=C%23&ots=71RneHgZUG&sig=eF3JXzHN435uVujrJX4oUkj66x8 (дата звернення: 07.11.2023).

36. Martelli A. *Python in a Nutshell*. " O'Reilly Media, Inc.", 2023. URL: https://www.google.com/books?hl=uk&lr=&id=2WSmEAAAQBAJ&oi=fnd&pg=PT28&dq=python&ots=oViXBVSxb6&sig=yxplfFPItxs2GzYihGzFXmrgHQg (дата звернення: 08.11.2023).

37. Bonet-Jover Alba. Exploiting discourse structure of traditional digital media to enhance automatic fake news detection. *Expert systems with applications*, 2021, 169: 114340. URL: https://www.sciencedirect.com/science/article/pii/S0957417420310277 (дата звернення: 09.11.2023).

38. Singhal S. Spotfake: A multi-modal framework for fake news detection. In: *2019 IEEE fifth international conference on multimedia big data (BigMM)*. IEEE, 2019. p. 39-47. URL: https://ieeexplore.ieee.org/abstract/document/8919302/ (дата звернення: 10.11.2023).

39. Zhang Chaowei, et al. Detecting fake news for reducing misinformation risks using analytics approaches. *European Journal of Operational Research*, 2019, 279.3: 1036-1052. URL: https://ieeexplore.ieee.org/abstract/document/8919302/ (дата звернення: 11.11.2023).

ДОДАТКИ

Додаток А. Технічне завдання

Вінницький національний технічний університет

Факультет менеджменту та інформаційної безпеки

Кафедра менеджменту та безпеки інформаційних систем

ЗАТВЕРДЖУЮ

Голова секції "Управління інформаційною безпекою" кафедри МБІС

д.т.н., професор

_____ Юрій ЯРЕМЧУК

"20" вересня 2023 р.

ТЕХНІЧНЕ ЗАВДАННЯ

до магістерської кваліфікаційної роботи на тему:

Удосконалення методу виявлення фейкової інформації у соцмережах засобами штучного інтелекту на основі мережі LSTM

08-72.МКР.013.00.000.ТЗ

Керівник магістерської кваліфікаційної роботи

к.т.н., проф., проф. каф. МБІС Азарова А.О.

_____

Вінниця – 2023 р.

1. Найменування та область застосування

Удосконалення методу виявлення фейкової інформації у соцмережах засобами штучного інтелекту на основі мережі LSTM. Область застосування: виявлення фейкової інформації у текстах новин.

2.Підстава для розробки

Розробка виконується на основі наказу ректора ВНТУ №247 від 18. 09. 2023 р.

3. Мета та призначення розробки

3.1 Мета розробки: Удосконалення методу виявлення фейкової інформації у соцмережах засобами штучного інтелекту на основі мережі LSTM та програмна розробка сервісу для практичної реалізації алгоритму.

3.2 Призначення: розроблений засіб призначений для виявлення фейкової інформації на основі удосконаленого методу.

4. Джерела розробки

4.1. Комплексні системи захисту інформації: [навчальний посібник] Ю. Є. Яремчук, П. В. Павловський, В. С. Катаєв, В. В. Сінюгін. Вінниця : ВНТУ, 2018. 118 с.

4.2. Бурячок В.Л. Політика інформаційної безпеки: підручник. / В.Л.Бурячок, Р.В.Грищук, В.О.Хорошко / За заг. ред. докт. техн. наук, проф. В.О. Хорошка. К.: ПВП «Задруга», 2014. 222 с.

4.3. Harrag Fouzi, Djahli Mohamed Khalil. Arabic fake news detection: A fact checking based deep learning approach. *Transactions on Asian and Low-Resource Language Information Processing*, 2022, 21.4: 1-34. URL: https://dl.acm.org/doi/abs/10.1145/3501401 (дата звернення: 25.10.2023).

4.4. Walker M. Python Data Cleaning Cookbook: Modern techniques and Python tools to detect and remove dirty data and extract key insights. Packt Publishing Ltd, 2020. URL: https://www.google.com/books?hl=uk&lr=&id=GMwOEAAAQBAJ&oi=fnd&pg=PP1&dq=clean+text+by+removing+HTML,+Python+characters&ots=_GKhTG-_LW&sig=PIKWgAyQkpEhFuS1W-fI2_xGRkA (дата звернення: 05.11.2023).

5. Вимоги до програми

5.1 Вимоги до функціональних характеристик:

5.1.1 Програмний засіб повинен мати зручний, легкий у використанні інтерфейс користувача;

5.1.2 Реалізація методу не повинна вимагати спеціальних ліцензійних програмних додатків;

5.1.3 Програмний засіб повинен виконувати процес автентифікації користувачів у системі.

