

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»
Факультет інформатики та обчислювальної техніки
Кафедра обчислювальної техніки

«На правах рукопису»
УДК 004.8

До захисту допущено:
В. О. Завідувача кафедри
 Михайло Новотарський
« » 2024 р

Магістерська дисертація
на здобуття ступеня магістра

за освітньо-професійною програмою «Комп'ютерні системи та мережі» зі
спеціальності 123 «Комп'ютерна інженерія»

на тему: «Автоматична верифікація новин через онлайн сервіси»

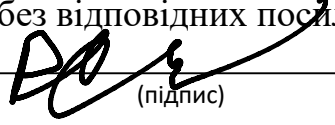
Виконав (-ла):
студент (-ка) II курсу, групи ІО-31мп
Ровишин Андрій Вячеславович

Керівник:
Асистент каф. ОТ, доктор філософії,
Шульга Максим Володимирович

Консультант з нормоконтролю:
асистент каф. ОТ, к.т.н.
Кулаков Олексій Юрійович

Рецензент:
Доцент кафедри ІСТ, к.т.н.
Галушко Дмитро Олександрович

Засвідчую, що у цій магістерській
дисертації немає запозичень з праць ін-
ших авторів без відповідних посилань.

Студент

(підпис)

Київ – 2024 року

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

Факультет Інформатики та обчислювальної техніки
(повна назва)

Кафедра Обчислювальної техніки
(повна назва)

Рівень вищої освіти – другий (магістерський)

Спеціальність 123. Комп'ютерна інженерія
(код і назва)

Освітньо-професійна програма Комп'ютерні системи та мережі

ЗАТВЕРДЖУЮ
В.о. Завідувача кафедри
Михайло НОВОТАРСЬКИЙ
(підпис) (ініціали, прізвище)

«___» _____ 2024 р.

ЗАВДАННЯ

на магістерську дисертацію студенту

Ровишину Андрію Вячеславовичу

(прізвище, ім'я, по батькові)

1. Тема дисертації “Автоматична верифікація новин через онлайн сервіси”
науковий керівник дисертації Шульга Максим Володимирович, асистент,
доктор філософії,

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом по університету від «08» Листопада 2024р. № 5016-с

2. Термін подання студентом дисертації 29.11.2024

3. Вихідні дані до проекту: технічне завдання, теоритичні дані, наукові
публікації

4. Зміст пояснювальної записки: огляд існуючих веб-сервісів, розгляд схем
їх реалізації, а також принципів роботи, пошук та аналіз їх основних переваг
та недоліків, проектування та розробка програмного забезпечення,
тестування програмного модулю та аналіз отриманих результатів

5. Перелік завдань, які потрібно розробити: огляд існуючих веб-сервісів,
розгляд схем їх реалізації, а також принципів роботи, проектування та
розробка програмного забезпечення, тестування програмного модулю та
аналіз отриманих результатів

6. Консультанти розділів дисертації*

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		Завдання видав	Завдання прийняв
Нормоконтроль	Русанова О.В.		
Нормоконтроль	Русанова О.В.		

7. Дата видачі завдання _____

Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Строк виконання етапів магістерської дисертації	Примітка
1	<i>Обговорення теми дисертації. Визначення предмету дослідження</i>	1.09.24-10.09.24	
2	<i>Дослідження існуючих рішень та проблем у досліджуваній області</i>	11.09.24-18.09.24	
3	<i>Побудова архітектурного рішення</i>	20.09.24-29.09.24	
4	<i>Програмна реалізація створеного алгоритму</i>	1.10.24-13.10.24	
5	<i>Програмування та загальне налагодження системи.</i>	15.10.24-30.10.24	
6	<i>Розробка стартам-проекту</i>	30.10.24-14.11.24	
7	<i>Оформлення магістерської дисертації.</i>	15.11.24-29.11.24	

Студент


(підпис)

Ровишин Андрій

Науковий керівник дисертації

(підпис)

Максим ШУЛЬГА

АНОТАЦІЯ

Ця магістерська дисертація зосереджується на технічних аспектах створення системи автоматичної верифікації новин із використанням онлайн-сервісів. У відповідь на актуальну проблему поширення фейкових новин дослідження охоплює огляд сучасних технологій та методів, які дозволяють автоматизувати процес перевірки достовірності інформації в медіа. Основна увага приділена аналізу методів машинного навчання та обробки природної мови (NLP), а також використанню метаданих та зовнішніх API для верифікації контенту. У роботі розроблено архітектуру системи, яка включає серверну і клієнтську частини, а також інтеграцію з популярними Python-бібліотеками (наприклад, NLTK, spaCy, scikit-learn) для аналізу текстових даних та навчання моделей. Метою дослідження є створення надійного інструменту для перевірки новин, який може ефективно ідентифікувати фальшиву інформацію та сприяти боротьбі з дезінформацією в онлайн-просторі. Розроблена система дозволяє користувачам швидко отримувати оцінку достовірності новин, забезпечуючи безпеку та конфіденційність даних, і має потенціал для подальшого впровадження у різних медіа-платформах.

ABSTRACT

This master's thesis focuses on the technical aspects of creating an automated news verification system using online services. In response to the urgent problem of fake news dissemination, the research provides an overview of current technologies and methods that enable the automation of information verification in the media. The main emphasis is placed on analyzing machine learning and natural language processing (NLP) methods, as well as utilizing metadata and external APIs for content verification. The work includes the development of a system architecture that comprises both server-side and client-side components, along with integration of popular Python libraries (such as NLTK, spaCy, and scikit-learn) for text data analysis and model training. The aim of the study is to create a reliable tool for news verification that can effectively identify false information and support the fight against disinformation in the online space. The developed system enables users to quickly obtain an assessment of news credibility, ensuring data security and privacy, and has the potential for further implementation across various media platforms.

РЕФЕРАТ

на магістерську дисертацію,

виконану на тему: Автоматична верифікація новин через онлайн сервіси

Студентом: Ровишин Андрій Вячеславович

Обсяг та структура роботи: Дана магістерська дисертація містить наступні розділи: вступ, чотири основні розділи та висновки. Загалом робота налічує 102 сторінки основного тексту, включає 40 ілюстрацій, 22 таблиці та 2 додатки. Для виконання роботи було використано 27 джерел літератури.

Актуальність роботи: Проблема поширення фейкових новин залишається критичною для суспільства, зокрема в умовах швидкого розвитку цифрових медіа. Вплив недостовірної інформації може мати серйозні наслідки для суспільної думки, економіки та навіть національної безпеки. Автоматизація процесу верифікації новин за допомогою сучасних онлайн сервісів та алгоритмів машинного навчання сприятиме швидкому та ефективному виявленню фейків, підвищуючи якість інформаційного середовища.

Мета роботи: Метою даної магістерської роботи є розробка системи, яка дозволяє автоматично перевіряти достовірність новин, використовуючи алгоритми машинного навчання, лінгвістичний аналіз тексту та метадані. Дана система також інтегрує онлайн сервіси для верифікації новин та забезпечує доступ до додаткових джерел для отримання об'єктивної інформації.

Завдання дослідження:

- Проведення аналізу основних проблем, пов'язаних з фейковими новинами, та огляд сучасних методів їхньої верифікації.
- Вивчення різних підходів до автоматизації процесу перевірки новин, зокрема з використанням алгоритмів обробки природної мови (NLP) та машинного навчання.
- Розробка архітектури системи для автоматичної верифікації новин та вибір оптимальних інструментів для її реалізації.
- Побудова бази даних для зберігання новин та результатів верифікації.

- Розробка програмного інтерфейсу для користувачів, що дозволить перевірити достовірність новин та отримати аналітичні дані.

Об’єкт дослідження: методи автоматичної верифікації новин через використання онлайн сервісів.

Предмет дослідження: технології та алгоритми машинного навчання, застосовані для виявлення та класифікації фейкових новин.

Практична цінність: Запропонована система для автоматичної верифікації новин дозволяє користувачам швидко перевіряти достовірність інформації, що значно підвищує рівень об’єктивності у медійному середовищі. Основні переваги розробленої системи:

Точність та надійність: Автоматизовані алгоритми машинного навчання ефективно ідентифікують фальшиві новини, зменшуючи ймовірність помилок.

Швидкість перевірки: Використання онлайн сервісів дозволяє проводити верифікацію новин у режимі реального часу, що особливо корисно для ЗМІ, журналістів та аналітиків.

Масштабованість: Система може бути легко адаптована для перевірки великих обсягів даних, що робить її придатною для використання у новинних агенціях та соціальних медіа.

Ключові слова: верифікація новин, фейкові новини, машинне навчання, обробка природної мови (NLP), онлайн сервіси, автоматизація.

Зміст

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ ТА СКОРОЧЕНЬ	10
ВСТУП	11
РОЗДІЛ 1. АНАЛІЗ ПРОБЛЕМИ ТА ПЕРЕГЛЯД ІСНУЮЧИХ РІШЕНЬ ДЛЯ АВТОМАТИЧНОЇ ВЕРИФІКАЦІЇ НОВИН	12
1.1 Вступ до проблематики фейкових новин та необхідності автоматичної верифікації	12
1.2. Основні виклики та проблеми, пов'язані з фальсифікацією новин	14
1.3. Огляд існуючих методів та інструментів для верифікації новин	15
1.3.1. Підходи на основі машинного навчання	16
1.3.2. Лінгвістичні методи аналізу тексту	17
1.3.3. Використання метаданих для перевірки джерела	18
1.4. Порівняння методів верифікації новин за точністю, швидкістю та масштабованістю	20
1.5. Огляд існуючих онлайн сервісів для верифікації новин	22
1.5.1. Сервіси для верифікації тексту (наприклад, Snopes, FactCheck)	22
1.5.2. Інструменти для аналізу медіа контенту (зображень та відео)	26
Висновки до розділу 1	29
РОЗДІЛ 2. ПРОЕКТУВАННЯ СИСТЕМИ ДЛЯ АВТОМАТИЧНОЇ ВЕРИФІКАЦІЇ НОВИН	31
2.1. Огляд архітектури системи для автоматичної верифікації новин	31
2.2. Вибір алгоритмів і методів для класифікації новин	34
2.2.1. Методи на основі обробки природної мови (NLP)	35
2.2.2. Моделі машинного навчання для класифікації фейкових новин	37
2.2.3. Використання нейронних мереж для виявлення фейкових новин	42
2.3. Вибір бази даних для зберігання новин, результатів верифікації та аналітичних даних	45
2.4. Інтеграція із зовнішніми сервісами для отримання додаткової інформації (API соціальних мереж, новинних агентств)	49
2.5. Забезпечення безпеки та конфіденційності в системі автоматичної верифікації	52
Висновки до розділу 2	56
РОЗДІЛ 3. РОЗРОБКА СИСТЕМИ АВТОМАТИЧНОЇ ВЕРИФІКАЦІЇ НОВИН	58

3.1. Структура та компоненти проекту	58
3.2. Розробка серверної частини проекту (Back-End)	63
3.2.1. Реалізація маршрутизації запитів для отримання та перевірки новин	63
3.2.2. Побудова структури бази даних для зберігання новин та результатів верифікації	67
3.3. Розробка клієнтської частини проекту (Front-End)	71
3.3.1. Інтерфейс для користувачів, що дозволяє перевірити новини на достовірність	71
3.3.2. Додаткові модулі аналітики та статистики для користувачів	74
3.4. Інтеграція Python-бібліотек для роботи з текстовими даними (наприклад, NLTK, spaCy) та навчання моделей (наприклад, scikit-learn, TensorFlow)	79
3.5. Тестування та налагодження системи	82
Висновки до розділу 3	89
РОЗДІЛ 4. РОЗРОБКА СТАРТАП-ПРОЕКТУ	92
4.1. Опис ідеї та цінності проекту автоматичної верифікації новин	92
4.2. Інформаційна картка проекту	94
4.3. Технічна інфраструктура	96
4.4. Технологічний аудит та аналіз конкурентного середовища	97
4.5. Стратегічний аналіз ринку та цільова аудиторія	99
4.6. Бізнес-модель проекту	101
4.6.1. Цінність продукту для користувачів та ринку	101
4.6.2. Інноваційність і перспективи розвитку проекту	101
4.6.3. Основні сегменти споживачів та їх потреби	102
4.7. Розробка маркетингової стратегії для просування сервісу	102
Висновки до розділу 4	103
ВИСНОВКИ	104
СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ	106
Додаток А Діаграма класів	108
Додаток Б Лістинг програми	109

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ ТА СКОРОЧЕНЬ

Скорочення, термін, позначення	Пояснення
ІІІ	Штучний інтелект
LLM	Великі мовні моделі
AFP	Agence France-Presse
NLP	Natural language processing – Обробка природної мови
API	Application Programming Interface
NLTK	Natural Language Toolkit

ВСТУП

В умовах сучасного інформаційного суспільства проблема поширення фейкових новин набула критичного значення. З розвитком соціальних мереж та онлайн-платформ швидкість передачі інформації зростає, що створило сприятливе середовище для розповсюдження неправдивих або викривлених даних. Фейкові новини стали потужним інструментом маніпуляції громадською думкою, що може мати негативні наслідки на соціальну стабільність, політичні процеси та економічну ситуацію в країні.

Автоматична верифікація новин за допомогою спеціалізованих онлайн-сервісів є важливим напрямом у боротьбі з фейковими новинами. Такі системи можуть аналізувати текстові дані, виявляючи в них ознаки неправдивої інформації, перевіряти достовірність джерела та визначати рівень довіри до певної публікації. Використання технологій машинного навчання, обробки природної мови (NLP), аналізу метаданих та інтеграція з зовнішніми інформаційними базами дозволяє автоматизувати процес перевірки новин, роблячи його швидшим та масштабованим.

Актуальність цього дослідження зумовлена необхідністю розробки інноваційних рішень для верифікації інформації, які зможуть допомогти користувачам швидко та ефективно ідентифікувати неправдиві новини. Метою даної роботи є створення системи автоматичної верифікації новин на базі Python з використанням сучасних бібліотек для аналізу тексту та навчання моделей. Дослідження охоплює аналіз існуючих підходів та інструментів для верифікації новин, проектування архітектури системи, вибір ефективних алгоритмів класифікації та реалізацію інтеграції з зовнішніми сервісами для підвищення точності перевірки.

Розробка такої системи має на меті сприяти підвищенню інформаційної безпеки суспільства, дозволяючи користувачам своєчасно виявляти фальсифіковані новини та зменшувати їхній вплив на громадську думку.

РОЗДІЛ 1.

АНАЛІЗ ПРОБЛЕМИ ТА ПЕРЕГЛЯД ІСНУЮЧИХ РІШЕНЬ ДЛЯ АВТОМАТИЧНОЇ ВЕРИФІКАЦІЇ НОВИН

У сучасному світі, де інформація поширюється з неймовірною швидкістю, проблема фейкових новин стає все більш актуальною. Відсутність достовірної інформації може призвести до серйозних наслідків, включаючи соціальні конфлікти, паніку та дезінформацію громадськості. З цієї причини виникає необхідність у розробці ефективних рішень для автоматичної верифікації новин. Буде проведено аналіз проблеми фейкових новин та перегляд існуючих підходів до їх автоматичної верифікації.

1.1 Вступ до проблематики фейкових новин та необхідності автоматичної верифікації

У сучасну еру інформаційних технологій, фейкові новини стали значущою проблемою, що впливає на різні аспекти суспільного життя. Стикаючись з великою кількістю дезінформації, суспільство потребує ефективних механізмів для її виявлення та запобігання. Ця проблема обумовлює необхідність створення автоматичних систем верифікації, які допоможуть забезпечити достовірність інформації та підвищити рівень інформаційної безпеки.

Фейкові новини – це свідомо вигадані чи маніпулятивно подані новини, які мають на меті дезінформувати аудиторію [1]. Вони часто використовуються для політичного впливу, фінансової вигоди або соціальної провокації. Поширення таких новин здійснюється за допомогою соціальних мереж, новинних порталів та інших медіа-платформ.

Фейкові новини можуть призводити до серйозних наслідків [2][3], таких що, показані на рис. 1.1.



Рис. 1.1. Проблеми та наслідки фейкових новин

Дезінформація громадськості – збільшення кількості недостовірної інформації ускладнює процес формування об'єктивного погляду на події.

Втрата довіри до медіа – постійне поширення фейкових новин підриває довіру до медіа-джерел.

Соціальна напруга – фейкові новини можуть викликати паніку, ненависть та соціальні конфлікти

Автоматична верифікація новин є критичним інструментом у боротьбі з фейковими новинами. Впровадження таких систем приводить до позитивних наслідків [4], що представлені на рис. 1.2.



Рис. 1.2. Позитивні риси впровадження автоматичної верифікації новин

Швидке виявлення фейкових новин – алгоритми штучного інтелекту здатні аналізувати величезні обсяги даних і швидко ідентифікувати недостовірну інформацію.

Підвищення довіри до медіа – автоматична верифікація допомагає забезпечити достовірність новинних повідомлень, що підвищує довіру до інформаційних джерел

Зменшення соціальної напруги – вчасне виявлення та запобігання поширенню фейкових новин сприяє стабільності в суспільстві.

Фейкові новини становлять серйозну загрозу для суспільства, підриваючи довіру до медіа та викликаючи соціальну напругу. У зв'язку з цим автоматична верифікація новин стає необхідністю. Використання сучасних технологій для перевірки достовірності інформації дозволяє ефективно боротися з дезінформацією та підтримувати інформаційну безпеку.

1.2. Основні виклики та проблеми, пов'язані з фальсифікацією новин

Фальсифікація новин є серйозною проблемою в сучасному інформаційному просторі. З поширенням цифрових технологій та соціальних мереж, процес створення та розповсюдження неправдивої інформації значно спростився, що призвело до широкомасштабних негативних наслідків для суспільства. Виклики, з якими стикаються як окремі особи, так і цілі організації, вимагають ефективних рішень для боротьби з фейковими новинами.

Виклики та проблеми детально розглянуті в таблиці 1.1

Таблиця 1.1

Виклики та проблеми фальсифікації новин

№	Виклики та проблеми	Пояснення
1	2	3
1	Швидкість поширення фейкових новин	Фейкові новини можуть поширюватися миттєво через соціальні мережі та інші онлайн-платформи, що робить їх важкими для контролю та зупинки [3]
2	Вплив на громадську думку	Фейкові новини можуть сильно вплинути на громадську думку, формуючи неправильне сприйняття подій та осіб. Це може призвести до соціального та політичного напруження. [2]

1	2	3
3	Економічні збитки	Компанії та організації можуть зазнати значних фінансових збитків внаслідок поширення фейкових новин, які можуть підривати їхню репутацію та довіру клієнтів [1]
4	Правові виклики	Визначення відповідальності за поширення фейкових новин є складним завданням через відсутність чітких правових рамок та міжнародного співробітництва [5]
5	Технологічні виклики	Фейкові новини стають все більш складними для виявлення через використання передових технологій, таких як діпфейки та автоматизовані боти [2]

Фальсифікація новин представляє серйозні виклики для суспільства, що вимагають комплексного підходу до їх вирішення. Ефективна боротьба з фейковими новинами повинна включати технічні, правові та соціальні заходи. Лише спільними зусиллями можна досягти зменшення впливу дезінформації та забезпечити інформованість громадськості.

1.3. Огляд існуючих методів та інструментів для верифікації новин

В епоху цифрових медіа та швидкого поширення інформації важливо мати надійні методи та інструменти для верифікації новин. Це допомагає забезпечити точність та достовірність інформації, яку споживають громадяни, журналісти та організації. Верифікація новин є ключовим етапом у боротьбі з дезінформацією та фейковими новинами.

Верифікація новин є важливим процесом у забезпеченні точності та достовірності інформації в сучасному світі. Використання різноманітних методів та інструментів допомагає зменшити вплив дезінформації та підвищити рівень довіри до медіа. Завдяки цим технологіям та платформам можна забезпечити, що громадяни отримують правдиву та об'єктивну інформацію.

1.3.1. Підходи на основі машинного навчання

Машинне навчання стає все більш важливим інструментом для верифікації новин. Використання алгоритмів та моделей машинного навчання дозволяє швидко та точно аналізувати великі обсяги даних, що допомагає виявити фейкові новини та підтвердити достовірність інформації

Підходи на основі машинного навчання показані на рис 1.3 та описані в таблиці 1.2.



Рис. 1.3. Підходи на основі машинного навчання

Таблиця 1.2

Підходи на основі машинного навчання

№	Підходи	Пояснення
1	2	3
1	Аналіз тексту	Машинне навчання застосовується для аналізу тексту новин за допомогою моделей, таких як нейронні мережі. Ці моделі можуть розпізнавати патерни, які вказують на потенційно фейкову інформацію. Зокрема, алгоритми вчать визначати структури мови та ключові слова, характерні для дезінформації [4]
2	Класифікація джерел	Алгоритми машинного навчання можуть аналізувати джерела новин і класифікувати їх за рівнем достовірності. Це включає аналіз історії публікацій, репутації джерела та взаємодії з іншими достовірними джерелами. Моделі можуть визначати, чи є певне джерело надійним або схильним до публікації фейкових новин [6]

1	2	3
3	Виявлення фейкових зображень та відео	Методи машинного навчання застосовуються для виявлення фейкових зображень та відео (діпфейків). Ці методи включають аналіз візуальних ознак, що можуть свідчити про маніпуляції, та використання алгоритмів для порівняння з оригінальними зображеннями чи відео [7]
4	Багатомовні моделі	Використання багатомовних моделей машинного навчання дозволяє верифікувати новини на різних мовах. Це важливо для глобальної боротьби з фейковими новинами, оскільки дезінформація може поширюватися у міжнародному контексті [8]

Методи на основі машинного навчання є потужними інструментами для верифікації новин, надаючи можливість швидкого та точного аналізу інформації. Завдяки цим технологіям можна значно зменшити вплив фейкових новин та забезпечити високу якість медіа-контенту.

1.3.2. Лінгвістичні методи аналізу тексту

Лінгвістичні методи аналізу тексту є важливим компонентом у процесі верифікації новин. Вони дозволяють виявити маніпулятивний контент та підозрілу інформацію через аналіз мови, структури тексту та використання лексичних одиниць. Ці методи значно підвищують точність визначення правдивості новинних повідомлень.