5.2 Вимоги до надійності:

5.2.1 Програмний засіб повинен працювати без помилок, у випадку виникнення критичних ситуацій необхідно передбачити виведення відповідних повідомлень;

5.2.2 Бази даних повинні бути налаштовані на автоматичне створення резервних копій;

5.2.3 Програмний засіб повинен виконувати свої функції.

5.3 Вимоги до складу і параметрів технічних засобів:

– процесор – Pentium 1500 МГц і подібні до них;

– оперативна пам’ять – не менше 512 Mb;

– середовище функціонування – операційна система сімейство Windows;

– вимоги до техніки безпеки при роботі з програмою повинні відповідати існуючим вимогам та стандартам з техніки безпеки при користуванні комп’ютерною технікою.

6. Вимоги до програмної документації

6.1 Обов’язкова поетапна інструкція для майбутніх користувачів, наведена у пункті 3.3

7. Вимоги до технічного захисту інформації

7.1 Необхідно забезпечити захист розроблюваного програмного засобу від несанкціонованого використання.

7.2 Неможливість отримання доступу незареєстрованих користувачів до онлайн-сервісу.

8. Техніко-економічні показники

8.1 Цінність результатів використання даного проекту повинна перевищувати витрати на його реалізацію.

8.2 Має бути реалізований таким чином, щоб підходити для використання широкого загалу.

9. Стадії та етапи розробки

№ з/п	Назва етапів магістерської кваліфікаційної роботи	Початок	Закінчення
1	Визначення напрямку магістерської роботи, формулювання теми	20.09.2023	25.09.2023
2	Аналіз предметної області обраної теми	26.09.2023	30.10.2023
3	Апробація отриманих результатів	31.10.2023	02.10.2023
4	Розробка алгоритму роботи	03.10.2023	17.10.2023
5	Написання магістерської роботи на основі розробленої теми	18.10.2023	10.11.2023
6	Розробка економічної частини	11.11.2023	23.11.2023
7	Передзахист магістерської кваліфікаційної роботи	24.11.2023	25.11.2023
8	Виправлення, уточнення, корегування магістерської кваліфікаційної роботи	26.11.2023	30.11.2023
9	Захист магістерської кваліфікаційної роботи	15.12.2023	15.12.2023

10. Порядок контролю та прийому

10.1 До приймання магістерської кваліфікаційної роботи надається:

– ПЗ до магістерської кваліфікаційної роботи;

– програмний додаток;

– презентація;

– відзив керівника роботи;

– відзив опонента

Технічне завдання до виконання прийняв _____ Пугач В. С.

Додаток Б. Лістинг програми

Інтерфейс програмного засобу

```
<!DOCTYPE html>
<html lang="uk">
<head>
    <title>Перевірка новин на фейк</title>
      <link rel="stylesheet" type="text/css" href="styles.css">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
</head>
<body>
    <div class="container">
        <header>
            <nav class="nav_bar">
                <div class="header_links">
                    <ul>
                        <li><a href="/">Головна</a></li>
                        <li><a href="/about">Про нас</a></li>
                        <li><a href="/help">F.A.Q.</a></li>
                        <li><a href="/feedback">Зворотний зв'язок</a></li>
                    </ul>
                </div>
                <div class="header_logo">
                    <a href="/" title="Перевірка новин на фейк">
                        <img class="header_logo" alt="Fake.Detect" title="Виявлення
фейкових новин за допомогою нейромереж" src="https://fact.ai/wp-
content/uploads/FactAI_Logo.svg">
                    </a>
                </div>
            </nav>
        </header>
        <main>
           </span><br> <h1>      FactAI: Перевірка новин</h1>
            <div class="content">
                <form action="/prediction" method="post">
                    <input class="textbox" type="text" name="title"
placeholder="     Заголовок новини">
                    <textarea class="message" name="text" placeholder="   Текст
новини" rows="4"></textarea>
                    <input class="textbox" type="text" name="url"
placeholder="     Посилання на джерело">
                        <button class="button"
type="submit">Перевірити</button>
                    <button class="button" type="submit">Додати до бази даних</button>
                </form>
            </div>
        </main>
        </span><br> <footer>
            <div class="footer_description">
    <h2>Наші переваги:</h2>
    <ul>
        <li><span class="bold">Автоматичне оновлення</span><br>База даних з новинами постійно
оновлюється, що дозволяє завжди коректно оцінювати інформацію, враховуючи мінливість стилів написання
неправдивих новин.</li>
        <li><span class="bold">Машинне навчання</span><br>Нейромережевий підхід з використанням
мережі LSTM забезпечує високу точність класифікації на україномовних текстах. Шляхом експериментів ми
обрали найкращі параметри мережі для роботи з українськими текстами.</li>
        <li><span class="bold">Унікальність розробки</span><br>Перший сервіс з підтримкою
української мови, здатний аналізувати україномовні та російськомовні тексти, побудований на власному
датасеті фейкових новин, що присутні в українському медіа-просторі. Наша система виявлення фейкової
інформації працює на основі потужної мережі LSTM, що дозволяє нам забезпечувати високу точність в роботі
з текстами та інформацією.</li>
    </ul>
</div>
        </footer>
```

```
    </div>
</body>
</html>
```

Стилі

```
/* Reset some default browser styles */
body, h1, h2, h3, p, ul, li {
    margin: 0;
    padding: 0;
}

/* Set a background color for the entire page */
body {
    background-color: #f0f0f0;
    font-family: Arial, sans-serif;
}

/* Style the header */
header {
    background-color: #333;
    color: #fff;
    padding: 10px;
    text-align: center;
}

.header_logo img {
    width: 150px;
    height: 150px;
    border: 0;
}

nav.nav_bar ul {
    list-style: none;
}

nav.nav_bar ul li {
    display: inline;
    margin-right: 20px;
}

nav.nav_bar a {
    text-decoration: none;
    color: #fff;
    font-weight: bold;
}

/* Style the main content */
.main {
    max-width: 800px;
    margin: 0 auto;
    padding: 20px;
    background-color: #fff;
    border: 1px solid #ddd;
    box-shadow: 0 0 5px rgba(0, 0, 0, 0.1);
}

h1 {
    font-size: 24px;
    margin-bottom: 20px;
    color: #333;
}

.content form input[type="text"], .content form textarea {
    width: 100%;
    padding: 10px;
    margin-bottom: 20px;
    border: 1px solid #ddd;
    border-radius: 4px;
}

.content form .button {
```

```
    background-color: #007bff;
    color: #fff;
    padding: 10px 20px;
    border: none;
    border-radius: 4px;
    cursor: pointer;
}

/* Style the footer */
footer {
    background-color: #333;
    color: #fff;
    padding: 20px;
    text-align: center;
}

.footer_description ul {
    list-style: none;
    padding: 0;
}

.footer_description h2 {
    font-size: 20px;
    margin-bottom: 10px;
}

.footer_description li {
    font-size: 16px;
    margin-bottom: 10px;
}

.bold {
    font-weight: bold;
}

/* Responsive design for small screens */
@media (max-width: 768px) {
    .main {
        padding: 10px;
    }

    .footer_description ul {
        text-align: left;
    }
}
```

Модуль нейромережі

```
import tensorflow as tf
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense

# Assuming training_embeddings and y_train are defined elsewhere

# Get input shape
input_shape = training_embeddings.shape
print(input_shape)

# Create a sequential model
model = Sequential()

# Add layers to the model
model.add(Dense(128, activation='sigmoid'))
model.add(Dense(128, activation='relu'))
model.add(Dense(2, activation='softmax'))

# Compile the model
model.compile(loss='binary_crossentropy', optimizer='adam', metrics=['acc'])

# Build the model with the specified input shape
model.build(input_shape)
model.summary()
```

```
# Set training parameters
num_epochs = 60
batch_size = 32

# Train the model
history = model.fit(training_embeddings, y_train, epochs=num_epochs,
validation_split=0.35, shuffle=False, batch_size=batch_size)
```

Модуль попередньої обробки тексту

```
import pandas as pd
from stop_words import get_stop_words
from nltk.corpus import stopwords
from nltk.stem.snowball import SnowballStemmer

data = pd.read_excel('fakenws.xlsx')
data = data[['title', 'text', 'label']]
data.columns = ['Title', 'Text', 'Label']
data_shape = data.shape
print(data_shape)

stop_words = get_stop_words('ukrainian', True)
print(stop_words)
my_stop_words = set(stopwords.words('ukrainian')).union(set(stop_words))


def remove_stop_words(text):
    words = utils.to_unicode(text).split()
    return " ".join(w for w in words if w not in my_stop_words)

data['Text'] = data['Text'].apply(remove_stop_words)
print(remove_stop_words("В україні скасовують інвалідність правити фронт більше людей"))
print(data['Text'])

stemmer = SnowballStemmer("ukrainian")


def stemming(text):
    text = [stemmer.stem(word) for word in text.split()]
    return " ".join(text)

data['Text'] = data['Text'].apply(stemming)
data.head(10)
```