Основні лінгвістичні методи аналізу тексту показані на рис. 1.4.



Рис. 1.4. Основні лінгвістичні методи аналізу тексту

Лінгвістичні методи використовують аналіз лексики та семантики для виявлення фейкових новин. Слова та фрази, які часто зустрічаються в контексті недостовірних новин, можуть бути автоматично виявлені та перевірені. Наприклад, алгоритми можуть ідентифікувати сенсаційні заголовки або надмірно емоційний контент, що часто є характерними ознаками фейкових новин [9].

Лінгвістичні методи також включають аналіз стилю та тону тексту. Наприклад, тексти фейкових новин можуть містити більшу кількість граматичних помилок, використання зайвих займенників або надмірно ускладнені речення. Аналіз тону може виявити надмірну емоційність або упередженість, що нехарактерно для достовірних новин [10].

Важливим аспектом є також аналіз структури та зв'язності тексту. Недостовірні новини часто мають порушену логічну структуру або нечіткі зв'язки між частинами тексту. Лінгвістичні методи дозволяють виявляти такі аномалії та оцінювати когерентність тексту [11].

Лінгвістичні методи аналізу тексту є важливим інструментом у верифікації новин. Вони дозволяють виявити маніпулятивний контент, покращують розуміння тексту та забезпечують більш високу точність визначення достовірності інформації. Завдяки цим методам можна ефективно боротися з поширенням фейкових новин та підтримувати високий рівень інформаційної безпеки.

1.3.3. Використання метаданих для перевірки джерела

В епоху цифрових медіа та швидкого поширення інформації важливо мати методи для перевірки достовірності новин. Один з таких методів – використання метаданих, які допомагають визначити автентичність джерела інформації.

Метадані – це додаткова інформація про джерело, яка може включати дату публікації, автора, місце публікації та інші деталі. Використання метаданих для перевірки джерела допомагає визначити, чи є новина

оригінальною або ж чи була вона скопійована з іншого джерела. Наприклад, програмне забезпечення для перевірки метаданих може аналізувати URL-адреси, дати публікації та інші деталі, щоб визначити, чи є новина оригінальною або ж чи була вона скопійована з іншого джерела.

Функції використання метаданих показані на рис. 1.5.



Рис. 1.5. Функції використання метаданих

Метадані дозволяють відстежити точну дату та час публікації новин. Це допомагає визначити, чи є новина актуальною та своєчасною. Наприклад, новина з недавньою датою публікації може виявитися більш достовірною, ніж новина, яка була опублікована місяці або роки тому. Також можливе виявлення маніпуляцій з датами для створення ілюзії актуальності старих подій [12].

Метадані містять інформацію про автора новини, що дозволяє оцінити його репутацію та надійність. Наприклад, автори, які мають довгу історію достовірних публікацій, є більш надійними. Відомі журналісти та редактори зазвичай мають перевірені профілі, що підвищує довіру до їхніх матеріалів [13].

Метадані допомагають ідентифікувати першоджерело новини. Якщо новина була перепощена або скопійована з іншого ресурсу, це може викликати підозри щодо її достовірності. Аналіз URL-адреси та доменного імені дозволяє визначити репутацію сайту, на якому вперше з'явилася новина [14].

Інколи метадані включають геолокаційні дані, що дозволяє визначити місце створення або публікації новини. Це може бути корисним для перевірки

достовірності подій, описаних у новині, співставляючи місце події з зазначеним у метаданих [15].

Використання метаданих для перевірки джерела новин є потужним інструментом, який дозволяє значно підвищити точність верифікації. Аналіз дат публікації, авторства, джерела публікації та геолокаційних даних допомагає виявляти фейкові новини та забезпечувати довіру до отриманої інформації. Завдяки метаданим можна ефективно боротися з дезінформацією та підтримувати високу якість медіа-контенту.

1.4. Порівняння методів верифікації новин за точністю, швидкістю та масштабованістю

У сучасному цифровому просторі поширення фейкових новин стало значною проблемою, що вимагає ефективних інструментів для їх виявлення. Верифікація новин дозволяє швидко та точно відрізнити правдиву інформацію від недостовірної, допомагаючи споживачам уникати маніпуляцій та дезінформації. Існує низка методів автоматичної верифікації новин, які відрізняються за ключовими характеристиками: точністю, швидкістю обробки та здатністю масштабуватись на великі обсяги даних. Варто провести детальне порівняння таких методів, акцентуючи на їх перевагах та недоліках для ефективного застосування у різних середовищах.

Точність

Точність є ключовим показником якості методів верифікації новин, оскільки саме від неї залежить достовірність отриманих результатів. До найбільш точних методів можна віднести наступні:

- Методи на основі штучного інтелекту (ШІ): Використання алгоритмів машинного навчання та обробки природної мови дозволяє досягати високої точності у виявленні фейкових новин. Алгоритми глибокого навчання, такі як нейронні мережі або моделі трансформерів, показують високу здатність

розпізнавати тонкі маніпуляції у тексті, зокрема, аналізуючи лексичні та граматичні особливості, риторику, емоційні відтінки та структуру новин.

- Методи на основі ручної перевірки: Людська перевірка інформації, як правило, дає найвищий рівень точності, оскільки експерти здатні оцінити контекст новини, її логічну цілісність та відповідність іншим відомим фактам. Проте такий метод є надзвичайно трудомістким і залежить від кваліфікації аналітика.

Швидкість

Швидкість обробки є критично важливою характеристикою, особливо для новинних видань, де оперативність є ключем до успіху. Методи, що забезпечують високу швидкість верифікації, включають:

- Автоматизовані алгоритми: Автоматичні моделі ШІ можуть обробляти великі обсяги текстової інформації в реальному часі, дозволяючи верифікувати новини за кілька секунд. Це особливо корисно для новинних агрегаторів та онлайн платформ, де швидкість має вирішальне значення.
- Методи крос-перевірки джерел: Деякі методи використовують швидкі алгоритми для автоматичного порівняння новин із відомими надійними джерелами. Наприклад, платформи для автоматичної перевірки новин можуть швидко сканувати бази даних перевірених новин та виявляти розбіжності.

Масштабованість

Масштабованість означає здатність системи верифікації працювати з великими обсягами інформації без зниження продуктивності. До методів, які відрізняються високою масштабованістю, належать:

- Хмарні рішення та паралельні обчислення: Системи, які використовують хмарні технології та паралельні обчислення, мають високу здатність до масштабування. Наприклад, платформи Google Fact Check та Microsoft Azure можуть опрацьовувати тисячі запитів одночасно.
- Методи на основі великих мовних моделей (LLM): Моделі, такі як GPT чи BERT, можуть одночасно аналізувати великі обсяги тексту та генерувати

оцінки правдивості для багатьох джерел паралельно, зберігаючи при цьому високу точність навіть при обробці значних обсягів даних.

Автоматична верифікація новин через онлайн сервіси є потужним інструментом для боротьби з фейковими новинами, але вибір конкретного методу верифікації залежить від вимог до точності, швидкості та масштабованості. Методи на основі штучного інтелекту демонструють високу точність і швидкість, проте потребують значних обчислювальних ресурсів для масштабування. Хмарні рішення та паралельні обчислення забезпечують масштабованість, що дозволяє таким методам впроваджуватись у великих інфраструктурах новинних платформ. Таким чином, поєднання різних методів, що використовують сильні сторони кожного з підходів, є оптимальним шляхом для ефективної верифікації новин у сучасному інформаційному середовищі.

1.5. Огляд існуючих онлайн сервісів для верифікації новин

Сучасне інформаційне середовище характеризується швидким поширенням новин, однак із цим виникає й проблема перевірки їх достовірності. Враховуючи велике число недостовірних або навіть фейкових новин, важливим завданням стає автоматична верифікація інформації. Для цього були розроблені спеціалізовані онлайн сервіси, які допомагають користувачам перевіряти факти та джерела інформації. Вони сприяють підвищенню прозорості інформаційного потоку та допомагають боротися з дезінформацією.

1.5.1. Сервіси для верифікації тексту (наприклад, Snopes, FactCheck)

З розширенням цифрових медіа та зростанням кількості новинних джерел, важливо мати доступ до надійних сервісів для верифікації новин. Сервіси для верифікації тексту, такі як Snopes та FactCheck, відіграють важливу роль у боротьбі з фейковими новинами та дезінформацією.

Сервіси для верифікації тексту показані на рис. 1.6.

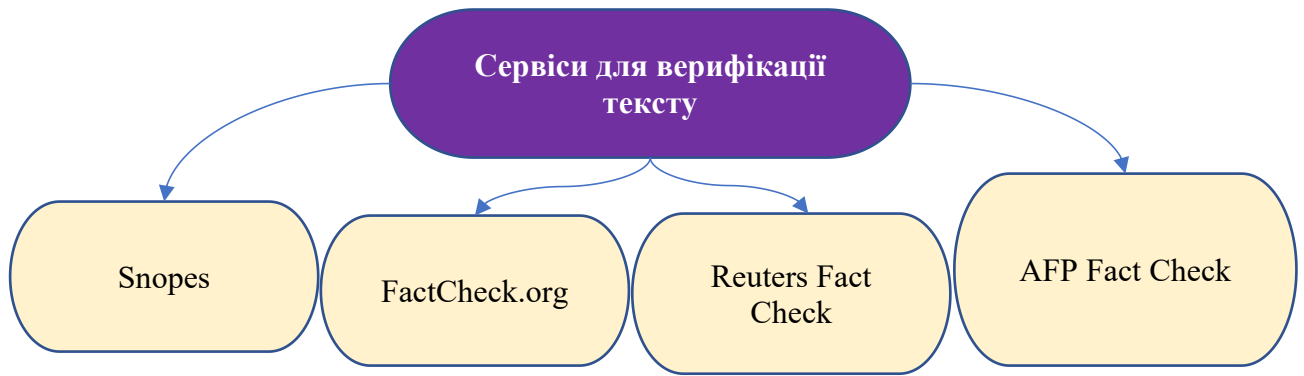


Рис. 1.6. Сервіси верифікації тексту

Snopes є одним з найвідоміших сервісів для верифікації новин. Він займається перевіркою різних інформаційних повідомлень, включаючи інтернет-меми, поширені вірування та новини. Snopes використовує команду експертів, які проводять детальні дослідження для визначення достовірності повідомлень [16][17] (рис. 1.7).

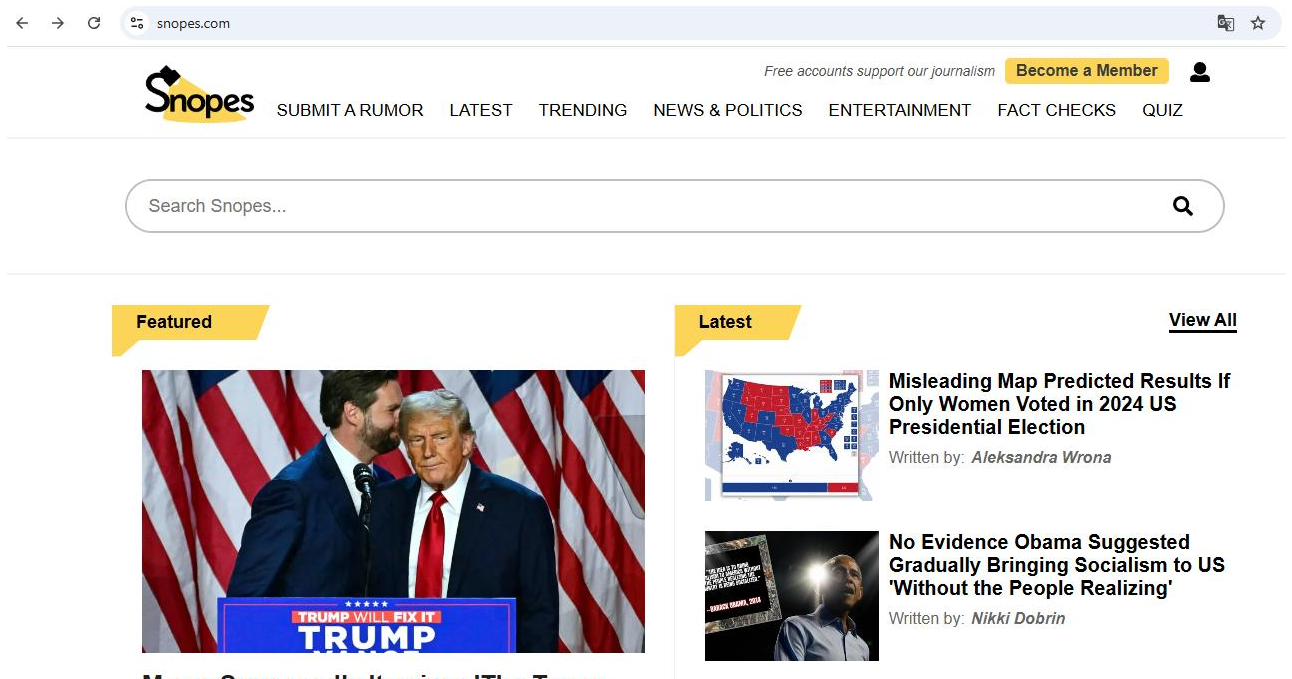


Рис. 1.7. Головна сторінка сайту Snopes.com [17]

FactCheck.org є некомерційним проектом, який спеціалізується на перевірці заяв політиків та новинних повідомлень. Він проводить аналіз заяв, документів та інших джерел інформації для визначення їхньої достовірності [18][19] (рис. 1.8).



Рис. 1.8. Головна сторінка сайту FactCheck.org [19]

Reuters Fact Check є частиною міжнародної медіакомпанії Reuters, яка займається перевіркою новинних повідомлень та боротьбою з дезінформацією. Він використовує команду факт-чекерів, які працюють в різних країнах для перевірки новинних повідомлень [20] (рис. 1.9).

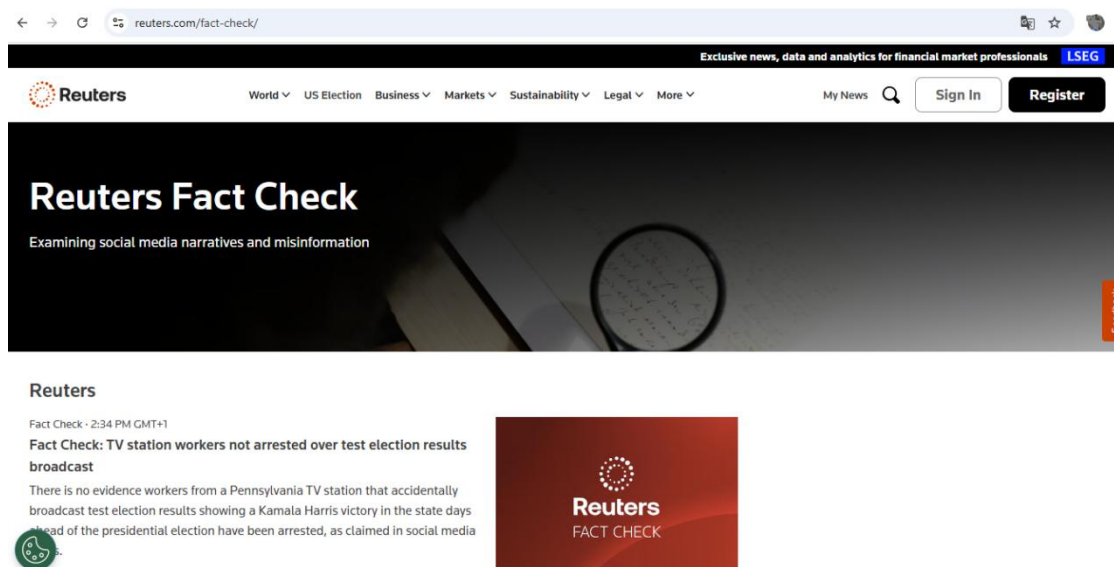


Рис. 1.9. Головна сторінка сайту www.reuters.com [20]

AFP Fact Check є частиною міжнародної медіакомпанії Agence France-Presse (AFP). Він спеціалізується на перевірці новинних повідомлень та боротьбі з дезінформацією, використовуючи команду експертів, які проводять детальні дослідження [21][22] (рис. 1.10).

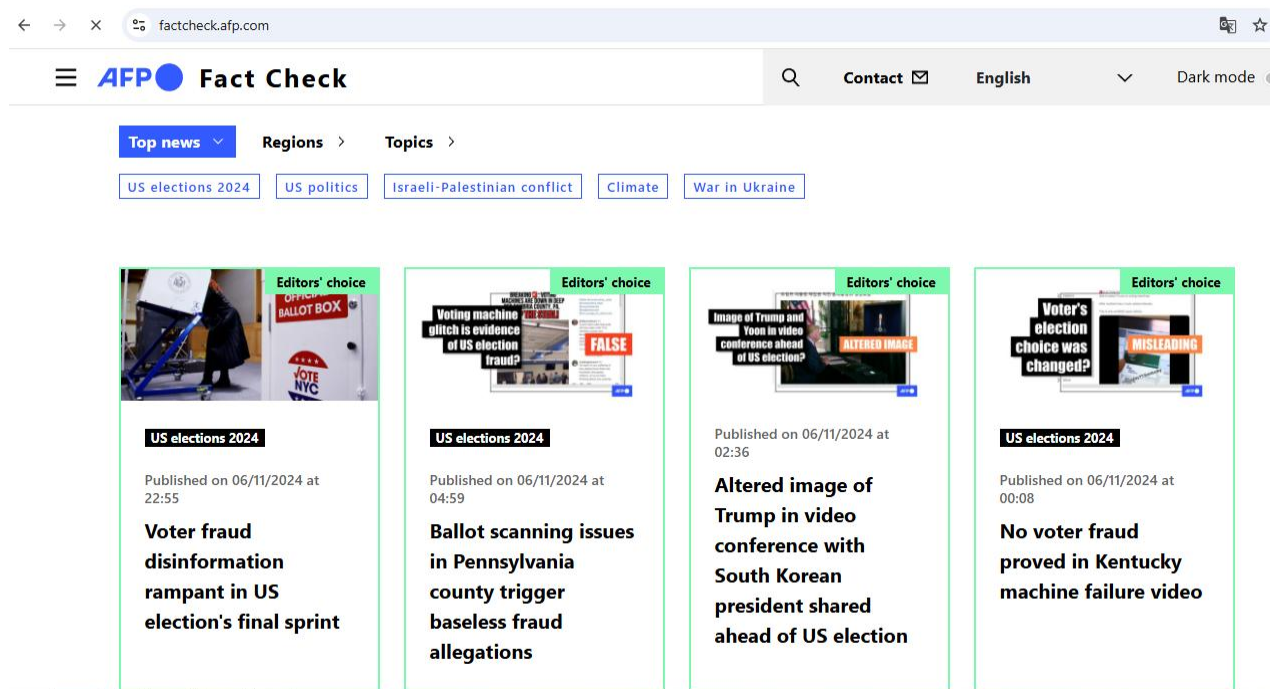


Рис. 1.10. Головна сторінка сайту factcheck.afp.com [22]

1.5.2. Інструменти для аналізу медіа контенту (зображень та відео)

У сучасному медіаполісі, де інформація рухається зі швидкістю світла, важливо мати засоби для верифікації новин та аналізу медіа контенту. Інструменти для аналізу зображень та відео допомагають зрозуміти контекст, виявити фальшиві новини та забезпечити точність інформації.

Інструменти для аналізу медіа контенту показані на рис. 1.11.

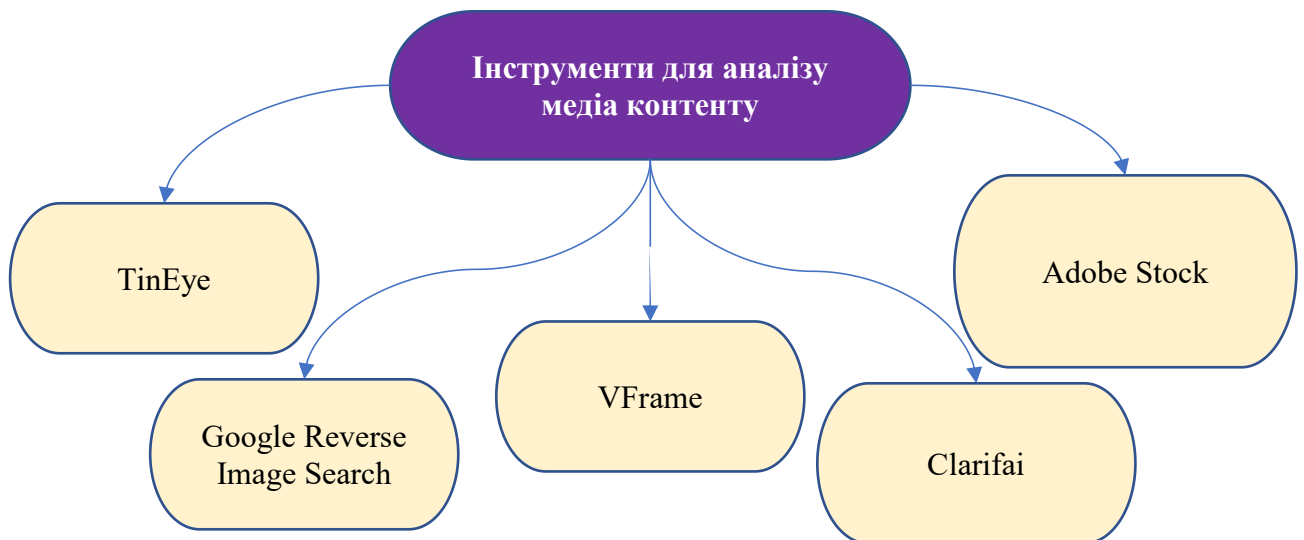


Рис. 1.7. Інструменти для аналізу медіа контенту

TinEye – цей інструмент дозволяє шукати зображення по всій Інтернету, виявляючи подібні зображення та їхні джерела [23] (рис. 1.12).