Модуль будування графіків

```
import matplotlib.pyplot as plt
from sklearn.feature_extraction.text import CountVectorizer

# Plotting accuracy
plt.plot(history.history['acc'])
plt.plot(history.history['val_acc'])
plt.title('Model Accuracy')
plt.ylabel('Accuracy')
plt.xlabel('Epoch')
plt.legend(['Test', 'Train'], loc='upper left')
plt.show()

# Plotting model loss
plt.plot(history.history['loss'])
plt.plot(history.history['val_loss'])
plt.title('Model Loss')
plt.ylabel('Loss')
plt.xlabel('Epoch')
plt.legend(['Test', 'Train'], loc='upper left')
plt.show()

# Plotting bar chart
plt.bar(10, FAKE_len, 3, label="FAKE")
plt.bar(15, REAL_len, 3, label="REAL")
plt.legend()
```

```
plt.ylabel('Count')
plt.title('Proportion')
plt.show()

# CountVectorizer
count_vectorizer = CountVectorizer()
count_vectorizer.fit(data[data['Label'] == 'FAKE']['Text'])
dictionary = count_vectorizer.vocabulary_.items()

vocab = []
count = []

for key, value in dictionary:
    vocab.append(key)
    count.append(value)

vocab_bef_stem = pd.Series(count, index=vocab)
vocab_bef_stem = vocab_bef_stem.sort_values(ascending=False)
top_vocab = vocab_bef_stem.head(15)
top_vocab.plot(kind='barh', figsize=(6, 8), xlim=(5707, 5725), color=['orange'])
plt.title('Top Words in Fake News')
plt.show()
```

Модуль передбачення

```
from flask import Flask, request, render_template, redirect, url_for, send_from_directory
from flask_wtf import Form
from wtforms import StringField, SubmitField, TextAreaField
from wtforms.validators import Required
import numpy as np
import pandas as pd
import tensorflow as tf
import tensorflow_hub as hub
from nltk.tokenize.treebank import TreebankWordDetokenizer
from stop_words import get_stop_words
from uk_stemmer import UkStemmer

# Disable GPU
os.environ['CUDA_VISIBLE_DEVICES'] = '-1'

# Flask app
app = Flask("name")
app.config['SECRET_KEY'] = 'hard to guess string'
app.config['SEND_FILE_MAX_AGE_DEFAULT'] = 0

# Set random seed
seed_value = 0
np.random.seed(seed_value)
random.seed(seed_value)
tf.random.set_random_seed(seed_value)

# Load pre-trained model
model = tf.keras.models.load_model('.../saved_model/my_model')

# Load word embedding module
module_url = "https://tfhub.dev/google/universal-sentence-encoder-multilingual/3"
embed = hub.load(module_url)

# Function to get embeddings
def getEmbeddings(newsText):
    with tf.compat.v1.Session() as session:
        session.run([tf.compat.v1.global_variables_initializer(),
tf.compat.v1.tables_initializer()])
        training_embeddings = session.run(embed([newsText]))
    return training_embeddings

# Function to make predictions
def getPrediction(text):
    prediction_result = model.predict(getEmbeddings(text)[0:1])
    result = (prediction_result > 0.6).astype(int)
    return result
```

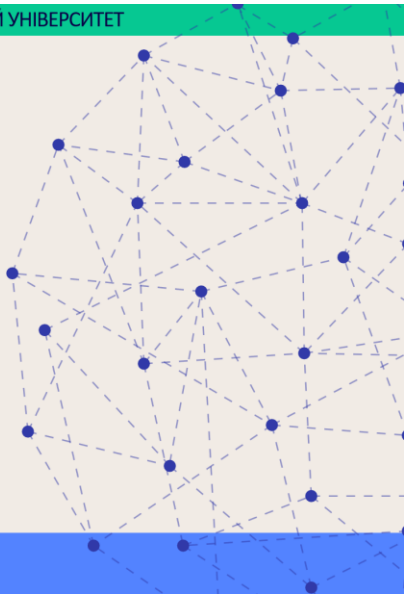
```
# Function to remove stop words
def stopWords(text):
    tokenizer = nltk.RegexpTokenizer(r"\w+")
    text_tokens = tokenizer.tokenize(text)
    tokens_without_sw = [word for word in text_tokens if not word in my_stop_words]
    filtered_text = TreebankWordDetokenizer().detokenize(tokens_without_sw)
    return filtered_text

# Function for stemming
def stemWords(text):
    stemmer = UkStemmer()
    text = [stemmer.stem_word(word) for word in text.split()]
    return " ".join(text)

# Function for text preprocessing
def textPreprocess(text):
    text = stopWords(text)
    text = stemWords(text)
    return text

# Function to get prediction label
def getPredictionLabel(prediction):
    MaxPosition = np.argmax(prediction)
    prediction_label = classes[MaxPosition]
    return prediction_label

# Flask form
class PostForm(Form):
    title = StringField('Заголовок новини ', validators=[Required()])
    text = TextAreaField('Текст новини ', validators=[Required()])
    url = StringField('URL новини ', validators=[Required()])
    submit = SubmitField('Перевірити')
    upload = SubmitField('Додати до бази даних')

# Run the app
if __name__ == "__main__":
    app.run(port=5000)
```

Додаток В. Ілюстративний матеріал

ВІННИЦЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ

МАГІСТЕРСЬКА КВАЛІФІКАЦІЙНА РОБОТА
на тему:

Удосконалення методу виявлення фейкової інформації у соцмережах засобами штучного інтелекту на основі мережі LSTM



ВИКОНАВ: СТ. ГР. 2КІТС-22М ПУГАЧ В. С.
Керівник: к.т.н., проф., проф. каф. МБІС - Азарова А. О.

НАУКОВА НОВИЗНА

Основним науковим результатом магістерської роботи є удосконалення процесу виявлення фейкової інформації засобами системного підходу та нейро-мережевого моделювання, а саме:

– удосконалено метод виявлення фейкових новин на основі налаштованої нейронної мережі LSTM, що, на відміну від існуючих аналогів, дозволяє виявити тонкі мовні ознаки, шаблони і характеристики фейків; крім того, застосування методів глибокого навчання під час оброблення природної мови уможливлює підвищення точності та надійності процесу ідентифікації неправдивої інформації у новинному контенті.

Робота вдосконаленого методу виявлення фейкових новин за допомогою штучного інтелекту на основі LSTM полягає у виконанні таких етапів.

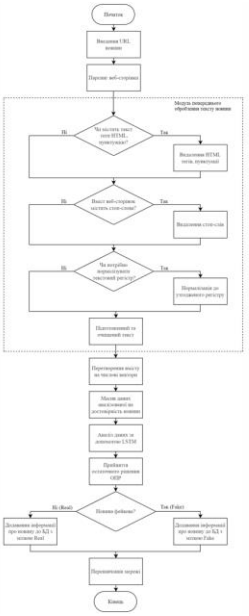
1. Збирання даних: це передбачає вилучення текстової інформації з різних веб-джерел, насамперед із соціальних мереж.
2. Попереднє оброблення: зібрані дані проходять етапи початкової попередньої обробки для підвищення їх якості та придатності для аналізу.
3. Інтелектуальний аналіз і прогнозування: для поглибленого аналізу та прогнозування автентичності вхідної інформації використовується нейронна модель, яка базується на мережах довготривалої короткочасної пам'яті (LSTM).
4. Прийняття остаточного рішення: результати аналізу штучного інтелекту надаються особі, що приймає рішення, яка остаточно оцінює правдивість інформації та вносить рішення в базу даних.
5. Перенавчання мережі: безперервне вдосконалення досягається шляхом періодичного перенавчання та тонкого налаштування мережі.

FACT AI

Для програмної реалізації удосконаленого методу виявлення фейкових новин використовувалася мова Python, оскільки вона має потужний арсенал бібліотек, що спроможні автоматизувати процес реалізації окремих модулів системи та надати доступ до потужних інструментів ШІ.

Отже, початок процесу збирання даних починається з введення URL новинної статті. Потім, застосовується парсинг веб-сторінки для вилучення відповідної інформації.

Враховуючи, що вміст веб-сторінки часто містить сторонні елементи, такі як спеціальні символи та теги HTML, етап попереднього оброблення стає обов'язковим. Щоб очистити текст шляхом видалення символів HTML, використовувалися вбудовані функції Python: `cleanText = re.compile(r'<[^>]+>|&[a-zA-Z0-9]+;')`



У цьому модулі уточнений текст піддається численним перетворенням, включаючи видалення стоп-слів, виділення основи слів (стеммінг), видалення знаків пунктуації та нормалізацію до узгодженого регістру. Щоб видалити стоп-слова, необхідно використовувати словники стоп-слів, що реалізується засобами бібліотеки Python:

```
nltk.download('stopwords')
stop_words = set(stopwords.words('ukrainian'))
```

Для нормалізації текстового регістру в IT використовується метод Python:

```
lowercase_text = text.lower()
```

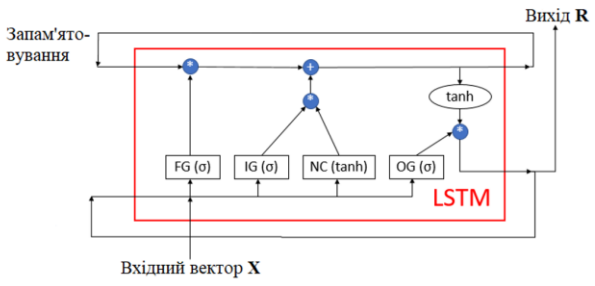
Для усунення пунктуації застосовано методом Python:

```
translator = str.maketrans('', '', string.punctuation)
cleaned_text = text.translate(translator)
```