Рис. 1.12. Головна сторінка сайту tineye.com [23]

Google Reverse Image Search: Цей сервіс дозволяє шукати зображення на всіх платформах Google, що допомагає знайти джерело та контекст зображення [24] (рис. 1.13).

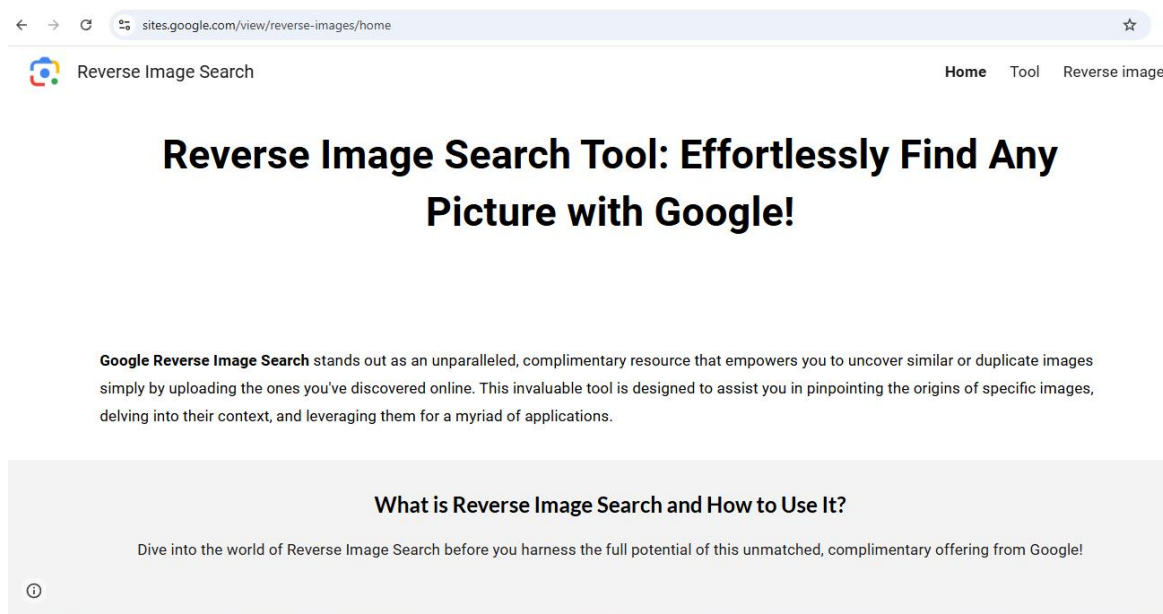


Рис. 1.13. Головна сторінка сайту reverse-images [24]

VFrame – цей інструмент спеціалізується на аналізі відео та зображень, використовуючи штучний інтелект для виявлення ключових моментів та тем [25] (рис. 1.14).

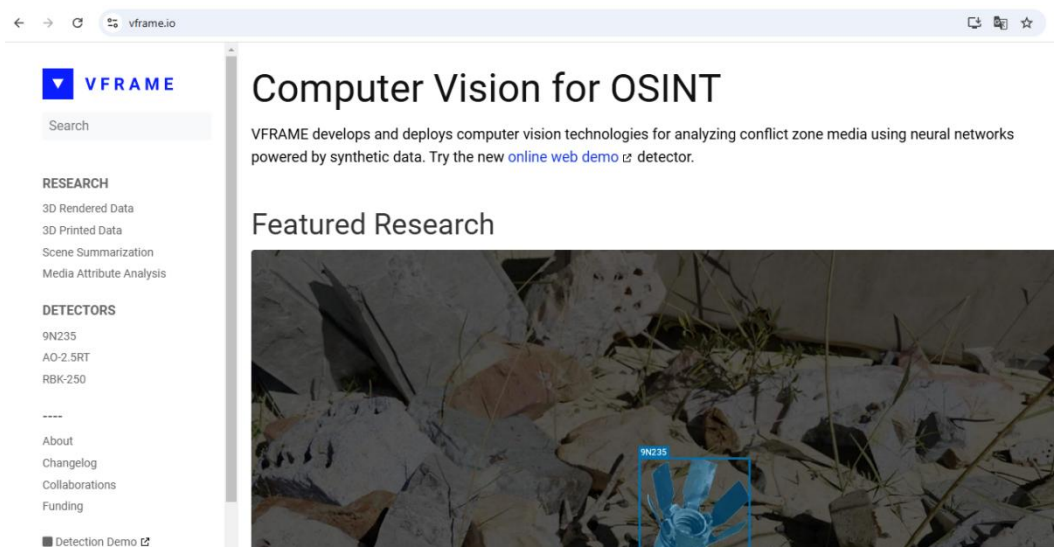


Рис. 1.14. Головна сторінка сайту VFrame [25]

Clarifai – цей сервіс використовує машинне навчання для аналізу зображень та відео, допомагаючи ідентифікувати об'єкти та сцени [26] (рис. 1.15).

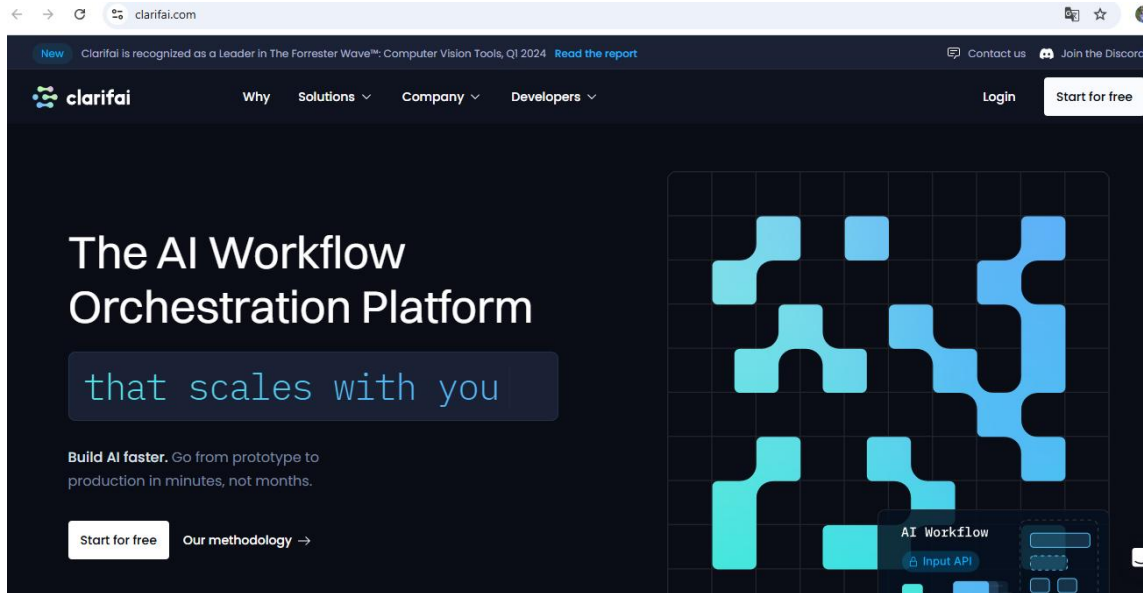


Рис. 1.15. Головна сторінка сайту Clarifai [26]

Adobe Stock – цей інструмент дозволяє шукати зображення та відео в бібліотеці Adobe, що допомагає знайти відповідні джерела та контекст [27] (рис. 1.16).

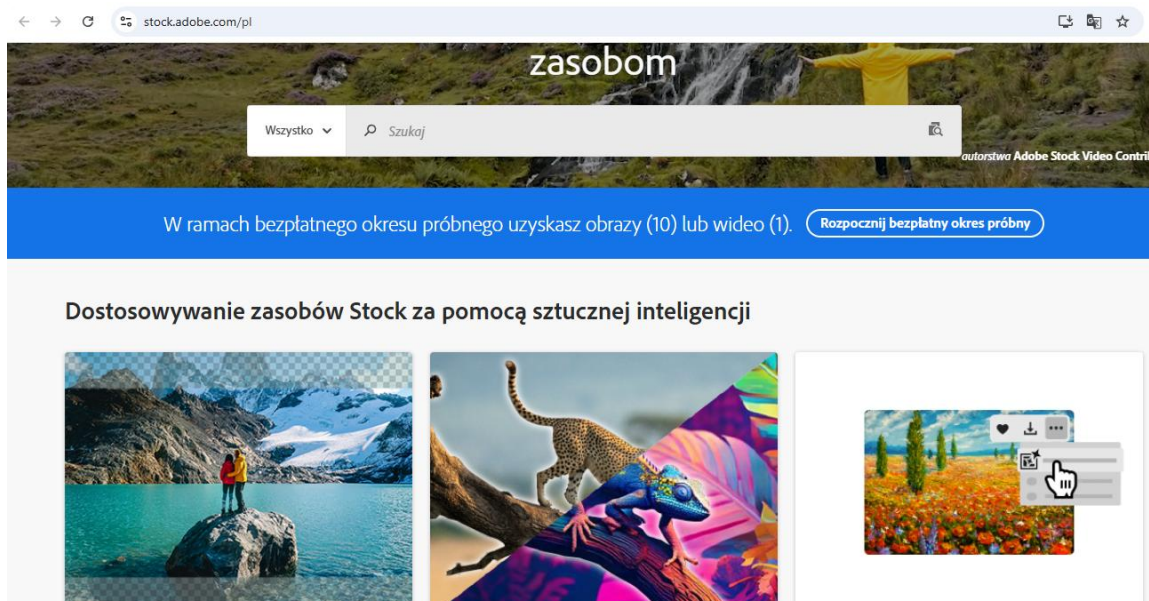


Рис. 1.16. Головна сторінка сайту Adobe Stock [26]

Використання інструментів для аналізу медіа контенту є важливим кроком у боротьбі з фальшивими новинами та забезпеченні точності інформації. Ці інструменти допомагають зрозуміти контекст, виявити джерела та забезпечити достовірність новин

Висновки до розділу 1

Було здійснено аналіз проблеми фальсифікації новин та необхідності автоматичної верифікації інформації, що є надзвичайно важливим в умовах сучасної інформаційної епохи. Фейкові новини становлять серйозну загрозу для суспільства, оскільки вони можуть маніпулювати громадською думкою, посилювати соціальну напругу та навіть впливати на політичні процеси. Для боротьби з цією проблемою потрібні ефективні інструменти та методи, що дозволяють швидко і точно перевіряти достовірність інформації.

Розглянуті основні виклики, пов'язані з фальсифікацією новин, зокрема, швидкість поширення неправдивої інформації, складність ідентифікації джерел та контексту, а також тенденція використання штучних методів для створення перекрученої інформації. Виявлення фейкових новин потребує комплексного підходу, що включає в себе використання як технологічних, так і лінгвістичних методів.

Огляд існуючих методів верифікації новин показав, що найбільш перспективними є підходи, які поєднують машинне навчання та лінгвістичний аналіз. Машинне навчання дозволяє створювати моделі, здатні автоматично виявляти неточності та маніпуляції в текстах, тоді як лінгвістичні методи можуть допомогти у виявленні стилістичних та семантичних ознак, характерних для фейкових новин. Використання метаданих також є важливим інструментом для перевірки достовірності джерел та підтвердження фактів.

Порівняння методів верифікації за різними критеріями вказує на важливість балансу між точністю, швидкістю та масштабованістю технологій. Ідеальний метод повинен бути достатньо швидким для оперативної реакції, але водночас забезпечувати високу точність результатів.

Особливу увагу було приділено огляду існуючих онлайн сервісів, що вже активно використовуються для перевірки новин. Сервіси, такі як Snopes та FactCheck, ефективно справляються з перевіркою текстів на наявність фейкової інформації, в той час як інструменти для аналізу медіа контенту дозволяють визначати автентичність зображень та відео, що є важливою частиною верифікації в умовах поширення мультимедійних фальсифікацій.

Отже, розділ дає чітке уявлення про необхідність застосування автоматичних інструментів для верифікації новин, наводить існуючі підходи та сервіси, що забезпечують ефективну перевірку інформації, та формулює основні напрямки для подальших досліджень і вдосконалення в цій сфері.

РОЗДІЛ 2.

ПРОЕКТУВАННЯ СИСТЕМИ ДЛЯ АВТОМАТИЧНОЇ ВЕРИФІКАЦІЇ НОВИН

У сучасному інформаційному середовищі значна кількість новинних матеріалів та статей публікується онлайн щоденно, і їхня достовірність залишається ключовим питанням як для журналістів, так і для звичайних користувачів мережі Інтернет. Завдяки розвитку технологій штучного інтелекту з'явилася можливість створити системи, які автоматично аналізують новинний контент і допомагають користувачам оцінити його на предмет достовірності. Важливим нюансом проекту є архітектура такої системи автоматичної верифікації новин, та опис її основних компонентів та логіку функціонування. Увагу приділено структурі алгоритмів, які здійснюють перевірку фактів, аналіз емоційності тексту, виявлення маніпулятивних елементів, а також процесу надання рекомендацій щодо оцінки достовірності новинного матеріалу.

2.1. Огляд архітектури системи для автоматичної верифікації новин

Автоматична система верифікації новин є багатофункціональною структурою, що поєднує аналіз тексту, визначення маніпуляційних маркерів і надання оцінок на основі наявності достовірних джерел та емоційних елементів. Система розроблена на базі сучасних методів обробки природної мови (NLP), з використанням моделей глибокого навчання для аналізу текстового контенту. Ключові елементи архітектури цієї системи представлені на рис. 2.1.



Рис. 2.1. Ключові елементи архітектури системи автоматизованої верифікації новин

Підготовка текстових даних для аналізу

Цей етап включає завантаження новинного матеріалу з URL, обробку заголовка та основного тексту статті. Кожна з цих частин оцінюється окремо, щоб визначити загальну достовірність новини. Заголовок зазвичай має сильний вплив на перше враження читача і може містити маніпуляційні елементи, тож він оцінюється із застосуванням окремої ваги.

Аналіз заголовка і основного тексту

Використовуючи модель машинного навчання, система генерує оцінки для кожного класу, наприклад "достовірний" та "фейковий". Оцінки згодом зважуються — 30% для заголовка та 70% для основного тексту, що дозволяє зробити акцент на змістовному аналізі всієї статті, при цьому враховуючи значущість заголовка. Результати аналізу заголовка і тексту зберігаються у вигляді оцінок, що відображають ймовірність достовірності кожної частини.

Додаткові фактори достовірності

Система використовує кілька лінгвістичних характеристик для оцінки достовірності:

- *Наявність дат, чисел, цитат і джерел.* Присутність конкретних дат, числових даних, цитат та посилань на джерела свідчить про більш високий рівень достовірності тексту. Такі елементи збільшують довіру до матеріалу, бо вони зазвичай вказують на фактологічну складову.
- *Оцінка емоційності тексту.* Емоційний рівень може свідчити про маніпуляційний характер тексту. Використання емоційних слів, як "жахливий", "приголомшливий", "терміново", може вказувати на наявність маніпулятивних елементів у тексті.

Розрахунок фінальної оцінки достовірності

Після обробки основного тексту і заголовка з використанням вагових коефіцієнтів та врахування додаткових факторів система обчислює загальну оцінку достовірності. Додаткові фактори впливають на підсумковий результат у позитивну сторону, якщо в тексті є дати, числові дані, цитати і джерела. Водночас емоційна забарвленість знижує остаточну оцінку достовірності, що дозволяє знизити ризик поширення неправдивої інформації.

Виявлення маніпулятивних маркерів

Система здійснює аналіз на наявність маніпулятивних виразів та зворотів у тексті. На цьому етапі система перевіряє, чи є в тексті конкретні вирази, які часто використовуються для привернення уваги чи створення у читача емоційного сприйняття ситуації. Маніпулятивні маркери обмежують об'єктивність інформації і є частим елементом фейкових новин.

Формування вердикту

Фінальний результат представлено у вигляді вердикту, що містить один із трьох рівнів достовірності: "ймовірно правда", "сумнівна достовірність" та "ймовірно фейк". Додатково, на основі отриманих оцінок, система надає докладні відомості про причини вердикту, що містять перелік факторів, таких як наявність конкретних дат, цитат, чисел, а також рівень емоційності.

Рекомендації для користувача

Система надає рекомендації на основі підсумкової оцінки достовірності. Якщо результат аналізу вказує на низьку достовірність, користувачу пропонується перевірити інформацію в офіційних джерелах, а також утриматися від поширення матеріалу без додаткової перевірки. Якщо достовірність є сумнівною, рекомендується провести додаткову перевірку та порівняти новину з іншими джерелами. При високій оцінці достовірності система дозволяє рекомендувати контент до поширення, але також вказує на необхідність перевірки додаткових джерел.

Таким чином, автоматична система верифікації новин будується на комплексному аналізі тексту, що враховує не лише загальні характеристики контенту, але і його специфічні лінгвістичні параметри та маніпуляційні елементи. Основна мета системи — надати користувачам зручний інструмент для оперативної перевірки новинного контенту. Завдяки цьому підходу користувачі отримують можливість краще орієнтуватися у інформаційному просторі, підвищуючи власну медіаграмотність та критичне мислення.

2.2. Вибір алгоритмів і методів для класифікації новин

Розробка автоматизованої системи для верифікації новин є важливим етапом у забезпеченні надійності інформації, доступної для широкого загалу. Така система має бути здатною ефективно оцінювати достовірність джерел та виявляти потенційно маніпулятивний чи неправдивий контент. Вибір алгоритмів та методів для класифікації новин є ключовим аспектом, оскільки саме вони визначають точність, швидкість та ефективність системи в процесі оцінки новинного матеріалу.

2.2.1. Методи на основі обробки природної мови (NLP)

Сучасні технології обробки природної мови (Natural Language Processing, NLP) мають значний потенціал для автоматичної оцінки достовірності текстів, зокрема новин. В епоху швидкого розповсюдження інформації, зростає потреба у розробці систем, які можуть швидко та об'єктивно оцінювати правдивість і можливу маніпулятивність повідомлень. Однією з таких систем є платформа для верифікації новин, яка використовує методи NLP, машинне навчання та розпізнавання маркерів маніпуляції для автоматичної оцінки достовірності повідомлень.

Система автоматичної верифікації новин використовує нейронну модель для аналізу текстових даних, що дозволяє враховувати різні аспекти достовірності та маніпулятивності тексту, що представлені в таблиці 2.1.

Таблиця 2.1

Аспекти достовірності та маніпулятивності тексту

№	Аспекти	Опис
1	2	3
1	Аналіз заголовку і тексту	Основою системи є модель нейронної мережі, яка здатна обробляти заголовок і текст новини окремо, після чого комбінувати їх оцінки Система аналізує заголовок для виявлення ознак маніпуляцій, перебільшень або емоційно забарвлених фраз, які можуть впливати на загальне сприйняття тексту. Такий підхід сприяє підвищенню точності оцінок, адже, як правило, маніпуляції починаються саме з заголовків Заголовок і текст отримують оцінки достовірності, які потім об'єднуються з різними ваговими коефіцієнтами, що дозволяє системі враховувати різні фактори впливу заголовка і змісту статті

1	2	3
2	Розрахунок достовірності на основі додаткових факторів	Система використовує допоміжний модуль для оцінки додаткових факторів, що впливають на достовірність тексту. Враховуються такі елементи: – <i>Дати</i>: аналізує наявність дат, що сприяє підвищенню достовірності, оскільки конкретні дати зазвичай надають тексту більше правдивості – <i>Числа та статистичні дані</i>: наявність числових даних часто свідчить про те, що інформація може бути об'єктивною – <i>Цитати</i>: присутність прямих цитат є індикатором достовірності, оскільки цитати відображають слова реальних осіб чи джерел – <i>Джерела</i>: згадки офіційних джерел, організацій чи посадових осіб надають тексту більше обґрунтованості – <i>Емоційний рівень</i>: емоційна забарвленість тексту знижує його достовірність, оскільки часто маніпуляції мають емоційний характер
3	Коригування оцінки достовірності	Оцінка достовірності піддається коригуванню на основі позитивних та негативних факторів. Наприклад, наявність чисел, цитат і джерел підвищує оцінку достовірності, тоді як емоційність тексту знижує її Такі коригування дозволяють враховувати нюанси інформації, зокрема її науковість, фактичність і стиль викладу
4	Аналіз на маніпуляцію та визначення маніпулятивних маркерів	Модель також може виявляти ознаки маніпуляції, аналізуючи текст на предмет маркерів, таких як перебільшення, заклики до негайної дії, надмірна емоційність або інші мовні звороти, що можуть впливати на читача Якщо такі ознаки знайдені, система знижує загальну оцінку достовірності тексту та попереджає користувача про можливу маніпуляцію

1	2	3
5	Формулювання вердикту	Система автоматично класифікує новину за трьома рівнями достовірності: – «Ймовірно правда» — при високій достовірності – «Сумнівна достовірність» — коли новина вимагає додаткової перевірки – «Ймовірно фейк» — при низькій достовірності Кожен із вердиктів доповнюється деталями, що надають користувачеві додаткову інформацію про причини класифікації, наприклад, емоційний характер тексту або відсутність джерел
6	Генерація рекомендацій	Залежно від рівня достовірності та наявності маніпулятивних ознак, система надає рекомендації щодо подальших дій. Наприклад, вона може рекомендувати користувачеві перевірити інформацію у надійних джерелах або утриматися від поширення новини без додаткової перевірки Цей модуль також може наголошувати на перевірці конкретних деталей, таких як джерела або цитати, що робить систему зручним інструментом для об'єктивної оцінки новин

Розроблена система для автоматичної верифікації новин, заснована на методах NLP, здатна ефективно аналізувати текст новин, оцінювати їх достовірність та виявляти маніпуляції. Інтеграція додаткових факторів оцінки, таких як наявність дат, цитат і джерел, забезпечує всебічний підхід до верифікації новин, що значно підвищує її надійність.