Далі, необхідно перетворити оброблений вміст документа на числові вектори. Для цього застосуємо метод Word2Vec. Застосуємо бібліотеку Gensim для ефективного способу оброблення Word2Vec у Python:

```
from gensim.models import Word2Vec
from nltk.tokenize import word_tokenize
```

У модулі інтелектуального аналізу навчена мережа LSTM оцінює вміст вхідного вектору X, що представляє собою закодований масив даних аналізованої на достовірність новини. Дані, представлені у вигляді векторів слів, потрапляють до модуля нейромережі LSTM, яка генерує класифіковані дані. Класифіковані дані дозволяють LSTM оцінити достовірність новини, класифікуючи її як "Real" (достовірна) або "Fake" (фейкова). Розглянемо структуру мережі LSTM (на рис.) для оцінки достовірності новин у соціальних мережах. Тут на вхід подаються **X** – вектор вхідних даних, а не виході отримуємо **R** – вихідний результат ("Real" або "Fake").



Етап інтелектуального аналізу та прогнозування для удосконаленого методу

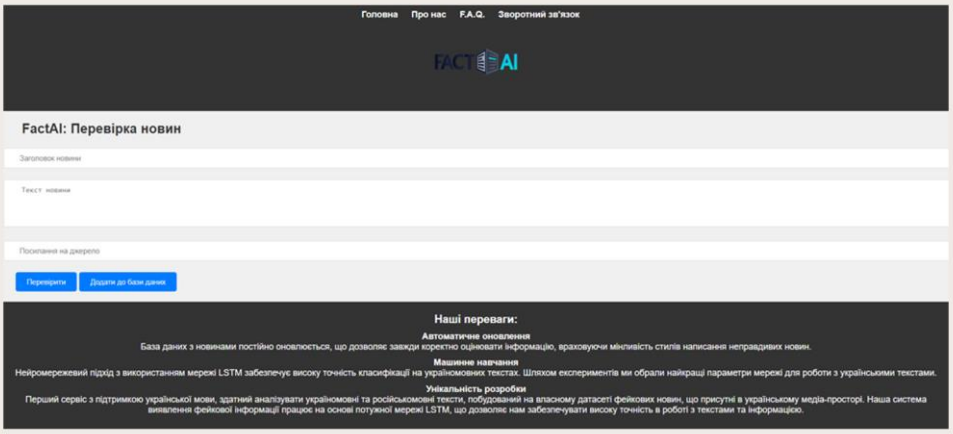
Отриманий результат передається ОПР для прийняття остаточного рішення і додавання його до бази даних із відповідним ярликом – "Real" або "Fake". Отже, цей процес є важливим у навчанні мережі LSTM, що спричинюється відповідними модифікаціями, внесеними ОПР до бази даних. Такий підхід уможливлює розширення набору навчальних даних, що підвищує точність подальшої роботи ІТ, сприяючи більш надійному виявленню фейкових новин із часом.

title	text	label	url
Шабунін та Шерембей планують зверн	Засновник Центру протидії корупції Віталій Шабулін та Дмитро Шеремб	Fake	https://ukr-news.vn.u
Україна не отримає репарацій через т	Україна не отримає репарацій від Росії, а її неповнолітні громадяни не м	Fake	https://www.my-worl
Україна готує підрив Кременчуцької Г	Отримано дані про мінування і можливий підрив Кременчуцької ГЕС з м	Fake	https://infoteka.biz.u
В Україні скасовують інвалідність, що	Чиновники пропонують реформувати Медико-соціальну експертну ком	Fake	http://www.presentn
На Сумщині формується нове з'єднан	24 жовтня передплатник одного з українських антифашистських каналів	Fake	https://news-kharkov

Збереження інформації у базі даних новин

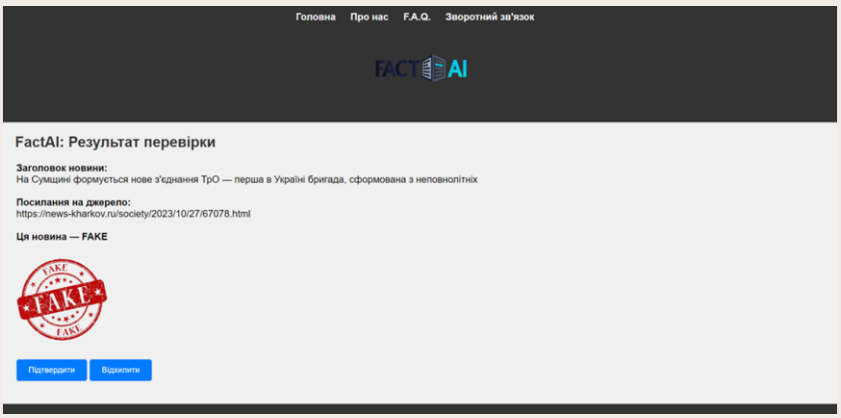
Таким чином, запропоновано удосконалений метод вивлення фейків, спроможний виділяти текстовий контент із веб-статей новин, проводити попереднє оброблення тексту, здійснювати інтелектуальний аналіз і прогнозування, надавати операторам можливість вносити виправлення і самостійно оцінювати достовірність новин. Слід зауважити, що такий підхід дозволяє постійне навчання мережі на основі інформації, отриманої від ОПР.