2.2.2. Моделі машинного навчання для класифікації фейкових новин

Автоматизація процесу верифікації новин сьогодні є важливим кроком у боротьбі з фейковою інформацією. Виявлення фейкових новин вимагає не тільки аналізу тексту та визначення ключових факторів достовірності, але й здатності враховувати маніпулятивний контекст і зміст повідомлення. Необхідно розглянути процес проектування системи автоматичної верифікації новин на основі машинного навчання. Для цього слід детально розглянути використання нейронних мереж, алгоритмів обробки тексту та аналізу

маніпуляцій, які дозволяють оцінити достовірність новинного контенту, підвищуючи об'єктивність інформаційного простору.

Основні етапи аналізу достовірності новин

Процес оцінки достовірності новин за допомогою машинного навчання в системі передбачає кілька ключових етапів (рис. 2.2).

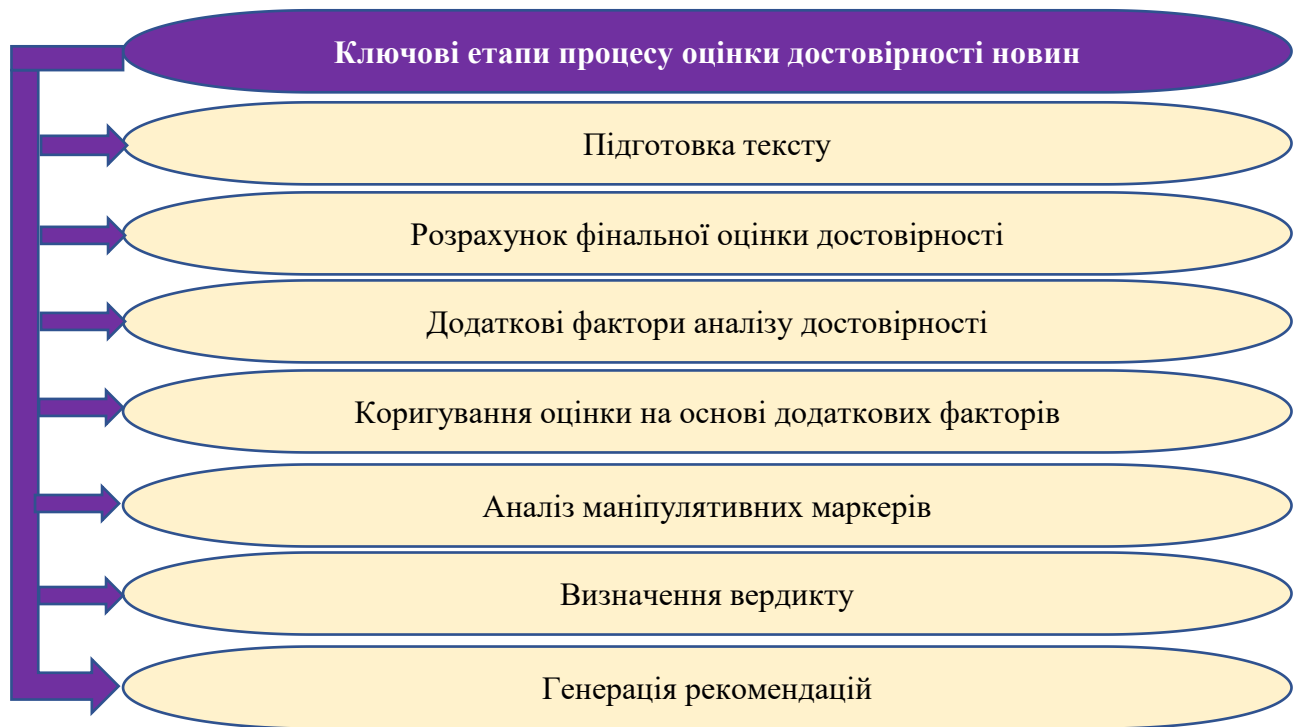


Рис. 2.2. Ключові етапи процесу оцінки достовірності новин

Підготовка тексту

Система починає з аналізу тексту та заголовка новини. Код передбачає вилучення заголовка та тексту з об'єкта `content`, що забезпечує зручне форматування тексту та його підготовку для подальшого оброблення. Після вилучення тексту здійснюється передобробка даних для подальшої їх обробки моделлю нейронної мережі.

При реалізації коду враховано, що `title_result` та `text_result` відповідають результатам аналізу достовірності заголовку та тексту, відповідно. Для цього застосовано попередньо натреновану модель (`self.model`), яка здатна оцінити

ймовірність достовірності або фейковості новини на основі текстових характеристик.

Розрахунок фінальної оцінки достовірності

У системі передбачено, що заголовок і текст новини мають різну вагу під час оцінки достовірності. На прикладі значень $\text{title_weight} = 0.3$ та $\text{text_weight} = 0.7$ встановлено, що основний акцент робиться на змісті новини. Відповідно до наданого коду, фінальна оцінка достовірності новини (credibility_score) формується за зваженою сумою оцінок достовірності заголовку та тексту.

Такий підхід виправданий, оскільки заголовки часто можуть бути емоційно насиченими або привертати увагу, тому їхня достовірність може бути нижчою, ніж у основного тексту. Однак заголовок все ж враховується у процесі класифікації.

Додаткові фактори аналізу достовірності

Система додатково аналізує новини на наявність певних показників достовірності, які подані в таблиці 2.2.

Таблиця 2.2

Додаткові фактори аналізу достовірності

№	Фактори	Опис
1	Дати	пошук точних дат, які можуть свідчити про хронологічну відповідність подій у тексті. Наявність дати підвищує довіру до новини
2	Числа та статистика	числові дані, відсотки та конкретні суми підвищують достовірність тексту, оскільки створюють враження обґрунтованості фактів
3	Цитати	система шукає пряму мову, яка свідчить про наявність конкретних джерел інформації. Цитати з офіційних джерел або особистостей додають довіри до матеріалу
4	Джерела	наявність посилань на офіційні джерела чи цитати підвищує рівень достовірності
5	Рівень емоційності	аналіз рівня емоційних висловлювань дозволяє оцінити маніпулятивний потенціал тексту. Висока емоційність тексту знижує його достовірність

Коригування оцінки на основі додаткових факторів

Результати аналізу додаткових факторів використовуються для коригування оцінки достовірності. Додаткові позитивні фактори, такі як наявність джерел або цитат, підвищують остаточну оцінку, тоді як висока емоційність знижує її. Використовується зважений підхід, щоб балансувати вплив кожного фактора на остаточний результат.

Аналіз маніпулятивних маркерів

Окрім факторів достовірності, система також визначає наявність маніпулятивних зворотів у тексті, які можуть свідчити про можливу упередженість або спробу впливати на емоції читача. Маркери маніпуляції знижують загальну оцінку достовірності новини, в результаті чого фінальна оцінка залишається у межах, що свідчать про потенційну упередженість тексту.

Визначення вердикту

На основі отриманої оцінки достовірності система формує вердикт. Варіанти вердикту представлені в таблиці 2.3

Таблиця 2.3

Варіанти вердикту

№	Вердикт	Пояснення
1	 Ймовірно правда	для новин, що мають високу оцінку достовірності
2	 Ймовірно фейк	для новин, які мають низьку оцінку
3	 Сумнівна достовірність	для текстів, оцінка яких вказує на необхідність додаткової перевірки

Кожен з вердиктів доповнюється деталями, які описують, на чому базується рішення системи. Ці деталі дозволяють користувачам краще зрозуміти, чому система прийняла саме таке рішення.

Генерація рекомендацій

На основі аналізу новини система формує рекомендації для користувача. Наприклад, для новин з низьким рівнем достовірності система радить

перевірити інформацію в офіційних джерелах та утриматись від поширення, доки не буде підтвердження. Рекомендації також надаються для новин із нейтральним вердиктом або текстів з маніпулятивними зворотами, щоб користувачі могли перевірити джерела та бути уважними до потенційних маніпуляцій.

Процес автоматичної верифікації новин є складним і багатокомпонентним завданням, яке потребує інтеграції різноманітних методів обробки тексту та аналізу достовірності. В описаному коді розроблено алгоритм, який оцінює новину з різних перспектив: достовірність джерел, наявність маніпуляцій та ключових текстових маркерів. Така модель машинного навчання є цінним інструментом для медіаграмотності та забезпечення якості інформаційного простору.

2.2.3. Використання нейронних мереж для виявлення фейкових новин

Метою підсистеми є автоматизація процесу перевірки достовірності новинних матеріалів на основі аналізу тексту, виявлення маніпулятивних прийомів, розрахунку рівня достовірності, а також надання рекомендацій. Основою є багаторівневий підхід до аналізу змісту новини: окремо аналізуються заголовок і текст, оцінюються додаткові фактори достовірності, враховуються маніпулятивні мовні засоби та коригуються оцінки достовірності з урахуванням різних аспектів тексту.

Нейронні мережі, а саме трансформери (наприклад, моделі BERT, RoBERTa, GPT), останнім часом довели свою високу ефективність у задачах обробки природної мови. Вони здатні обробляти величезні масиви текстових даних і виявляти приховані шаблони та аномалії в текстах, що є надзвичайно корисним для аналізу достовірності інформації. На основі методу можна відзначити низку кроків, що є характерними для системи перевірки новин за допомогою нейронних мереж.

Структура процесу верифікації новин показана на рис. 2.3.



Рис. 2.3. Структура процесу верифікації новин

Підготовка тексту новини для аналізу

На першому етапі, система отримує текст новини для аналізу через вказане посилання (URL), що дозволяє витягти заголовок та основний текст статті. Попередня обробка даних передбачає очищення заголовку та тексту від можливих зайвих пробілів і некоректних символів. Важливо, щоб вхідні дані мали відповідний формат, що мінімізує можливість хибних оцінок у подальшій обробці.

Аналіз заголовку і тексту новини окремо

Система здійснює окремий аналіз заголовку та тексту, щоб врахувати їх специфіку у контексті можливих маніпуляцій. Заголовок має високий рівень впливу на сприйняття новини, тому при оцінці достовірності заголовку

приділяється увага можливому емоційному забарвленню та використанню маніпулятивних прийомів. На цьому етапі також використовується модель, що надає ймовірності для різних класів (наприклад, “правда” і “фейк”), дозволяючи зробити попередній висновок про достовірність кожного з елементів окремо.

Обчислення фінальної оцінки достовірності новини

Для визначення загальної достовірності новини застосовуються вагові коефіцієнти до оцінок заголовку і тексту, що становлять 0.3 та 0.7 відповідно. Важливість тексту як основного джерела інформації оцінюється вище, що дозволяє системі приділяти більше уваги змісту новини. Фінальна оцінка обчислюється шляхом множення ймовірностей достовірності заголовку та тексту на їх відповідні вагові коефіцієнти та підсумування результатів, що дозволяє отримати інтегровану оцінку, яка є базовою для подальшої корекції.

Визначення додаткових факторів для оцінки достовірності

Система застосовує кілька додаткових факторів для детальнішої оцінки достовірності, зокрема:

- *Наявність дат*: текст перевіряється на присутність конкретних дат і часових позначок, які зазвичай підвищують достовірність новини.
- *Наявність числових даних*: числові показники та статистика додають об’єктивності тексту, оскільки можуть бути перевірені або підтверджені.
- *Присутність цитат*: цитати надають автентичності, особливо якщо йдеться про конкретних осіб або організації.
- *Посилання на джерела*: відсутність або наявність посилань на конкретні джерела є важливим фактором, який впливає на оцінку достовірності.
- *Рівень емоційного забарвлення*: висока емоційність може свідчити про маніпулятивність тексту, тому її рівень враховується як негативний чинник.

Коригування оцінки достовірності на основі додаткових факторів

Коригування основної оцінки достовірності відбувається через додавання позитивних або віднімання негативних балів, отриманих у попередньому кроці. Додаткові позитивні бали за наявність дат, чисел, цитат або джерел підвищують

достовірність, тоді як емоційне забарвлення навпаки знижує її. Завдяки цьому процесу система досягає більшої точності у визначенні вірогідності новини.

Виявлення маніпулятивних мовних засобів

Система здійснює перевірку тексту на наявність маніпулятивних маркерів, таких як емоційно насичені слова, заклики до негайної дії або використання надзвичайної лексики. Виявлення таких мовних засобів сприяє зменшенню фінальної оцінки достовірності та попереджає користувача про можливе маніпулювання.

Формування остаточного вердикту та рекомендацій

На основі фінальної оцінки достовірності система надає один із трьох вердиктів: “Ймовірно правда”, “Ймовірно фейк” або “Сумнівна достовірність”. При цьому до кожного вердикту додаються специфічні деталі, що пояснюють, чому було прийняте саме таке рішення. Додатково надаються рекомендації для подальших дій: чи слід перевірити інформацію в інших джерелах, утриматись від поширення або, навпаки, можна вважати інформацію достовірною.

Таким чином, розроблена система автоматичної верифікації новин на основі нейронних мереж дозволяє ефективно аналізувати текстові матеріали, враховуючи різні фактори, які впливають на оцінку достовірності. Використання технологій штучного інтелекту дозволяє здійснювати перевірку новин автоматично, зменшуючи ризики дезінформації та підвищуючи обізнаність користувачів про потенційно фейковий контент.

2.3. Вибір бази даних для зберігання новин, результатів верифікації та аналітичних даних

У процесі створення системи автоматичної верифікації новин важливим етапом є вибір бази даних для зберігання контенту новин, результатів верифікації та пов'язаних аналітичних даних. База даних має забезпечувати ефективне управління великими обсягами текстової інформації та підтримувати

роботу системи з низькою затримкою, швидким доступом до даних, можливістю масштабування та надійністю зберігання.

Основні вимоги до бази даних для зберігання новин

База даних повинна відповідати кільком критичним вимогам, пов'язаним з обробкою текстової інформації та аналітичних результатів. Перш за все, вона має дозволяти зберігання великих обсягів неструктурованих даних, оскільки новинні статті представлені в текстовому форматі, що може містити заголовки, текстові блоки, дати, посилання на джерела тощо. Крім того, важлива наявність можливостей фільтрації та пошуку для зручного доступу до конкретних новин за ключовими словами або показниками верифікації. Щоб забезпечити коректне функціонування алгоритмів штучного інтелекту, база даних також повинна підтримувати швидкий доступ до інформації та гнучке масштабування, що дозволить обробляти постійно зростаючий обсяг інформації.

Переваги використання реляційних та нереляційних баз даних

Реляційні бази даних підходять для зберігання даних із структурованою інформацією, що добре адаптується для ведення реєстру новин, збереження інформації про дату публікації, URL-адресу джерела, а також основних метрик аналізу та верифікації. Серед популярних реляційних баз даних, які можуть бути використані для таких цілей, варто виділити MySQL та PostgreSQL, які забезпечують надійне зберігання та мають високу швидкість вибірки інформації. Однак, для аналізу великих обсягів неструктурованих текстів вони можуть бути не настільки гнучкими через обмеженість у підтримці складної текстової обробки.

Нереляційні бази даних такі як MongoDB або Elasticsearch, пропонують гнучкішу структуру, що дозволяє ефективно зберігати неструктуровані дані. Це особливо корисно для системи, яка працює з великим обсягом тексту та вимагає високої швидкості пошуку. Наприклад, Elasticsearch дозволяє швидко індексувати текстові дані, здійснювати пошук за ключовими словами та фразами, а також легко розширюється за допомогою масштабування у горизонтальному напрямку. Така гнучкість стає важливою для підтримки роботи системи автоматичної верифікації, яка вимагає швидкого доступу до великих обсягів текстових даних для обробки, аналізу та оновлення інформації.

Розподілене зберігання та масштабованість

Система автоматичної верифікації новин може обробляти великі обсяги даних, тому вибір бази даних з можливістю масштабування є необхідним. Нереляційні бази даних, такі як MongoDB та Elasticsearch, підтримують розподілене зберігання, що забезпечує балансування навантаження між серверами та можливість масштабування системи. Це особливо важливо для інтеграції з потоковими джерелами інформації та новин, коли необхідно постійно оновлювати та аналізувати нові дані. Наприклад, MongoDB дозволяє створювати репліки та масштабувати базу шляхом додавання нових вузлів, що допомагає підтримувати продуктивність при збільшенні обсягу даних та запитів до системи.

Безпека зберігання та обробки інформації

Оскільки система верифікації працює з великою кількістю новинних матеріалів, важливо забезпечити захист даних, включно з механізмами автентифікації та контролю доступу. Реляційні бази даних, такі як PostgreSQL, пропонують вбудовані засоби для налаштування різних рівнів доступу до даних, шифрування з'єднань та захисту від несанкціонованого доступу. Крім того, необхідно забезпечити резервне копіювання даних, щоб уникнути втрати інформації у випадку збоїв.

Вибір бази даних є критично важливим для забезпечення стабільної та ефективної роботи системи автоматичної верифікації новин. Залежно від вимог до структури, масштабу та швидкості доступу до інформації, можна вибрати реляційні або нереляційні рішення. Реляційні бази даних, такі як PostgreSQL, надають можливість структурованого зберігання та високий рівень безпеки, тоді як нереляційні бази даних, такі як Elasticsearch та MongoDB, забезпечують більшу гнучкість, швидкість пошуку та масштабованість, що робить їх оптимальними для зберігання великих обсягів текстових даних та оперативного аналізу.

2.4. Інтеграція із зовнішніми сервісами для отримання додаткової інформації (API соціальних мереж, новинних агентств)

Інтеграція із зовнішніми сервісами є важливим елементом системи автоматичної верифікації новин, адже саме завдяки їй можливо отримати додаткові дані, які підвищують точність аналізу та достовірність результатів. API (Application Programming Interface) дозволяють системам взаємодіяти одна з одною через інтернет, автоматично обмінюючись інформацією без втручання користувачів. Сервіси соціальних мереж і новинних агентств надають свої дані через API, що дає змогу новинній системі швидко верифікувати контент, оцінити його на основі популярності, поширеності та зворотного зв'язку користувачів, а також врахувати репутацію джерела.

Таким чином, важливо розглянути аспекти інтеграції, процеси отримання та обробки інформації через API, значення їхнього використання у контексті верифікації новин, а також переваги й потенційні ризики цього підходу.

Інтеграція зовнішніх джерел через API соціальних мереж та новинних агентств виконує дві головні функції:

- *Отримання актуальної інформації* — системи через API можуть отримувати контекстні дані про новини, їхній зміст, часові аспекти та охоплення аудиторії. Це дозволяє скоротити час на верифікацію новини, особливо якщо вона активно обговорюється на різних платформах.
- *Визначення популярності та надійності джерела* — аналізуючи соціальну взаємодію, зокрема лайки, поширення та коментарі, система отримує відомості про суспільний резонанс новини та її репутацію.

Цей підхід є надзвичайно ефективним, оскільки дозволяє системі одразу звернутися до авторитетних зовнішніх джерел і використати інформацію для об'єктивного оцінювання достовірності, а також уникнути потреби в людському аналізі на початковому етапі.

Застосування API соціальних мереж у процесі верифікації новин

API соціальних мереж, таких як Twitter, Facebook або Instagram, дозволяють збирати і аналізувати великі обсяги даних щодо обговорень, поширень та коментарів до певної новини. Можливості застосування API соціальних мереж у процесі верифікації новин представлені в таблиці 2.4.

Таблиця 2.4

Можливості застосування API соціальних мереж у процесі верифікації новин

№	Можливості	Пояснення
1	Аналізувати популярність новини	визначати рівень активності користувачів на різних платформах. Якщо новина активно поширюється, ймовірність її достовірності вища, особливо якщо вона підтримується офіційними сторінками чи профілями відомих експертів
2	Враховувати емоційну реакцію та контекст	API соціальних мереж надають дані, які допомагають виявити емоційні реакції користувачів, що може слугувати індикатором надійності чи маніпулятивності новини
3	Моніторинг репутації джерела	якщо новина публікується сторінками або особистостями з високою репутацією, її ймовірність виявитися фейковою значно знижується. Соціальні мережі надають відомості про перевірені профілі, що автоматично додає новині ваги

Застосування API новинних агентств для підтвердження достовірності

API новинних агентств, як-от Reuters, BBC, Associated Press, дозволяють отримувати верифіковані факти, які можуть бути використані для перевірки ключових моментів новини. Можливості застосування API новинних агентств подані в таблиці 2.5.

**Можливості застосування API новинних агентств для підтвердження
достовірності**

№	Можливості	Пояснення
1	Перевірка фактичної інформації	підтвердження дат, місць подій, учасників тощо. Якщо інформація збігається з даними з авторитетного новинного агентства, достовірність новини підвищується
2	Перевірка ексклюзивності новини	якщо новина є унікальною та не має підтвердження з інших джерел, виникає більше сумнівів у її надійності
3	Визначення маніпуляцій	якщо новинні агентства описують подію з меншим рівнем емоційного забарвлення, а аналізована новина використовує емоційні звороти, це може свідчити про маніпулятивний характер

Переваги та ризики використання API у верифікації новин

Переваги подані в таблиці 2.6

Переваги використання API у верифікації новин

№	Переваги	Опис
1	Швидкість та ефективність	інтеграція з API дає змогу оперативно отримувати й аналізувати великі обсяги інформації, що суттєво знижує час на перевірку новини
2	Об'єктивність	використання зовнішніх джерел забезпечує об'єктивну оцінку на основі актуальних даних і суспільного резонансу
3	Підвищення точності	перевірка надійності новини через кілька незалежних джерел мінімізує ризик поширення фейків

Ризики подані в таблиці 2.7.

Таблиця 2.7

Ризики використання API у верифікації новин

№	Ризики	Опис
1	Залежність від якості даних	неточні чи неактуальні дані з API можуть негативно впливати на результат оцінки
2	Ризик надмірної емоційності	соціальні мережі часто характеризуються емоційними реакціями, що може викривити загальну картину новини
3	Конфіденційність та безпека	оскільки інформація передається через API, це вимагає захисту конфіденційних даних користувачів, а також належного управління доступом до API

Інтеграція з API соціальних мереж та новинних агентств є незамінною складовою сучасних систем верифікації новин. Вона дозволяє швидко і точно отримати важливу інформацію для підтвердження або спростування новини. Ретельно розроблена система, яка використовує API з дотриманням принципів безпеки та об'єктивності, забезпечує ефективну верифікацію і сприяє боротьбі з дезінформацією. Однак, попри всі переваги, важливо враховувати можливі ризики, пов'язані з емоційним контекстом та конфіденційністю, що потребує відповідальної реалізації й контролю якості.

2.5. Забезпечення безпеки та конфіденційності в системі автоматичної верифікації

Забезпечення безпеки та конфіденційності є важливими аспектами в розробці та експлуатації будь-якої системи, зокрема таких складних і чутливих, як системи автоматичної верифікації новин. Оскільки верифікація новин передбачає обробку великих обсягів інформації з Інтернету, включаючи особисті дані та потенційно конфіденційну інформацію, важливо вжити заходів для захисту як цих даних, так і самих користувачів.

Захист особистих даних користувачів

Основною вимогою при розробці систем автоматичної верифікації є дотримання принципів конфіденційності та захисту особистих даних користувачів. Система повинна бути здатною забезпечити безпеку таких даних, як:

- *IP-адреси та інші ідентифікаційні дані користувачів.* Оскільки система не повинна збирати або обробляти зайву інформацію, важливо обмежити збирання тільки тих даних, які є необхідними для аналізу новин та взаємодії з користувачами. Для цього слід впровадити відповідні технології анонімізації.
- *Забезпечення анонімності.* При використанні URL-адреси новини для аналізу, система не повинна зберігати особисті або чутливі дані користувачів. Всі запити до сервера повинні здійснюватися без ідентифікації користувача, щоб забезпечити максимальний рівень анонімності.

Для досягнення цієї мети слід використовувати такі підходи, як анонімізація даних і обмеження доступу до чутливої інформації тільки для авторизованих осіб.

Шифрування даних

Важливим аспектом безпеки є захист інформації, яка передається через Інтернет. Всі запити до сервера та передача даних між користувачем та системою повинні бути зашифровані, щоб уникнути витоків або несанкціонованого доступу до цієї інформації. Для цього використовуються методи представлені в таблиці 2.8.

Методи шифрування даних

№	Методи	Опис
1	Шифрування HTTPS	Всі дані, що передаються між клієнтським додатком і сервером, повинні бути зашифровані за допомогою SSL/TLS протоколу (HTTPS). Це дозволяє забезпечити конфіденційність даних користувачів і захистити їх від можливих атак, таких як "man-in-the-middle"
2	Зашифровані збереження даних	Всі конфіденційні дані, що зберігаються на сервері, повинні бути зашифровані для запобігання несанкціонованого доступу. Це важливо, якщо система зберігає будь-які дані, пов'язані з користувачами чи їхніми запитами

Відсутність збирання персональних даних

Одним із основних принципів безпеки є мінімізація збирання персональних даних. Система автоматичної верифікації новин повинна бути розроблена таким чином, щоб не зберігати будь-яку інформацію, яка б могла ідентифікувати користувача. Всі дії, які не вимагають особистої ідентифікації, повинні виконуватись без збору даних, таких як:

- Імена користувачів
- Логіни та паролі
- Ідентифікаційні номери або дані, що дозволяють ідентифікувати особу користувача

Замість цього, система повинна фокусуватися на аналізі загальних текстових даних, що надаються користувачами (наприклад, URL новини або текст новини) без збирання будь-яких даних, що стосуються особи.

Взаємодія з зовнішніми сервісами та API

Якщо система автоматичної верифікації новин використовує сторонні сервіси або API для збору даних або здійснення аналізу, необхідно забезпечити захист даних при взаємодії з цими сервісами. Для цього повинні бути виконані заходи приведені в таблиці 2.9.

Заходи захисту даних при взаємодії з зовнішніми серверами або API

№	Заходи	Опис
1	Довірені джерела	Використовувати тільки ті зовнішні сервіси та API, які відповідають стандартам безпеки та конфіденційності даних. Наприклад, API для збору новин або перевірки фактів повинні бути перевірені на наявність вразливих місць та потенційних загроз
2	Перевірка доступу та авторизації	Всі запити до зовнішніх сервісів повинні бути авторизованими, щоб запобігти зловмисним спробам доступу до API. Важливо використовувати механізми авторизації, такі як API-ключі або OAuth

Контроль доступу до системи

Ще одним важливим аспектом безпеки є контроль доступу до системи та її компонентів. Це включає в себе як користувацький доступ, так і доступ до внутрішніх частин системи, де зберігаються дані або виконується обробка запитів:

- *Авторизація та аутентифікація.* Для доступу до адміністративних або чутливих частин системи необхідно використовувати надійні методи аутентифікації (наприклад, двофакторну аутентифікацію), що забезпечують додатковий рівень безпеки.
- *Обмеження доступу.* Кожному користувачеві слід надавати тільки мінімально необхідні права для виконання своїх функцій. Наприклад, користувачам не слід надавати доступ до адміністрування або змінення налаштувань системи, якщо це не є необхідним.

Безпека та конфіденційність даних є критично важливими аспектами при розробці систем автоматичної верифікації новин. Забезпечення захисту персональних даних користувачів, шифрування переданої інформації, а також запобігання несанкціонованому доступу до чутливої інформації є основними заходами для забезпечення безпеки. Крім того, важливо гарантувати, що система взаємодіє тільки з перевіреними зовнішніми сервісами та використовує надійні методи контролю доступу до своїх компонентів. За допомогою цих

заходів можна забезпечити високий рівень безпеки та конфіденційності в процесі автоматичної верифікації новин.

Висновки до розділу 2

Проектування системи автоматичної верифікації новин базується на використанні сучасних технологій обробки природної мови (NLP) і машинного навчання для класифікації новин за їх достовірністю. Одним з основних елементів архітектури є застосування моделей глибокого навчання, таких як трансформери, для аналізу текстів заголовків та основного тексту новин. Зокрема, за допомогою попередньо навчених моделей здійснюється оцінка достовірності новини шляхом виведення ймовірності її правдивості на основі ряду критеріїв, таких як використання дат, статистичних даних, цитат та джерел. Це дозволяє системі автоматично робити висновки про надійність джерела та ймовірність маніпуляцій.

Вибір методів класифікації фейкових новин, зокрема через NLP та машинне навчання, є ефективним підходом, оскільки він дозволяє обробляти великі обсяги даних і адаптувати модель до різноманітних форм подачі новин. Алгоритми машинного навчання здатні враховувати тонкі нюанси в текстах, що дозволяє виявляти навіть найменші ознаки маніпуляцій чи фальсифікацій. Це підвищує точність системи в порівнянні з традиційними методами перевірки фактів. У свою чергу, нейронні мережі, зокрема ті, що засновані на архітектурі трансформерів, здатні забезпечити високу якість розпізнавання тексту та його контексту, що є важливим аспектом для детального аналізу новин.

Особливу увагу було приділено інтеграції з зовнішніми сервісами, що дозволяє системі отримувати додаткову інформацію для перевірки новин. Взаємодія з API соціальних мереж, новинних агентств та інших джерел підвищує точність оцінки достовірності новини, оскільки дозволяє порівнювати її з іншими джерелами та відстежувати потенційні фейки в реальному часі.

Інтеграція таких даних виявляється важливою для верифікації джерел інформації та можливих маніпуляцій.

Зважаючи на важливість захисту персональних даних та конфіденційності користувачів, система автоматичної верифікації новин також повинна забезпечувати високий рівень безпеки. Зокрема, важливо враховувати захист від зловмисних дій, таких як фальшиві запити до API або маніпулювання результатами аналізу, що може призвести до поширення недостовірної інформації. В системі повинні бути впроваджені заходи щодо захисту даних користувачів та забезпечення анонімності, зокрема через використання шифрування та інших технологій безпеки.

Таким чином, проектування системи для автоматичної верифікації новин передбачає використання найсучасніших технологій в галузі обробки природної мови та машинного навчання, а також інтеграцію з зовнішніми джерелами для підвищення точності аналізу. Крім того, безпека та конфіденційність є критично важливими аспектами, що потребують особливої уваги при розробці таких систем. В результаті, система верифікації новин здатна значно покращити процес перевірки достовірності інформації, що є важливим кроком у боротьбі з фейковими новинами в умовах сучасного інформаційного середовища.

РОЗДІЛ 3.

РОЗРОБКА СИСТЕМИ АВТОМАТИЧНОЇ ВЕРИФІКАЦІЇ НОВИН

Система має на меті забезпечити ефективну перевірку новинних джерел і текстів на предмет маніпуляцій, емоційних забарвлень і інших факторів, що можуть впливати на об'єктивність і точність інформації. Основні компоненти, що розглядаються в даному розділі, включають процеси збору даних, аналізу тексту, застосування алгоритмів машинного навчання, а також формування рекомендацій для кінцевого користувача на основі результатів верифікації.

Метою розробленої системи є підвищення рівня медіаграмотності серед користувачів шляхом автоматичного оцінювання новин з подальшим наданням рекомендацій щодо їх достовірності. Створена система інтегрує різноманітні методи перевірки, зокрема аналіз заголовків, виявлення маніпулятивних маркерів, а також додаткові фактори, такі як наявність джерел, дат і цитат. Технічний підхід до верифікації новин дозволяє значно полегшити і прискорити процес перевірки інформації, що, у свою чергу, сприяє більш обґрунтованому споживанню новин і зниженню впливу фейкових новин на аудиторію.

3.1. Структура та компоненти проекту

Розробка системи автоматичної верифікації новин є важливим кроком у боротьбі з поширенням недостовірної інформації. Оскільки в умовах цифрової епохи зростає обсяг новинних потоків, їх швидка перевірка за допомогою автоматичних інструментів є надзвичайно важливою для забезпечення достовірності інформації. Було розглянуто основні компоненти розробленої системи, її архітектуру та функціональність.

Опис структури

Система верифікації новин складається з кількох основних компонентів, кожен з яких виконує свою функцію у процесі автоматичного аналізу інформації. Ці компоненти представлені на рис. 3.1 та описані в таблиці 3.1.



Рис. 3.1. Основні компоненти системи верифікації новин

Таблиця 3.1

Основні компоненти системи верифікації новин

№	Компоненти	Опис
1	2	3
1	Клас NewsVerificationSystem	це основний клас, який відповідає за управління верифікацією новин. Він містить усю логіку для збору, аналізу та оцінки інформації, отриманої з новинних статей
2	Метод analyze_url	цей метод є основною точкою входу для користувача. Він відповідає за отримання новинного контенту з заданого URL та подальший аналіз тексту. Метод викликає інші допоміжні функції для виконання необхідних перевірок і оцінок
3	Методи для оцінки достовірності	після отримання тексту новини система здійснює кілька кроків для перевірки її достовірності, таких як: – Аналіз заголовку та основного тексту через модель машинного навчання

		– Оцінка на основі різних факторів, таких як наявність дат, цифр, цитат, емоційного забарвлення та інші Перевірка на маніпулятивні звороти в тексті, які можуть впливати на сприйняття інформації
--	--	---

1	2	3
4	Моделі машинного навчання	для визначення ймовірної правдивості або фальшивості інформації використовується попередньо натренована модель, яка на основі аналізу тексту визначає правдоподібність новини
5	Аналіз маніпуляцій та емоційних компонентів	за допомогою регулярних виразів система перевіряє наявність маніпуляційних зворотів і емоційних слів, що можуть вказувати на спроби маніпуляції читачем
6	Визначення результату верифікації	на основі отриманих даних система виводить висновок про достовірність новини: "Ймовірно правда", "Ймовірно фейк", або "Сумнівна достовірність". Кожен вердикт має детальні пояснення, що допомагають користувачу краще зрозуміти результати аналізу
7	Рекомендації	на основі результатів аналізу система також генерує рекомендації щодо того, як користувач може перевірити новину або як уникати поширення недостовірної інформації

Основні етапи роботи системи

Система верифікації новин має чітко визначену послідовність дій, яка включає етапи представлені на рис. 3.2 та описані в таблиці 3.2.

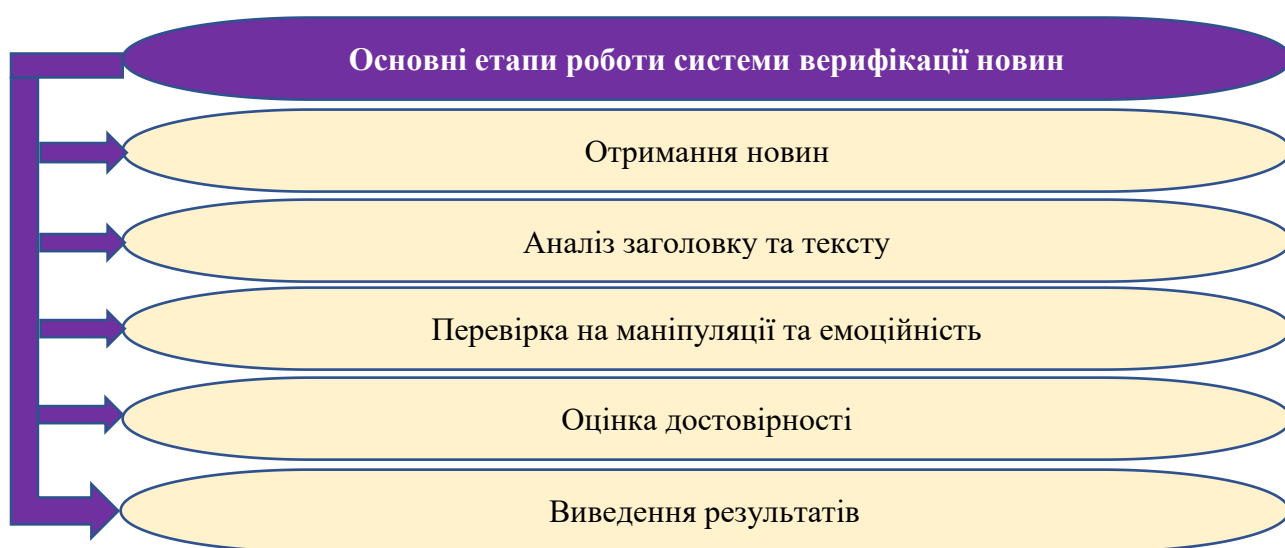


Рис. 3.2. Основні етапи роботи системи верифікації новин

Основні етапи роботи системи верифікації новин

№	Етапи	Опис
1	Отримання новини	користувач надає URL новини, яку потрібно проаналізувати. Система звертається до цього ресурсу та витягує з нього текст новини, включаючи заголовок і основний контент
2	Аналіз заголовка та тексту	спочатку відбувається аналіз заголовку новини з використанням машинного навчання. Після цього система аналізує основний текст, оцінюючи його за кількома критеріями: наявність фактів, джерел, емоційних виразів тощо
3	Перевірка на маніпуляції та емоційність	одночасно з аналізом тексту система виявляє маніпулятивні маркери та оцінює рівень емоційного забарвлення новини
4	Оцінка достовірності	на основі зібраних даних формується фінальна оцінка достовірності новини, яка враховує всі перевірені фактори. Розрахунок включає як текстовий, так і емоційний аналіз
5	Виведення результатів	користувач отримує детальний звіт про новину, включаючи її оцінку достовірності, список маніпулятивних зворотів, якщо вони є, а також рекомендації щодо подальших дій

Розроблена система автоматичної верифікації новин забезпечує високий рівень точності в оцінці достовірності новинного контенту. Завдяки використанню штучного інтелекту, а також комплексному підходу до аналізу різноманітних аспектів тексту, система може не лише визначити правдивість новини, але й виявити потенційні маніпуляції чи емоційні впливи. Це робить її корисним інструментом для користувачів, які бажають отримати швидку та надійну перевірку новин, перш ніж поширювати їх серед інших.

3.2. Розробка серверної частини проекту (Back-End)

Розробка серверної частини є важливим етапом у створенні системи автоматичної верифікації новин. Серверна частина відповідає за обробку запитів від користувача, аналіз текстового контенту новин, оцінку їх достовірності та генерування відповідних рекомендацій. Слід детально розглянути основні компоненти серверної частини, зокрема структуру класу, що забезпечує аналіз новин, методи для обробки запитів та логіку оцінки достовірності інформації.

3.2.1. Реалізація маршрутизації запитів для отримання та перевірки новин

У сучасних умовах інформаційної перенасиченості важливим аспектом є оперативна перевірка достовірності новин, що отримують значний розголос. Технології автоматизації, зокрема використання штучного інтелекту (ШІ), набувають особливої актуальності у сфері верифікації інформації. Одним із способів реалізації таких технологій є створення онлайн-сервісів для перевірки новин. Слід детально розглянути процес маршрутизації запитів для отримання та перевірки новин у рамках системи верифікації, яка використовує штучний інтелект для аналізу тексту, визначення його достовірності та надання рекомендацій.

Реалізація маршрутизації запитів

Маршрутизація запитів у контексті автоматичної верифікації новин передбачає ефективне та структуроване отримання інформації з новинних джерел, її аналіз та видачу результатів. У наданому коді для реалізації цієї функціональності використовується система на базі фреймворку Streamlit, яка надає користувачам можливість ввести URL новини для подальшого аналізу. Ключові етапи реалізації цього процесу представлені на рис. 3.3.

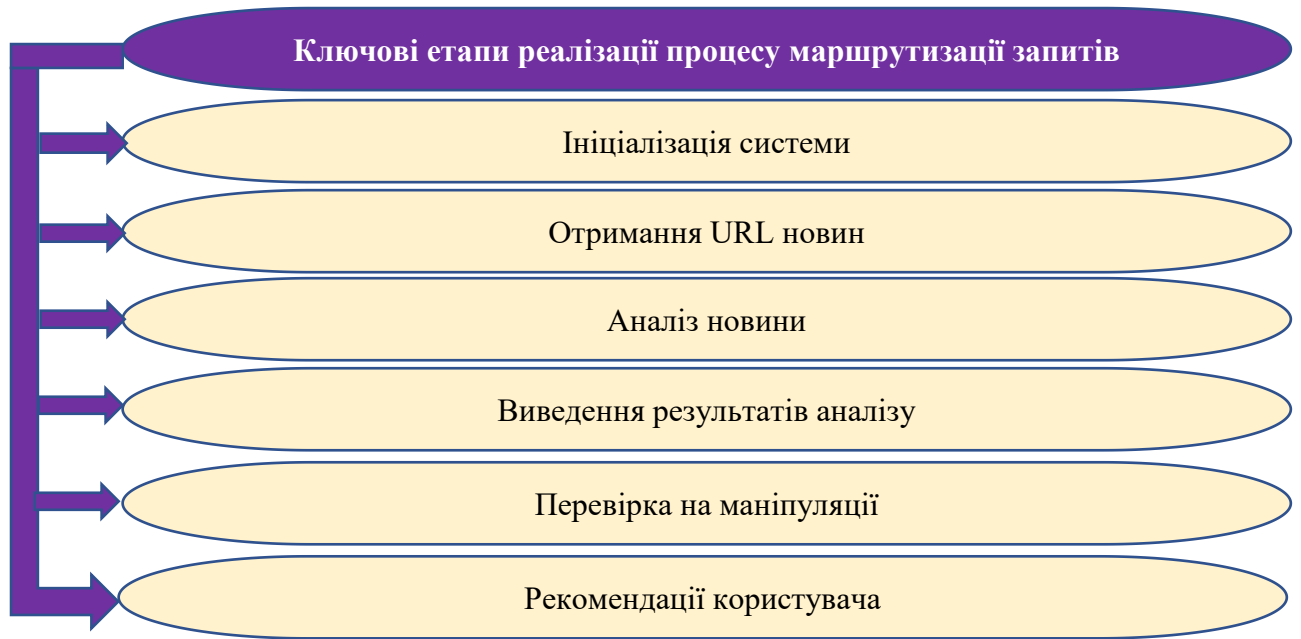


Рис. 3.3. Ключові етапи реалізації процесу маршрутизації запитів

Ініціалізація системи: У першому кроці програма перевіряє, чи існує вже збережений об'єкт системи верифікації в сесії користувача. Якщо об'єкт не знайдено, система ініціалізує новий екземпляр об'єкта `NewsVerificationSystem` (рис. 3.4):

```
if 'system' not in st.session_state:
    with st.spinner("🚀 Ініціалізація системи..."):
        st.session_state.system = NewsVerificationSystem()
```

Рис. 3.4. Фрагмент ініціалізації нового екземпляру об'єкта

Це дозволяє уникнути повторного створення об'єкта системи під час кожного запиту, що значно покращує продуктивність і скорочує час відповіді.

Отримання URL новини (рис. 3.5): Користувач вводить URL новини через текстове поле:

```
url = st.text_input(
    "URL новини:",
    placeholder="https://example.com/
news"
)
```

Рис. 3.5. Отримання URL новин

URL виступає як основний вхідний параметр для системи, що дозволяє здійснити подальший аналіз статті.

Аналіз новини (рис. 3.6): Коли користувач натискає кнопку "Аналізувати", система передає URL до методу `analyze_url` об'єкта `NewsVerificationSystem`. У рамках цього методу здійснюється завантаження новини, її обробка, перевірка достовірності та оцінка на основі моделі штучного інтелекту:

```
results = st.session_state.system.analyze_url(url)
```

Рис. 3.6. Аналіз новин

Важливо зазначити, що цей крок включає в себе кілька етапів:

- Збирання тексту новини за вказаним URL.
- Обробка заголовку та основного тексту новини за допомогою моделі ШІ.
- Розрахунок показників достовірності новини, включаючи аналіз маніпуляцій, емоційності та наявності додаткових факторів (дат, цифр, цитат і джерел).

Виведення результатів аналізу: Після обробки запиту система надає користувачеві детальну інформацію про новину, зокрема:

- Заголовок новини.
- Дата аналізу.
- Оцінка достовірності, ймовірність фейку та рівень маніпуляцій.