АНАЛІЗ РОБОТИ ІС, ЇЇ ТЕСТУВАННЯ ТА ПЕРЕВІРКА АДЕКВАТНОСТІ



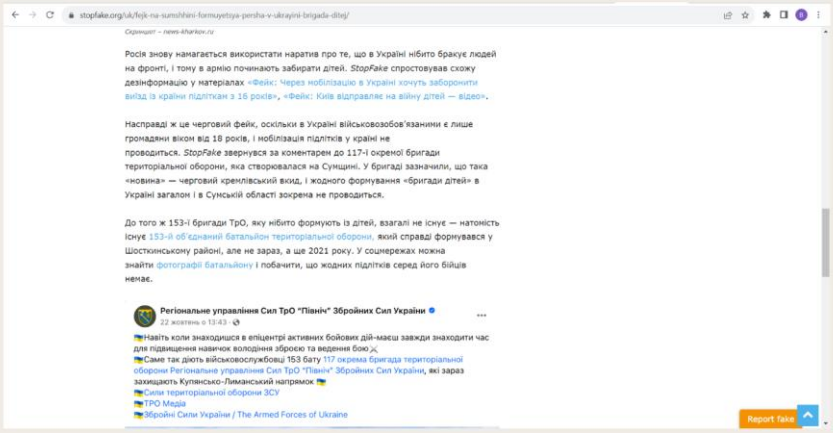
Інтерфейс головної сторінки для реалізації удосконаленого методу виявлення фейкових новин

ДОВЕДЕННЯ АДЕКВАТНОСТІ РОЗРОБЛЕНОГО ПІДХОДУ



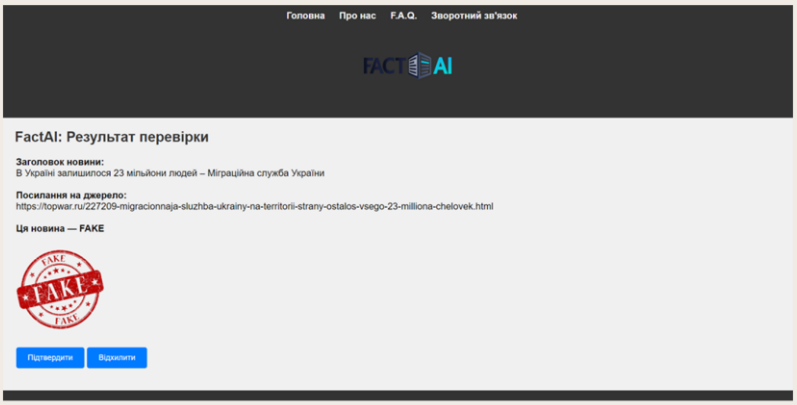
Результат перевірки новини 1 запропопонованим методом

ДОВЕДЕННЯ АДЕКВАТНОСТІ РОЗРОБЛЕНОГО ПІДХОДУ



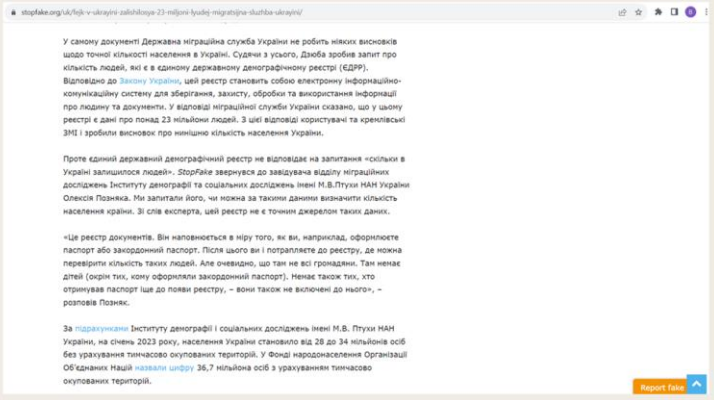
Результат перевірки новини 1 офіційним каналом Stopfake