- Детальні рекомендації щодо перевірки новини.
- Виявлені маніпуляції та емоційні вирази у тексті.

```
st.subheader("📄 Знайдена стаття")
st.write(f"***Заголовок:** {content['title']}")
st.write(f"***Дата аналізу:** {content['date']}")

st.subheader("📊 Результати аналізу")
st.markdown(f"### {analysis['verdict']}")

col1, col2, col3 = st.columns(3)
with col1:
    st.metric("Достовірність", f"{analysis['metrics']['credibility']:.1%}")
with col2:
    st.metric("Ймовірність фейку", f"{analysis['metrics']['fake']:.1%}")
with col3:
    st.metric("Рівень маніпуляцій", f"{analysis['metrics']['manipulation']:.1%}")
```

Виведення результатів аналізу виглядає наступним чином (рис. 3.7):

Рис. 3.7. Виведення результатів аналізу

Перевірка на маніпуляції: Одним із ключових аспектів маршрутизації запитів є перевірка тексту новини на маніпулятивні маркери. Виявлені маніпуляції знижують рівень достовірності новини, що також відображається в результатах аналізу. Виявлені маніпулятивні звороти виводяться окремим списком (рис. 3.8):

```
if analysis['manipulation_markers']:
    st.subheader("⚠ Виявлені маніпулятивні звороти")
    for marker in analysis['manipulation_markers']:
        st.markdown(f"- {marker}")
```

Рис. 3.8. Виведення виявлених маніпуляцій

Рекомендації для користувача: Система надає персоналізовані рекомендації в залежності від результатів аналізу новини. Зокрема, якщо новина має низьку достовірність або виявлені маніпуляції, користувач отримує попередження (рис. 3.9):

```
st.subheader("💡 Рекомендації")
for rec in analysis['recommendations']:
    st.markdown(rec)
```

Рис. 3.9. Виведення рекомендацій

Реалізація маршрутизації запитів для отримання та перевірки новин в онлайн-сервісі є важливим кроком до забезпечення швидкої та ефективної верифікації інформації в умовах сучасного медіапейза. Використання моделей ІІІ для аналізу тексту новин дозволяє не лише визначити достовірність контенту, але й виявити маніпуляції, оцінити емоційний тон та надати користувачам необхідні рекомендації щодо подальшої перевірки джерел. Технологічна реалізація процесу маршрутизації запитів у вигляді інтерактивного інтерфейсу, що включає всі етапи збору, аналізу та видачі результатів, дозволяє значно підвищити ефективність верифікації новин у реальному часі.

3.2.2. Побудова структури бази даних для зберігання новин та результатів верифікації

Процес автоматичної верифікації новин потребує ефективного зберігання даних, що дозволяє швидко отримати доступ до статей, їх результатів аналізу та рішень щодо достовірності. Для забезпечення стабільної та продуктивної роботи системи верифікації новин необхідно розробити структуру бази даних, яка зможе не лише зберігати всі необхідні дані, але й підтримувати масштабованість та швидкість доступу. Було описано основні принципи та структуру бази даних для зберігання новин, а також результатів їх верифікації.

Структура бази даних

База даних для зберігання новин і результатів верифікації повинна містити кілька основних таблиць для організації та зберігання даних. Кожна таблиця має бути оптимізована для зберігання специфічної інформації, що

стосується новин, їх аналізу та додаткових факторів достовірності. Структура включає наступні основні таблиці:

Таблиця новин (news)

Таблиця новин зберігає основну інформацію про новину, таку як заголовок, текст, дата публікації, а також ідентифікатор URL. Ці дані необхідні для першочергового аналізу. Опис полів таблиці новин представлений в таблиці 3.3.

Таблиця 3.3

Опис таблиці новин

№	Поля	Опис
1	id	унікальний ідентифікатор новини
2	title	заголовок новини
3	text	текст новини
4	url	посилання на новину
5	publication_date	дата публікації

Ця таблиця є основною для зберігання даних про новини та взаємодіє з іншими таблицями через ключові поля.

Таблиця результатів верифікації (verification_results)

Ця таблиця зберігає результати верифікації новин, зокрема оцінку достовірності, рівень маніпуляцій, а також відповідний вердикт. Вона також містить зовнішній ключ для зв'язку з таблицею новин. Опис полів поданий в таблиці 3.4.

Таблиця 3.4

Опис таблиці результатів верифікації

№	Поля	Опис
1	id	унікальний ідентифікатор результату верифікації
2	news_id	зовнішній ключ, що посилається на таблицю новин
3	credibility_score	оцінка достовірності
4	fake_probability	ймовірність того, що новина є фейком
5	manipulation_score	оцінка маніпуляцій

6	verdict	(результат верифікації — "Ймовірно правда", "Ймовірно фейк", "Сумнівна достовірність")
7	recommendations	рекомендації для користувача, наприклад, для перевірки джерел або уникнення поширення

Таблиця результатів верифікації є важливою частиною, оскільки забезпечує збереження результатів для подальшого аналізу та виведення користувачеві рекомендацій.

Таблиця маніпуляційних маркерів (manipulation_markers)

В окремій таблиці можна зберігати виявлені маніпуляційні маркери, які визначаються під час аналізу новини. Ці маркери дозволяють додатково оцінити рівень маніпуляцій у тексті новини, що допомагає визначити достовірність. Опис полів – таблиця 3.5.

Таблиця 3.5

Опис таблиці маніпуляційних маркерів

№	Поля	Опис
1	id	унікальний ідентифікатор маркера
2	verification_result_id	зовнішній ключ, що посилається на таблицю результатів верифікації
3	marker	конкретний маніпуляційний маркер, наприклад, "незрозумілі джерела", "емоційний накал"

Ця таблиця допомагає розділити та зберігати окремо маніпуляції, що виявлені в новинах.

Таблиця додаткових факторів достовірності (credibility_factors)

Окремо варто зберігати таблицю для додаткових факторів достовірності, таких як наявність дат, чисел, цитат або джерел, а також рівень емоційності тексту. Це дозволяє більш детально відслідковувати фактори, які впливають на оцінку новини. Поля таблиці можуть описані в таблиці 3.6.

Таблиця 3.6

Опис таблиці додаткових факторів достовірності

№	Поля	Опис
---	------	------

1	id	унікальний ідентифікатор фактора
2	verification_result_id	зовнішній ключ, що посилається на таблицю результатів верифікації
3	has_dates	наявність дат у новині
4	has_numbers	наявність чисел або статистичних даних
5	has_quotes	наявність цитат
6	has_sources	наявність джерел
7	emotional_level	рівень емоційності в тексті

Збереження цих даних дозволяє детальніше коригувати оцінку достовірності на основі конкретних характеристик новини.

Таблиця користувачів (users)

Для підтримки системи, в якій можна буде відстежувати активність користувачів, варто зберігати таблицю користувачів, щоб мати змогу вести статистику аналізу новин. Поля цієї таблиці описані в таблиці 3.7.

Таблиця 3.7

Опис таблиці користувачів

№	Поля	Опис
1	id	унікальний ідентифікатор користувача
2	username	ім'я користувача
3	email	електронна пошта
4	role	роль користувача, наприклад, "адміністратор", "користувач"

Користувачі можуть мати різні рівні доступу до системи верифікації новин, і ця таблиця дозволяє ефективно керувати їх доступом.

Правильна організація структури бази даних є критично важливою для успішного функціонування системи верифікації новин. Від точності зберігання та швидкості доступу до результатів аналізу залежить ефективність роботи системи. В описаній структурі бази даних кожен елемент відіграє важливу роль у забезпеченні цілісності даних, надаючи можливість зберігати новини, їх результати верифікації, маніпуляційні маркери та додаткові фактори достовірності. Це дозволяє не лише здійснювати якісний аналіз новин, але й

підтримувати масштабованість та швидкий доступ до необхідної інформації для подальшого аналізу та прийняття рішень.

3.3. Розробка клієнтської частини проекту (Front-End)

В умовах сучасного інформаційного перевантаження та стрімкого поширення дезінформації, важливість перевірки новин на достовірність стає надзвичайно актуальною. Для забезпечення швидкої та ефективної верифікації контенту, у рамках цього проекту була розроблена система автоматичної перевірки новин через онлайн-сервіс. Клієнтська частина цього проекту є інтерфейсом для кінцевих користувачів, який дозволяє зручно та оперативно отримувати оцінки достовірності новин, а також рекомендації для подальших дій.

3.3.1. Інтерфейс для користувачів, що дозволяє перевірити новини на достовірність

Розробка інтерфейсу для користувачів є одним із ключових етапів у створенні зручної та ефективної системи верифікації новин. Інтерфейс повинен бути простим у використанні, інтуїтивно зрозумілим та доступним для різних категорій користувачів. Основною метою є забезпечення швидкого доступу до інструментів для перевірки новин, а також надання користувачеві вичерпної інформації про надійність того чи іншого джерела інформації.

Основні функції інтерфейсу показані в таблиці 3.8.

Таблиця 3.8

Основні функції користувацького інтерфейсу

№	Функції	Опис
1	Введення URL новини	користувач може вставити URL новини для автоматичної верифікації
2	Оцінка достовірності	система надає оцінку достовірності новини на основі аналізу її змісту

3	Рекомендації та маніпуляції	користувач отримує рекомендації щодо перевірки новини та інформацію про можливі маніпуляції в тексті
4	Деталі про статтю	користувач може переглянути додаткові деталі та контекст новини, що дають змогу краще зрозуміти підстави для оцінки її достовірності

Інтерфейс базується на використанні веб-технологій та передбачає інтеграцію з серверною частиною, що здійснює перевірку новин.

Введення URL новини

У користувача є можливість вставити URL новини в текстове поле. Після цього, за допомогою кнопки «Аналізувати», система здійснює запит до серверної частини для обробки новини.

Аналіз новини

Після введення URL та натискання кнопки «Аналізувати» відбувається взаємодія з сервером, на якому працює модель штучного інтелекту для обробки тексту. Система здійснює ряд перевірок:

- Аналіз заголовка та тексту новини з використанням нейронної мережі для визначення ймовірності достовірності.
- Виявлення маніпуляційних маркерів, таких як емоційний контекст чи невірні ствердження.
- Оцінка додаткових факторів, таких як наявність дат, цифр, цитат чи джерел інформації.

Оціночний процес включає коригування балів на основі ваги кожного елементу (наприклад, тексту або заголовка), а також врахування маніпуляцій.

Результати аналізу

Після виконання перевірки, система виводить результати:

- Заголовок та дата статті – для швидкої ідентифікації новини.
- Основний вердикт – ймовірно правдива, ймовірно фейк чи сумнівна достовірність.
- Метрики достовірності – відображення рівня достовірності новини, ймовірності фейку та рівня маніпуляцій.
- Виявлені маніпулятивні звороти – список фраз або висловів, що можуть вказувати на маніпуляції.
- Рекомендації – пропозиції для користувача щодо перевірки новини через додаткові джерела.

Інтерактивний розділ для тексту новини

Користувач може натискати на кнопку для перегляду повного тексту новини, якщо це необхідно для більш глибокого аналізу.

Сайдбар з інформацією про систему

У бічній панелі відображається важлива інформація про систему:

- Опис використаних технологій (штучний інтелект, аналіз маніпуляцій, оцінка джерел).
- Технічний статус та інформація про пристрій, на якому виконується система (GPU чи CPU).

Кнопки для оновлення та перенавчання моделі

Для адміністраторів системи передбачена можливість перезавантажити модель або перенавчити її за необхідності. Це дає змогу забезпечити постійну актуальність алгоритмів.

Відображення результатів у вигляді графічних елементів

Для зручності сприйняття користувач отримує візуальне представлення результатів через графічні елементи, такі як діаграми та метрики, що робить аналіз більш доступним і зрозумілим.

Розроблений інтерфейс для перевірки новин на достовірність є зручним та інтуїтивно зрозумілим інструментом для кінцевих користувачів. Він дозволяє швидко отримати детальну інформацію про новину, оцінити її достовірність та виявити можливі маніпуляції. Завдяки інтеграції з потужною серверною частиною, система забезпечує точні та надійні результати аналізу. Важливою перевагою є доступність для різних типів користувачів і можливість отримання рекомендацій щодо подальших дій, що робить систему надзвичайно корисною для боротьби з дезінформацією в реальному часі.

3.3.2. Додаткові модулі аналітики та статистики для користувачів

У сучасному інформаційному просторі, де новини поширюються через численні цифрові канали, виникає необхідність у технологіях, що забезпечують перевірку достовірності інформації. Одним із важливих елементів таких систем є автоматична верифікація новин, що дає змогу швидко оцінити надійність публікацій. Важливо розглянути модуль аналізу новин, що включає кілька додаткових функцій для детального оцінювання інформації, таких як аналіз

додаткових факторів, коригування оцінки достовірності та виявлення маніпуляційних елементів. Ці модулі є важливими для надання користувачам точних і обґрунтованих результатів аналізу новин.

Аналіз додаткових факторів достовірності

Першим важливим етапом є розробка і впровадження методу, який дозволяє автоматично аналізувати додаткові фактори, що можуть впливати на достовірність новини. У системі використовуються п'ять основних категорій факторів, зокрема наявність дат, чисел, цитат, джерел та рівень емоційності тексту. Для кожної категорії тексту, що аналізується, застосовуються регулярні вирази (regex), які дозволяють виявляти ключові елементи, що можуть свідчити про серйозність чи маніпулятивність публікації.

Зокрема, аналіз наявності дат дозволяє визначити, чи містить текст чіткі вказівки на час, що дозволяє перевірити актуальність новини. Пошук чисел і статистики дає змогу перевірити, чи містить новина конкретні дані, що підвищують її достовірність. Наявність цитат та посилань на джерела дозволяє переконатися в обґрунтованості поданої інформації. Наявність емоційно забарвлених слів у тексті служить індикатором можливого маніпулювання емоціями читача.

Зібрані дані коригуються і нормалізуються до певних значень, що дозволяє зрозуміти, які аспекти новини можуть підвищувати чи знижувати її достовірність.

Схема аналізу додаткових факторів достовірності зображена на рис. 3.10.



Рис. 3.10. Схема аналізу додаткових факторів достовірності

Коригування оцінки достовірності на основі додаткових факторів

Коли базова оцінка достовірності новини отримана на основі аналізу заголовку та основного тексту, наступним кроком є коригування цієї оцінки з урахуванням додаткових факторів. Зокрема, позитивні фактори, такі як наявність дат, чисел, цитат і джерел, підвищують достовірність новини, додаючи певну кількість балів до базової оцінки. Натомість емоційність тексту може знизити достовірність, тому що часто маніпулятивні новини використовують емоційно заряджені слова для того, щоб вплинути на сприйняття читача.

Математично це коригування здійснюється шляхом додавання позитивних коригувань і віднімання негативних. Оцінка варіюється від 0.1 до 0.9, що дозволяє створити більш точну картину достовірності новини.

Схема коригування оцінки достовірності на основі додаткових факторів зображена на рис. 3.11.



Рис. 3.11. Схема коригування оцінки достовірності на основі додаткових факторів

Аналіз маніпуляційних маркерів

Маніпуляція в тексті є однією з основних причин для зниження достовірності новини. Виявлення маніпуляційних зворотів є важливим кроком у визначенні надійності джерела. Система використовує спеціальні методи для виявлення маніпуляційних маркерів, таких як некоректне застосування емоційних термінів або маніпуляційне подання фактів.

За допомогою цих маркерів система здатна визначити, чи використовуються у тексті маніпуляції, що можуть змінити сприйняття новини у читача. Цей етап також коригує фінальну оцінку достовірності новини, знижуючи її за наявності маніпуляційних елементів.

Схема аналізу маніпуляційних маркерів показана на рис. 3.12

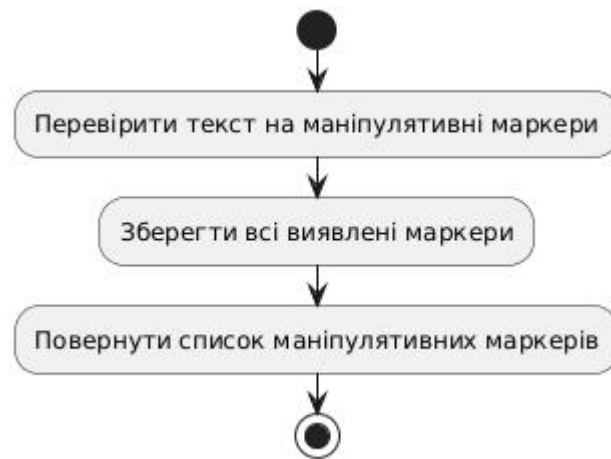


Рис. 3.12. Схема аналізу маніпуляційних маркерів

Генерація рекомендацій на основі результатів аналізу

Одним із важливих аспектів системи є надання рекомендацій користувачам після виконання аналізу. Це дає змогу зрозуміти, що саме потрібно робити з інформацією, яку система оцінює як сумнівну чи потенційно маніпулятивну.

Рекомендації поділяються на три основні категорії в залежності від оцінки достовірності новини: для новин із низькою достовірністю пропонуються заходи для додаткової перевірки, для новин середнього рівня достовірності надаються поради для ретельнішого аналізу, а для новин з високим рівнем достовірності рекомендується перевірити додаткові джерела для підтвердження інформації. Такі рекомендації дозволяють користувачам не лише отримати оцінку новини, але й діяти відповідно до результатів її аналізу.

Схема генерації рекомендацій на основі результатів аналізу представлена на рис. 3.13.

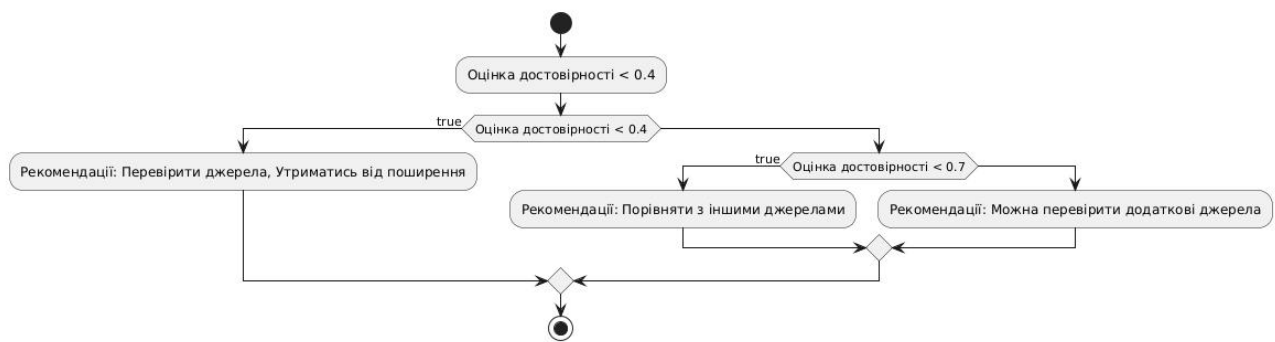


Рис. 3.13. Схема генерації рекомендацій на основі результатів аналізу

Розглянуті модулі аналітики та статистики в системі автоматичної верифікації новин забезпечують всебічний підхід до оцінки достовірності інформації. Завдяки впровадженню аналізу додаткових факторів, коригуванню оцінки на основі маніпуляцій та емоційності, а також виявленню маніпуляційних маркерів, система здатна не лише давати точну оцінку, але й допомагати користувачам у прийнятті обґрунтованих рішень щодо сприйняття новин. Це підвищує ефективність боротьби з фейками та дезінформацією, сприяючи збереженню довіри до медіа.

3.4. Інтеграція Python-бібліотек для роботи з текстовими даними (наприклад, NLTK, spaCy) та навчання моделей (наприклад, scikit-learn, TensorFlow)

У сучасному світі обробка текстових даних і автоматизація процесів верифікації інформації є надзвичайно важливими аспектами для забезпечення достовірності новин. Автоматична верифікація новин через онлайн-сервіси, як показано в наданому коді, включає використання кількох потужних Python-бібліотек для аналізу тексту, оцінки його достовірності та побудови моделей машинного навчання для подальшого аналізу. Одними з таких бібліотек є NLTK, spaCy для обробки тексту, а також scikit-learn та TensorFlow для навчання моделей.

Інтеграція бібліотеки NLTK для обробки тексту

NLTK (Natural Language Toolkit) є однією з найбільш популярних бібліотек Python для обробки природних мов. Вона надає широкий набір інструментів для обробки текстових даних, таких як токенизація, лемматизація, парсинг та інші функції. У контексті верифікації новин, NLTK може бути корисною для підготовки та попередньої обробки тексту перед застосуванням моделей для оцінки достовірності.

У коді, наданому для аналізу новин, можна побачити використання регулярних виразів (за допомогою бібліотеки `re`) для пошуку конкретних елементів тексту, таких як дати, числа, цитати і джерела. NLTK може бути додатково інтегрована для більш детальної обробки текстових даних, наприклад:

- Токенизація: розбиття тексту на окремі слова чи фрази, що дозволяє точно визначати ключові слова.
- Лемматизація: приведення слів до базової форми для кращого розуміння та аналізу.
- Виділення іменованих сутностей (NER), таких як імена, географічні назви, організації, що є корисним для оцінки наявності об'єктивної інформації.

Використання spaCy для обробки природної мови

spaCy є ще однією потужною бібліотекою для обробки природної мови. Вона орієнтована на швидку і ефективну роботу з великими обсягами текстових даних. У порівнянні з NLTK, spaCy надає більшу точність та швидкість у виконанні основних завдань обробки тексту.

У контексті верифікації новин spaCy може бути використана для:

- Сегментації тексту на речення та абзаци, що допомагає розділити статтю на логічні частини для подальшого аналізу.
- Пошуку іменованих сутностей, що дозволяє автоматично визначати важливі елементи тексту (наприклад, організації чи події), які можуть бути корисними при перевірці достовірності.

- Аналізу граматичної структури тексту, що може допомогти виявити маніпулятивні чи неправдиві конструкції.

Завдяки можливостям spaCy, можна створити більш точні і детальні алгоритми для аналізу текстів новин, що дозволяють виявляти потенційно сумнівні чи фальшиві елементи.

Використання scikit-learn для навчання моделей

scikit-learn є однією з найпопулярніших бібліотек Python для машинного навчання. Вона надає велику кількість алгоритмів для класифікації, регресії та кластеризації даних. У контексті автоматичної верифікації новин, scikit-learn може бути використана для створення моделей, які будуть класифікувати текст новин як достовірний чи фейковий на основі ряду ознак.

Одним з основних етапів є створення моделі класифікації текстів за допомогою методів машинного навчання, таких як наївний байєсівський класифікатор, SVM (Support Vector Machine) чи дерева рішень. Для цього необхідно:

- Зібрати навчальний набір даних, який складається з різних новин, що мають мітки "достовірно" або "фейк".
- Використовувати техніки векторизації тексту, такі як TF-IDF або word2vec, щоб перетворити текст на числові вектори для подальшого використання в алгоритмах класифікації.
- Навчити модель на цих даних, а потім застосувати її до нових статей для визначення їх достовірності.

Використання TensorFlow для глибокого навчання

TensorFlow — це потужна бібліотека для глибокого навчання, яка дозволяє створювати нейронні мережі для аналізу складних структур даних. У разі верифікації новин, можна використовувати нейронні мережі для більш точного та комплексного аналізу тексту.

Моделі глибокого навчання, такі як рекурентні нейронні мережі (RNN), LSTM (Long Short-Term Memory) або трансформери, можуть бути використані для:

- Виявлення контексту та смислових зв'язків у тексті, що дозволяє моделі краще розуміти зміст новини.
- Ідентифікації маніпуляцій та тонких нюансів у мові, які можуть бути пропущені за допомогою традиційних методів.

Такі моделі можуть бути навчання на великих наборах даних, що дозволяє їм досягти високої точності в оцінці достовірності новин.

Інтеграція з іншими інструментами для перевірки достовірності

На додаток до основних бібліотек, важливо інтегрувати систему з іншими інструментами для перевірки достовірності інформації, такими як сервіси для перевірки фактів (наприклад, FactCheck.org, PolitiFact), що дозволяє автоматично здійснювати запити до онлайн-баз даних для порівняння фактів.

Такі інструменти можуть бути використані в межах створеного коду для того, щоб підтвердити чи спростувати конкретні факти, які згадуються в тексті новини.

Інтеграція Python-бібліотек для обробки текстових даних і машинного навчання дає змогу створювати ефективні системи для автоматичної верифікації новин. Завдяки таким бібліотекам, як NLTK, spaCy, scikit-learn і TensorFlow, можна розробляти моделі, які не тільки класифікують новини як достовірні чи фейкові, але й враховують різноманітні фактори, такі як емоційний зміст, маніпуляції в тексті та відсутність або наявність джерел. Подальший розвиток таких систем дозволить значно підвищити ефективність боротьби з фейковими новинами та маніпуляціями в медіа.

3.5. Тестування та налагодження системи

Тестування та налагодження є невід'ємною частиною процесу розробки програмного забезпечення, особливо коли йдеться про систему, яка здійснює автоматичну верифікацію новин. Адже точність, надійність і ефективність роботи такої системи напряду залежать від її здатності правильно оцінювати достовірність інформації. Важливо розглянути процес тестування та

налагодження системи, яка використовує штучний інтелект для автоматичної перевірки новин, зокрема її ключові компоненти, стратегії тестування та вирішення можливих проблем.

Процес тестування та налагодження

Етапи процесу тестування та налагодження показані на рис.3.14.

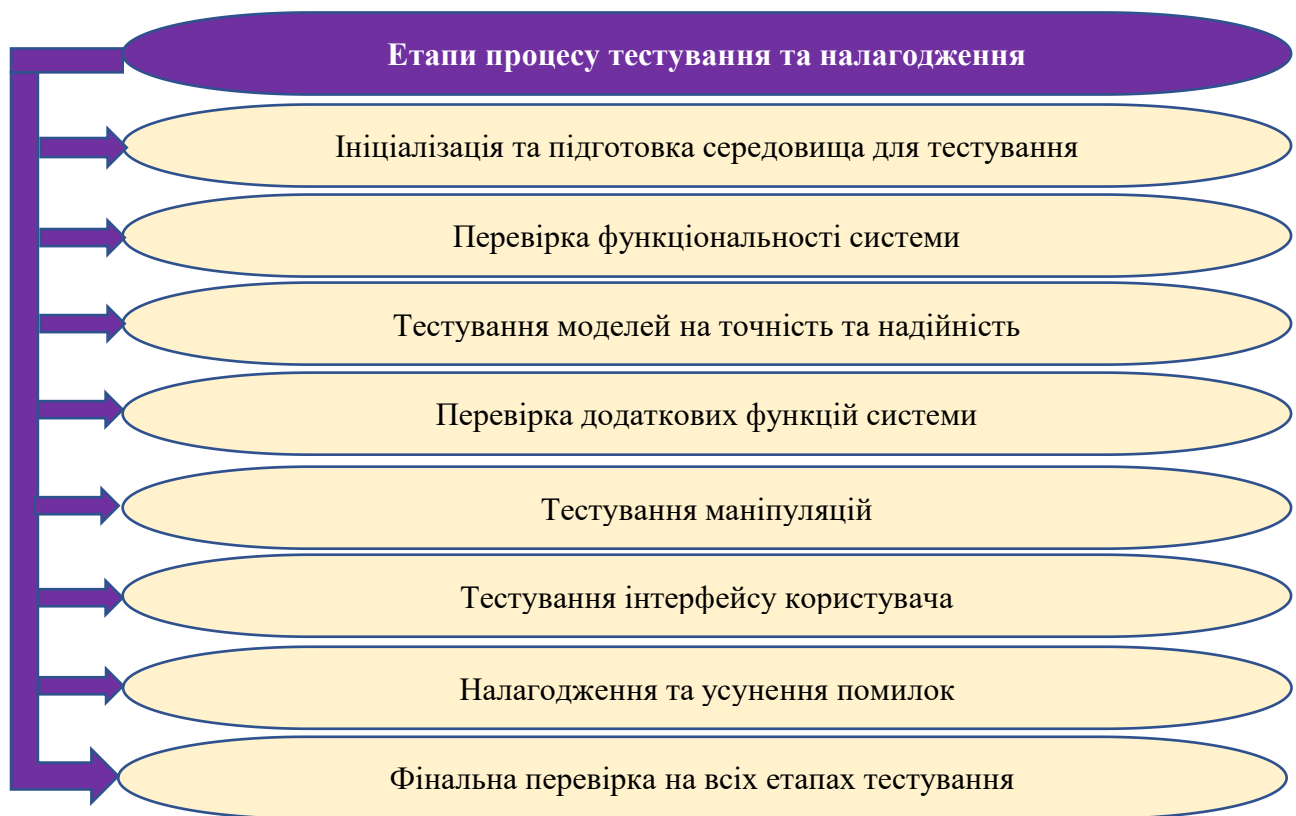


Рис. 3.14. Етапи процесу тестування та налагодження

Ініціалізація та підготовка середовища для тестування

Першим кроком у тестуванні є налаштування середовища, що забезпечує стабільну роботу системи. У представленому коді використовуються такі бібліотеки, як streamlit для створення інтерфейсу користувача та torch для виконання моделей машинного навчання, зокрема класифікації новин. Перед тестуванням необхідно перевірити, чи всі залежності налаштовані коректно, а також чи є доступ до обчислювальних потужностей (наприклад, до GPU), що може впливати на швидкість роботи моделі.

Перевірка функціональності системи

Одним із ключових етапів тестування є перевірка базових функцій системи, таких як аналіз тексту новини за допомогою моделі та формування результатів оцінки достовірності. Важливо перевірити, чи система коректно обробляє введений користувачем URL та чи правильно вона витягує текст з новини для подальшого аналізу. Тестування цих функцій можна проводити як вручну, вводячи різні URL, так і через автоматизовані юніт-тести, що дозволяють перевірити всі можливі сценарії роботи системи.

Тестування моделей на точність та надійність

У коді передбачено використання попередньо натренованої моделі для оцінки достовірності заголовка та тексту новини. Моделі для класифікації новин повинні бути перевірені на точність перед тим, як вони будуть використані в системі. Тестування передбачає перевірку на великій вибірці новин, де вказано, чи є вони правдивими або фейковими, що дозволить зрозуміти, наскільки ефективно модель справляється з поставленим завданням.

Перевірка додаткових функцій системи

Окрім основного аналізу достовірності новини, система також повинна враховувати додаткові фактори, такі як наявність емоційності, цитат, джерел та числових даних. Для цього проводиться тестування кожної з цих функцій через верифікацію шаблонів регулярних виразів, які використовуються для виявлення цих факторів. Тестування дозволяє оцінити, чи правильно система визначає наявність або відсутність конкретних елементів у текстах новин.

Тестування маніпуляцій

Система повинна виявляти маніпулятивні маркери в текстах новин, такі як емоційні заклики чи зміщення фокусу на певні аспекти інформації. Тестування цих функцій полягає в тому, щоб перевірити, чи система коректно ідентифікує маніпулятивні звороти та оцінює їх вплив на загальну достовірність новини.

Тестування інтерфейсу користувача

Важливим етапом є тестування інтерфейсу, що взаємодіє з користувачем. Оскільки система надає користувачам можливість вводити URL для аналізу новин, важливо перевірити, чи коректно відображаються результати аналізу, чи зручний інтерфейс і чи є всі необхідні функції для повної взаємодії з системою.

Налагодження та усунення помилок

Після тестування кожного компонента системи необхідно провести налагодження для виявлення та усунення можливих помилок. Під час тестування можуть виникати різні проблеми, наприклад, неправильна обробка тексту новини, некоректне відображення результатів аналізу або збої у роботі моделі. Важливо мати систему логування, що дозволить швидко ідентифікувати помилки, а також вивести корисну інформацію для виправлення ситуації.

Фінальна перевірка на всіх етапах тестування

Після налагодження всі компоненти системи повинні пройти фінальне тестування, де перевіряються всі функції системи в комплексі, що дасть змогу оцінити її загальну працездатність. Крім того, слід перевірити, як система реагує на різні варіанти текстів новин — як з наявністю, так і без наявності маніпуляцій чи емоційних зворотів, чи правильно оцінює достовірність новин за різними критеріями.

Покрокове тестування програми

Запуск програми – ініціалізація тензорів (рис. 3.15)

```
PowerShell для разработчиков
+ PowerShell для разработчиков
** Visual Studio 2022 Developer PowerShell v17.11.5
** Copyright (c) 2022 Microsoft Corporation
*****
PS D:\work_zulu\FakeNews\App> streamlit run App.py

You can now view your Streamlit app in your browser.

Local URL: http://localhost:8501
Network URL: http://192.168.1.3:8501

C:\Users\user\AppData\Local\Programs\Python\Python312\Lib\site-packages\torchvision\io\image.py:14: UserWarning: Failed to load image Python extension: '[WinError 127] Не найдена указанная процедура'If you don't plan on using image functionality from `torchvision.io`, you can ignore this warning. Otherwise, there might be something wrong with your environment. Did you have `libjpeg` or `libpng` installed before building `torchvision` from source?
  warn(
2024-11-12 16:41:12.692462: I tensorflow/core/util/port.cc:153] oneDNN custom operations are on. You may see slightly different numerical results due to floating-point round-off errors from different computation orders. To turn them off, set the environment variable `TF_ENABLE_ONEDNN_OPTS=0`.
2024-11-12 16:41:43.259565: I tensorflow/core/util/port.cc:153] oneDNN custom operations are on. You may see slightly different numerical results due to floating-point round-off errors from different computation orders. To turn them off, set the environment variable `TF_ENABLE_ONEDNN_OPTS=0`.
```

Рис. 3.15. Ініціалізація тензорів

Ініціалізація системи (рис 3.16)

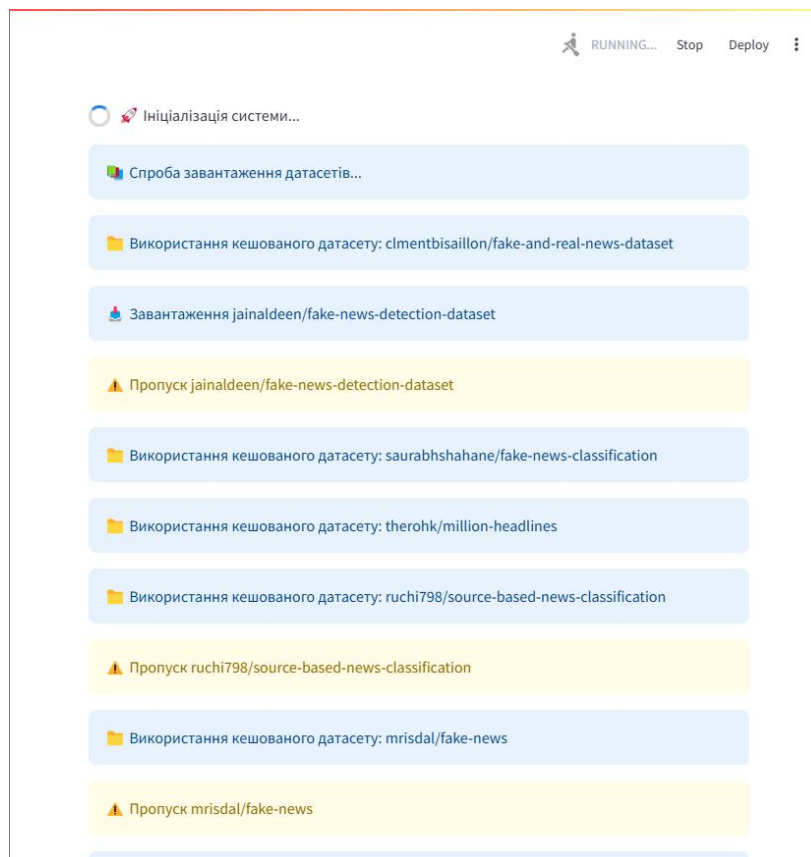


Рис 3.16. Ініціалізація системи

Завантаження датасетів (рис. 3.17)

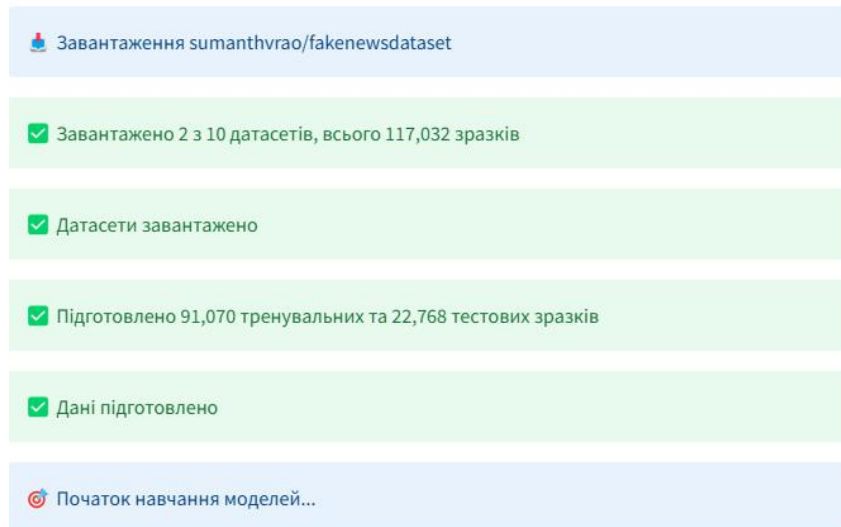


Рис. 3.17. Завантаження датасетів

Інформація про систему (рис. 3.18)

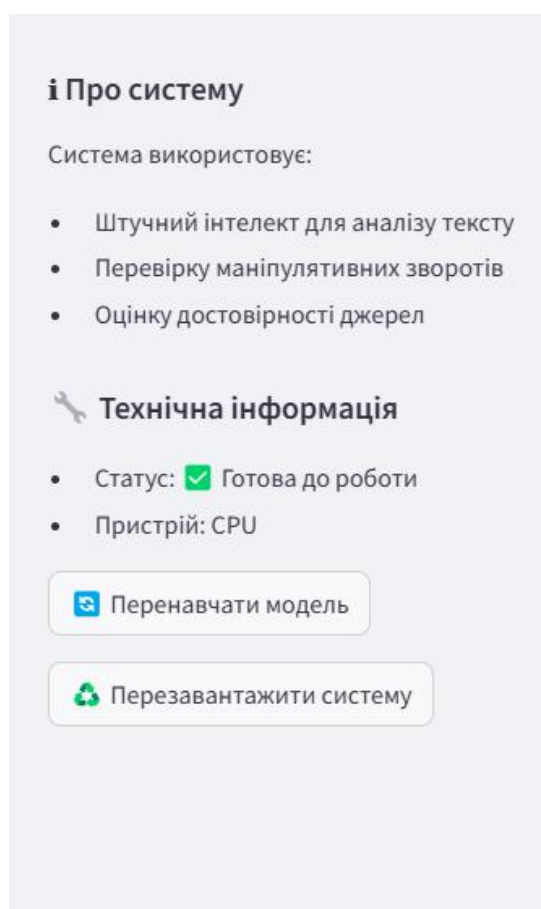


Рис. 3.18. Інформація про систему

Інтерфейс користувача (рис. 3.19)

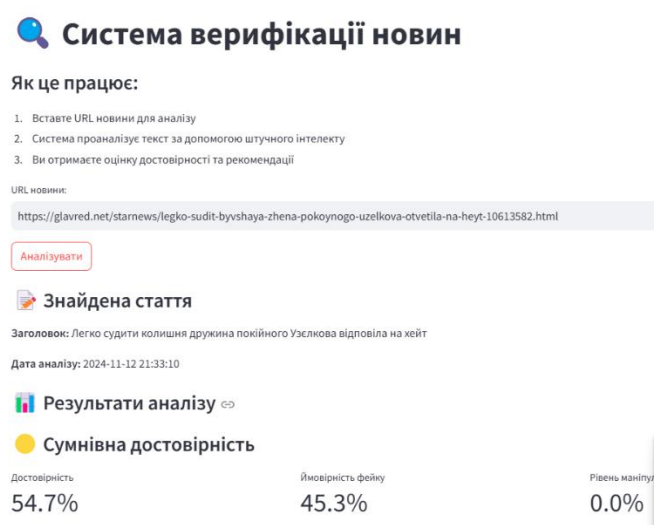


Рис. 3.19. Інтерфейс користувача

Аналіз статті новин (рис. 3.20)

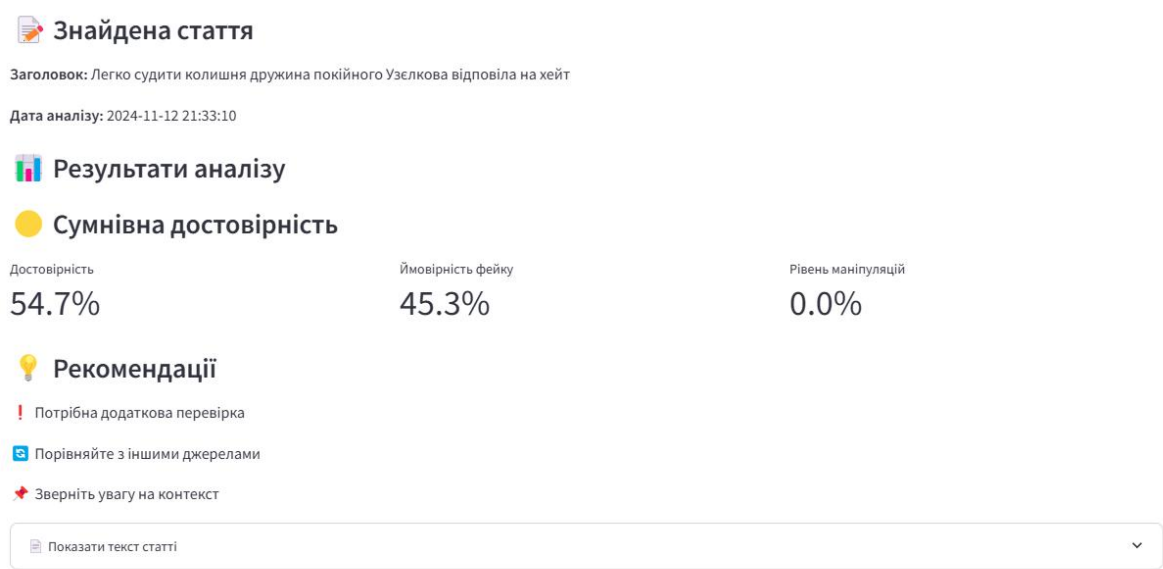


Рис. 3.20. Результати верифікації статті новин

Вивід рекомендацій (рис. 3.21)

Рекомендації

 Потрібна додаткова перевірка

 Порівняйте з іншими джерелами

 Зверніть увагу на контекст

 Показати текст статті

Ексдружина покійного українського боксера Вячеслава Узелкова Марина Боржемська зіткнулася з хвилею хейту в мережі. Причиною стало те, що вона нібито відвернулася від колишнього чоловіка, який був тяжко хворий, і те, що її приклад нібито наслідували їхні спільні діти. Тим часом, у день смерті спортсмена і сама Боржемська, і її син Роберт вшанували пам'ять чоловіка у своїх соцмережах. Марина також поділилася, що, незважаючи на всі труднощі та конфлікти в їхніх стосунках, вона вдячна Вячеславу за їхніх спільних дітей - сина Роберта і доньку Олівію. Пізніше вона зробила пост на своїй сторінці в Instagram, де відповіла на хвилю ненависті, яка полилася на її сім'ю. Багато хто знав Марину Боржемську, як дружину відомого боксера Вячеслава Узелкова. Однак справжня слава до дівчини прийшла 2016 року, коли вона стала тренером у популярному телевізійному проєкті Зважені та щасливі. У 1999 році пара одружилася, проте через 22 роки їхній шлюб дав тріщину. За словами Марини, причина розлучення в тому, що їхні шляхи просто розійшлися. А ось сам Слава заявив, що причиною розриву стало його невміння стримувати ревності. Марина сама виховує 14-річного сина Роберта та 11-річну доньку Олівію. Ми всі різні, і саме в цій різноманітності наша особливість. Іноді легко судити інших, не знаючи всіх деталей, і особливо дивують коментарі від тих, з ким я навіть не зустрічалася, але хто стверджує, ніби мене добре знає. У кожного з нас є своя правда, своє бачення і переживання, які формують наші переконання. Це нормально, адже наші життєві історії пишуться з дитинства і складаються з впливу батьків, близьких, соціуму та середовища, - пише вона. За словами Марини, на будь-якому етапі свого шляху люди зустрічають інших людей, і особливо важливо бути уважними до дітей, бо кожне слово, дія та жест мають наслідки. Діти - це наше майбутнє, вони виростуть і створять атмосферу у своїх сім'ях, у суспільстві, країні та світі, - додає вона. Боржемська також каже, що протягом цих років вона навчилася захищатися від негативу, усвідомлюючи, що кожне слово відображає внутрішній світ мовця. Але мої діти ще не готові до цього. Щастя, що в них є я і люди, які завжди стануть на захист,

Рис. 3.21. Рекомендації

Тестування та налагодження системи є ключовим етапом її розробки. Від цього процесу залежить ефективність і точність автоматичної верифікації новин. Система повинна бути здатна коректно оцінювати новини на основі заголовків, текстів та додаткових факторів, таких як емоційність і маніпуляції, а також мати зручний інтерфейс для користувачів. Важливо враховувати, що тестування має бути багатограним, охоплюючи всі компоненти системи і виявляючи будь-які недоліки до її фінального запуску.