ДОВЕДЕННЯ АДЕКВАТНОСТІ РОЗРОБЛЕНОГО ПІДХОДУ



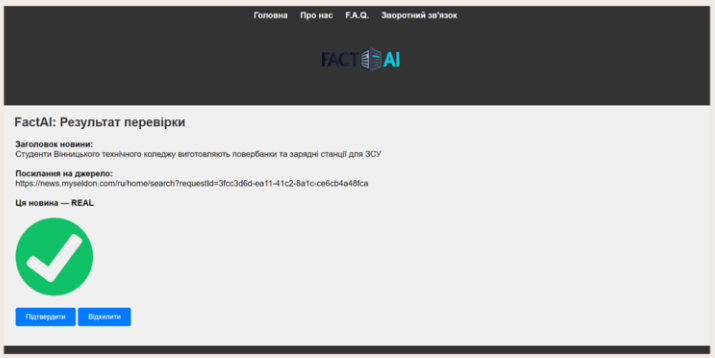
Результат перевірки новини 2 запропонованим методом

ДОВЕДЕННЯ АДЕКВАТНОСТІ РОЗРОБЛЕНОГО ПІДХОДУ



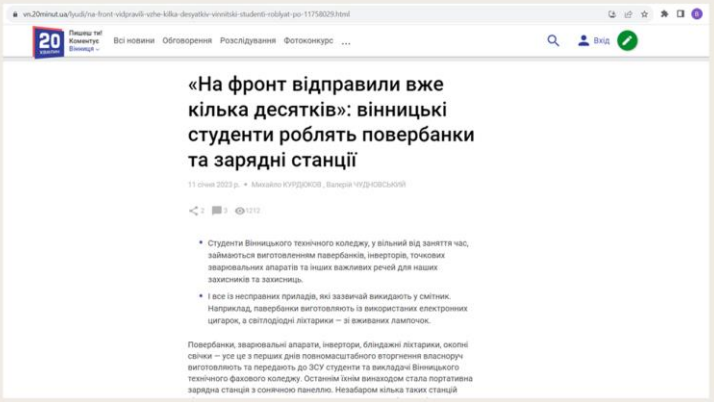
Результат перевірки новини 2 офіційним каналом Stopfake

ДОВЕДЕННЯ АДЕКВАТНОСТІ РОЗРОБЛЕНОГО ПІДХОДУ



Результат перевірки новини 3 запропонованим методом

ДОВЕДЕННЯ АДЕКВАТНОСТІ РОЗРОБЛЕНОГО ПІДХОДУ



Результат перевірки новини 3 офіційним каналом 20 хвилин Вінниця

ВИСНОВКИ

У результаті виконання завдань магістерської роботи було досягнено поставленої мети щодо вдосконалення методу виявлення фейкової інформації у соціальних мережах. Це було реалізовано завдяки впровадженню інструментів штучного інтелекту на основі мережі довгострокової короткочасної пам'яті (LSTM). Підхід на основі LSTM, як продемонстровано в цій роботі, довів свою ефективність, забезпечуючи точні результати та високу продуктивність.

Удосконалений метод має переваги в ключових сферах, таких як точність, швидкість оброблення, а також в аспекті вдосконалення механізмів зворотного зв'язку та можливостей інтеграції. Ці переваги підкреслюють позивні перспективи практичного застосування запропонованого методу, представ-леного його програмною ралізацією.

ПРОТОКОЛ
ПЕРЕВІРКИ КВАЛІФІКАЦІЙНОЇ РОБОТИ
НА НАЯВНІСТЬ ТЕКСТОВИХ
ЗАПОЗИЧЕНЬ

Назва роботи: Удосконалення методу виявлення фейкової інформації у соцмережах засобами штучного інтелекту на основі мережі LSTM
Тип роботи: магістерська кваліфікаційна робота
(БДР, МКР)

Підрозділ: Кафедра менеджменту та безпеки інформаційних систем
Факультет менеджменту та інформаційної безпеки
(кафедра, факультет)

Показники звіту подібності Unicheck

Оригінальність 99 % Схожість 1 %

Аналіз звіту подібності (відмітити потрібне):
1. Запозичення, виявлені у роботі, оформлені коректно і не містять ознак плагіату.
2. Виявлені у роботі запозичення не мають ознак плагіату, але їх надмірна кількість викликає сумніви щодо цінності роботи і відсутності самостійності її виконання автором. Роботу направити на розгляд експертної комісії кафедри.
3. Виявлені у роботі запозичення є недобросовісними і мають ознаки плагіату та/або в ній містяться навмисні спотворення тексту, що вказують на спроби приховування недобросовісних запозичень.

Особа, відповідальна за перевірку _____ Коваль Н.П.
(підпис) (прізвище, ініціали)

Ознайомлені з повним звітом подібності, який був згенерований системою Unicheck щодо роботи.

Автор роботи _____ Пугач В.С.
(підпис) (прізвище, ініціали)

Керівник роботи _____ Азарова А.О.
(підпис) (прізвище, ініціали)