Висновки до розділу 3

Розробка системи автоматичної верифікації новин через онлайн-сервіси є важливим кроком до забезпечення достовірної інформації для широкої аудиторії. Описаний процес створення цієї системи передбачає кілька ключових етапів, що взаємодіють між собою та забезпечують ефективне функціонування платформи для перевірки новин. Спроектowana система надає користувачам інструмент для оперативного аналізу інформації, виявлення маніпуляцій та оцінки достовірності джерел.

В основу розробки покладено поєднання сучасних технологій обробки природної мови та алгоритмів машинного навчання, що дозволяє здійснювати точний та швидкий аналіз новинного контенту. Серверна частина проекту, яка відповідає за обробку запитів користувачів і верифікацію текстів новин, була спроектована таким чином, щоб забезпечити масштабованість та високу ефективність в обробці великих обсягів даних. Крім того, для кожної новини, яка надходить на перевірку, система здійснює комплексний аналіз, враховуючи різні фактори, такі як наявність конкретних дат, статистичних даних, цитат, джерел інформації, а також рівень емоційності тексту. Це дозволяє не лише оцінити загальну достовірність контенту, але й виявити можливі маніпуляції та їх вплив на сприйняття новини.

Інтерфейс користувача, реалізований у вигляді онлайн-сервісу, є зрозумілим та доступним. Користувачі можуть легко вставити URL новини для її перевірки та отримати зручні і чіткі результати аналізу, що містять оцінки достовірності, ймовірність фейковості новини, рівень маніпуляцій і рекомендації щодо подальших дій. Це дозволяє не лише проводити оцінку конкретних новин, але й ознайомлюватися з виявленими маніпуляціями та емоційними прийомами, що використовуються в текстах.

Розробка бази даних, яка зберігає результати перевірок, стала необхідною для ведення статистики та аналізу трендів щодо достовірності новин. Збереження такої інформації дозволяє не лише оперативно надавати користувачам актуальні дані, але й забезпечити довготривалий моніторинг ефективності системи, оцінюючи точність та надійність її роботи. Додатково передбачено можливість перенавчання моделі для підвищення її ефективності за допомогою нових даних, що є важливим для адаптації до нових трендів у інформаційному просторі.

Використання сучасних бібліотек для обробки текстових даних, таких як NLTK та spaCy, а також популярних фреймворків для навчання моделей машинного навчання, дозволило побудувати систему, яка є як гнучкою, так і продуктивною. Врахування численних факторів, таких як емоційність тексту та

наявність маніпуляцій, дає змогу надати більш точну оцінку, а не просто «правда» чи «фейк», що є важливим для глибшого розуміння природи новинного контенту.

Важливим етапом є тестування та налагодження системи, оскільки лише на цьому етапі можна виявити всі можливі недоліки та провести їх коригування. Завдяки проведеному тестуванню система була налаштована на високу точність роботи, що дозволяє користувачам отримувати достовірні результати при мінімальних затримках.

Отже, система автоматичної верифікації новин є ефективним інструментом для боротьби з фейковими новинами та маніпуляціями в інформаційному просторі. Вона поєднує в собі передові методи обробки текстових даних, машинного навчання та оцінки достовірності, що дозволяє не тільки виявляти фейкову інформацію, але й допомагати користувачам формувати критичне ставлення до споживаних новин.

РОЗДІЛ 4.

РОЗРОБКА СТАРТАП-ПРОЕКТУ

У сучасному світі поширення фейкових новин є однією з найбільших загроз для суспільства, що впливає на формування громадської думки та прийняття важливих рішень. Автоматична верифікація новин, яка використовує сучасні технології штучного інтелекту та аналізу тексту, є важливим інструментом для боротьби з дезінформацією. Буде представлено концепцію стартап-проекту, спрямованого на створення системи автоматичної верифікації новин, а також технічний опис її реалізації. Особлива увага буде приділена процесу розробки основних компонентів системи, таким як алгоритми аналізу тексту, використання моделей машинного навчання для оцінки достовірності, а також розробка зручного інтерфейсу для кінцевих користувачів.

4.1. Опис ідеї та цінності проекту автоматичної верифікації новин

Ідея проекту полягає в створенні автоматизованої системи, яка дозволяє користувачам перевіряти достовірність новин у реальному часі за допомогою сучасних методів обробки природної мови (NLP) та машинного навчання. Система використовує потужні моделі штучного інтелекту для аналізу тексту новин, перевірки наявності маніпуляцій, оцінки емоційного забарвлення та перевірки фактичної точності інформації. Завдяки цьому, користувачі отримують не тільки загальну оцінку достовірності новини, але й детальні рекомендації щодо подальших дій, таких як перевірка джерел чи порівняння з іншими новинами.

Проект пропонує цінність у вигляді підвищення рівня медіаграмотності серед громадян та надає інструменти для швидкої і точнішої перевірки інформації. Ця система є корисною як для кінцевих користувачів, так і для медіа-компаній, оскільки допомагає зменшити поширення фейкових новин та

маніпуляцій, що можуть негативно впливати на суспільство та політичні процеси.

Технічна реалізація і алгоритми

Основною технологією для реалізації системи є використання штучного інтелекту, зокрема, попередньо натренованих мовних моделей, які можуть здійснювати оцінку текстів на достовірність та виявляти маніпуляції. Основні компоненти системи представлені в таблиці 4.1.

Таблиця 4.1

Основні компоненти системи

№	Компоненти	Опис
1	2	3
1	Аналіз заголовків та тексту	Модель штучного інтелекту здійснює оцінку заголовка та основного тексту новини на основі попередньо натренованих моделей для виявлення потенційної достовірності. Цей процес здійснюється шляхом порівняння з еталонними даними, де системі надається оцінка достовірності за шкалою від 0 до 1
2	Визначення маніпуляцій	Для оцінки маніпуляцій використовуються алгоритми, які визначають наявність маніпулятивних зворотів і оцінюють рівень емоційного забарвлення тексту. Виявлення маніпуляцій також здійснюється на основі текстових патернів, таких як сенсаційні вирази, що можуть свідчити про маніпуляцію або спробу маніпулювання емоціями читача
3	Аналіз додаткових факторів	Система також аналізує такі додаткові фактори, як наявність точних дат, числових даних, цитат та джерел. Ці фактори додають вагу до оцінки достовірності новини, оскільки наявність чіткої і перевіреної інформації свідчить про вищу ймовірність достовірності
4	Коригування оцінки	Після того, як базова оцінка достовірності була отримана на основі аналізу заголовка та тексту, система застосовує коригування, враховуючи фактори, такі як емоційне забарвлення тексту або наявність маніпуляцій
5	Визначення вердикту та рекомендації	На основі отриманих результатів, система генерує фінальний вердикт щодо достовірності новини. У разі необхідності, користувач отримує рекомендації щодо подальших дій, таких як перевірка джерел або додаткове порівняння новини з іншими ресурсами

Інтерфейс користувача

Важливою частиною проекту є розробка зручного інтерфейсу для користувачів, який забезпечує швидкий доступ до результатів аналізу. Система має простий веб-інтерфейс, де користувач може вставити URL новини та отримати швидкий аналіз. Додатково, система генерує відгуки про новину, вказуючи на можливі маніпуляції, рекомендації щодо перевірки та основні характеристики, які можуть допомогти у визначенні достовірності.

Проект автоматичної верифікації новин має важливу соціальну цінність у боротьбі з дезінформацією. Він дозволяє підвищити рівень медіаграмотності серед широкої аудиторії, надаючи інструменти для перевірки достовірності інформації та виявлення маніпуляцій. Технічна реалізація проекту базується на використанні штучного інтелекту та машинного навчання, що дозволяє здійснювати точний та ефективний аналіз новин у реальному часі. Автоматизація процесу верифікації є потужним інструментом для медіа-індустрії та суспільства в цілому, сприяючи зниженню рівня поширення фейкових новин.

4.2. Інформаційна картка проекту

Таблиця 4.2

Інформаційна карта проекту

№	Характеристика	Опис
1	2	3
1	Назва проекту	"Система автоматичної верифікації новин" (News Verification System)
2	Автори проекту	Ровишин Андрій Вячеславович
3	Коротка анотація	Система автоматичної верифікації новин є інструментом для аналізу новинних статей з метою оцінки їх достовірності. Використовуючи передові моделі машинного навчання, система оцінює такі параметри, як точність заголовків, наявність маніпуляційних зворотів, емоційний фон і рівень достовірності джерел. Окрім того, система надає рекомендації для користувачів щодо

		подальшого перевірки інформації та виявлення можливих маніпуляцій
--	--	---

Продовження таблиці 4.2

1	2	3
4	Термін реалізації	1 рік
5	Необхідні фінансові ресурси	<i>Заробітні плати працівникам (щомісячно):</i> Розробники (2 особи): \$3.000 кожен = \$6.000 Тестувальник: \$1.500 Проджект-менеджер: \$2.000 Бізнес-аналітик: \$1.500 Маркетолог: \$1.200 <i>Всього щомісяця: \$11.200</i> <i>Хмарні технології для складних обчислень: \$10.000</i> <i>Програмне забезпечення: \$3.000</i> <i>Обладнання для працівників: \$4.000</i> <i>Рекламне просування: \$3.000</i> <i>Загальний бюджет: \$31.200</i>
6	Необхідні матеріальні ресурси	Комп'ютерні монітори та портативні ноутбуки для тестування
7	Необхідні людино-ресурси	Для успішної реалізації проєкту необхідно залучити команду з 10-15 фахівців, серед яких: – 5 програмістів, що спеціалізуються на машинному навчанні та обробці природної мови. – 2 аналітики даних, що будуть працювати з результатами верифікації і забезпеченням точності системи. – 1 менеджер проєкту для координації роботи та управління командою. – 2-3 тестувальники для перевірки працездатності системи і забезпечення її стабільної роботи. – 2 маркетолога для просування проєкту та залучення користувачів.
8	Опис проблеми, яку вирішує проєкт	У сучасному світі поширення фейкових новин, маніпуляцій та дезінформації стало серйозною проблемою, яка впливає на громадську думку і політичні процеси. Люди все частіше стають жертвами маніпуляцій і неправдивих новин, що може призвести до соціальних і політичних наслідків. Проєкт має на меті надати користувачам інструмент для об'єктивної оцінки новин, що дозволить їм розрізняти достовірну інформацію від маніпуляцій або фейків

1	2	3
9	Головні цілі та завдання	<i>Основна мета проєкту:</i> Розробити та впровадити автоматизовану систему верифікації новин для підвищення рівня довіри до інформації в Інтернеті <i>Завдання:</i> – Розробити модель машинного навчання для аналізу новин. – Створити алгоритм для виявлення маніпуляційних маркерів. – Забезпечити точність і швидкість аналізу новин в реальному часі. – Надати користувачам рекомендації щодо перевірки інформації. – Розширити базу даних новинних джерел і інформації для покращення точності аналізу.
10	Очікувані результати	Оскільки проєкт орієнтований на використання новітніх технологій штучного інтелекту, очікується, що система зможе досягти високої точності в аналізі новин та виявленні маніпуляцій. Очікуваними результатами є: – <i>Зниження поширення фейкових новин</i> серед користувачів. – <i>Підвищення обізнаності</i> користувачів про маніпуляції в медіа. – <i>Покращення довіри до новинних джерел</i>, які проходять перевірку через систему. – <i>Зростання популярності проєкту</i> серед користувачів і можливість масштабування системи на глобальний рівень.

4.3. Технічна інфраструктура

Frontend: Для розробки користувацького інтерфейсу обрано фреймворк Streamlit, який дозволяє швидко створювати веб-додатки на Python.

Backend: В якості основної серверної технології використовується Python разом із фреймворками, що підтримують аналіз тексту та штучний інтелект.

База даних: Для зберігання метаданих статей та результатів їх аналізу можна використовувати MongoDB.

Аутентифікація: Для забезпечення безпечного доступу до системи та захисту персональних даних користувачів пропонується використовувати JSON Web Tokens (JWT).

Хмарні сховища: Для зберігання великих медіафайлів, таких як зображення або відео, може бути використано одне з провідних хмарних сховищ, наприклад AWS або Google Cloud Storage.

4.4. Технологічний аудит та аналіз конкурентного середовища

1. *Оцінка технічної спроможності:* У рамках проекту створення системи автоматичної верифікації новин, команда має необхідний рівень експертизи в галузі веб-розробки та розробки штучного інтелекту. Використання передових моделей машинного навчання для аналізу тексту та штучного інтелекту дозволяє ефективно оцінювати достовірність новин. Крім того, команда має досвід у розробці інтерактивних веб-платформ із використанням фреймворків для швидкої обробки великих обсягів даних.
2. *Вибір технологій та фреймворків:* Для реалізації проекту використовуються сучасні технології, зокрема Python для машинного навчання та веб-розробки, фреймворки на кшталт Streamlit для створення інтерфейсу, а також бібліотеки для обробки природної мови, зокрема Hugging Face. Ці технології забезпечують високу ефективність, швидкодію та масштабованість системи, дозволяючи обробляти великі обсяги текстових даних у реальному часі.
3. *Безпека та конфіденційність:* Враховуючи чутливість обробки новин, система передбачає використання найкращих практик для забезпечення безпеки даних. Всі операції з обробкою та аналізом даних виконуються з високим рівнем конфіденційності, а доступ до особистої інформації користувачів обмежено. Приділено увагу захисту від зовнішніх загроз та дотриманню стандартів захисту даних.

4. *Масштабування та витривалість*: Система проектується таким чином, щоб легко масштабуватися при збільшенні кількості користувачів. Враховано вимоги до витривалості, і забезпечено належне функціонування навіть при високих навантаженнях, що є важливим аспектом для платформи, яка повинна забезпечувати оперативну перевірку новин у реальному часі.
5. *Хмарні технології*: Для зберігання даних і виконання обчислень розглядається використання хмарних сервісів. Це дозволяє забезпечити гнучкість та зручність у масштабуванні інфраструктури, знижуючи витрати на локальні сервери і надаючи можливість швидкої обробки даних без перевантажень.
6. *Забезпечення сумісності*: Платформа оптимізована для роботи на різних пристроях і операційних системах, включаючи десктопи та мобільні телефони, що забезпечує зручність для користувачів з різними умовами доступу до інтернету.
7. *Тестування та оптимізація*: Система передбачає детальне тестування всіх компонентів, включаючи алгоритми машинного навчання, для виявлення помилок та оптимізації швидкодії. Запроваджено постійний моніторинг продуктивності та адаптацію алгоритмів для забезпечення стабільної роботи при різних навантаженнях.
8. *Реалізація монетизації*: Платформа надає можливості для реалізації моделей монетизації через підписки або преміум-функції для розширених можливостей перевірки новин, що дозволить забезпечити сталий розвиток проекту.
9. *Стратегія розширення*: Платформа передбачає розширення функціональності, зокрема, додавання нових інструментів для покращення точності верифікації, впровадження алгоритмів для виявлення маніпуляцій та використання більш широких даних для аналізу достовірності.
10. *Оцінка загальних витрат*: Враховуючи поточні вимоги до розвитку системи, проведено оцінку загальних витрат на розробку, запуск та подальшу

підтримку проекту. Оцінка вартості включає витрати на розробку, хмарні сервіси, тестування та маркетинг.

- 11.*Прогнозування термінів:* Проект має реалістичні терміни розробки та запуску, з урахуванням необхідних етапів тестування та удосконалення алгоритмів.
- 12.*Оцінка ризиків:* Для оцінки можливих ризиків було проведено аналіз безпеки, можливої неефективності моделей верифікації та наявності маніпуляцій. Розроблено стратегії для зниження ризиків, зокрема, через постійне вдосконалення моделей та оновлення бази даних.
- 13.*Висновки:* Після проведеного технологічного аудиту, система автоматичної верифікації новин виявилася технічно здійсненою та перспективною. Приділено належну увагу безпеці, масштабуванню та оптимізації, що дозволяє розглядати проект як ефективний інструмент для покращення достовірності новин в Інтернеті.

4.5. Стратегічний аналіз ринку та цільова аудиторія

Огляд Галузі: Сфера автоматичної верифікації новин за допомогою онлайн-сервісів активно розвивається. Серед конкурентів можна виділити сервіси, які використовують штучний інтелект для оцінки достовірності інформації та виявлення маніпуляцій. Ключовими характеристиками конкурентів є застосування алгоритмів машинного навчання для аналізу контенту та надання рекомендацій щодо перевірки новин.

Цільова Аудиторія: Цільовою аудиторією є користувачі, які шукають надійні інструменти для перевірки новин у реальному часі. Це можуть бути журналісти, редактори новинних ресурсів, а також широка аудиторія, яка хоче отримати достовірну інформацію та уникнути поширення фейкових новин.

Тенденції та Інновації: В останні роки активно впроваджуються новітні технології обробки природних мов і машинного навчання, що дозволяє підвищити точність аналізу та швидкість обробки інформації. Інноваційні

рішення включають інтеграцію додаткових факторів для оцінки достовірності, таких як емоційний рівень тексту та наявність маніпуляцій.

Ринкова Лакуна: На ринку спостерігається дефіцит сервісів, що поєднують комплексний аналіз новин, враховуючи різноманітні фактори достовірності, маніпуляції та емоційність. "Система верифікації новин" може зайняти нішу завдяки інтеграції кількох методів перевірки та наданню рекомендацій для користувачів.

Монетизація та Бізнес-Модель: Потенційні шляхи монетизації включають впровадження преміум-планів для розширених можливостей аналізу, рекламу для новинних платформ, а також партнерські угоди з медіа та фактчекінговими організаціями.

Партнерські Відносини: Для подальшого розвитку системи можливе партнерство з новинними агентствами, технологічними компаніями, а також організаціями, що займаються боротьбою з фейковими новинами.

Ризики та Вирішення: Основними ризиками є недостатня точність алгоритмів або можливість маніпуляцій з боку користувачів. Для мінімізації ризиків необхідно постійно вдосконалювати модель, а також налаштовувати систему на обробку нових типів даних.

Оцінка Обсягу Ринку: Потенційний ринок для системи верифікації новин включає численні медіа, журналістів, а також кінцевих користувачів, що шукають достовірну інформацію. Оцінка обсягу ринку передбачає значні можливості для зростання, враховуючи зростання попиту на подібні інструменти.

Після проведеного аналізу ринку можна стверджувати, що система верифікації новин має потенціал для успішного розвитку. Унікальні можливості, що поєднують точність перевірки та додаткові фактори для аналізу достовірності, дозволяють виділити її серед конкурентів.

4.6. Бізнес-модель проекту

Бізнес-модель проекту "Автоматична верифікація новин через онлайн сервіси" передбачає створення цінності для користувачів, надаючи їм доступ до технології штучного інтелекту для аналізу новинних статей. Система забезпечує перевірку достовірності інформації, виявлення маніпуляцій та надання рекомендацій щодо поширення новин. Інноваційний підхід дозволяє користувачам швидко оцінювати надійність джерел, що є важливим для підвищення рівня медіаграмотності. Офіційні партнери та співпраця з медіа-організаціями забезпечують сталий розвиток і зростання проекту.

4.6.1. Цінність продукту для користувачів та ринку

Програмне забезпечення для автоматичної верифікації новин надає користувачам зручну платформу для перевірки достовірності інформації з новинних ресурсів. Воно дозволяє здійснити аналіз тексту, виявити маніпулятивні елементи та оцінити надійність джерел. Система автоматично генерує рекомендації щодо подальших дій та пропонує детальний розбір статей, допомагаючи користувачам приймати обґрунтовані рішення щодо достовірності новин. Це значно підвищує ефективність перевірки інформації для широкого кола користувачів, від журналістів до звичайних громадян.

4.6.2. Інноваційність і перспективи розвитку проекту

Проект автоматичної верифікації новин є інноваційним, оскільки впроваджує штучний інтелект для оцінки достовірності інформації в реальному часі. Така система, що поєднує аналіз тексту, виявлення маніпуляцій та оцінку джерел, є унікальною на ринку і відповідає сучасним вимогам боротьби з фейковими новинами. Розвиток технології дозволяє постійно вдосконалювати модель для ще більш точного визначення достовірності інформації та адаптації до нових маніпуляційних технік, що відкриває перспективи для широкого використання в медіа-сфері та інших галузях.

4.6.3. Основні сегменти споживачів та їх потреби

1. *Журналісти та медіа-аналітики*, які мають потребу у швидкому перевірці достовірності новин та мінімізації маніпуляцій у публікаціях.
2. *Користувачі соціальних мереж*, які шукають інструменти для оцінки достовірності новин перед їх поширенням серед своїх підписників.
3. *Освітні установи та дослідники*, які використовують сервіси для навчання студентів аналізу медіа та розвитку критичного мислення щодо інформації.
4. *Представники державних органів та організацій*, які потребують інструменти для моніторингу медіа-простору та перевірки фактів в публічних новинах.
5. *Інтернет-компанії та платформи*, які хочуть забезпечити користувачам надійний інструмент для перевірки новин, що поширюються через їх сервіси.

4.7. Розробка маркетингової стратегії для просування сервісу

Маркетингова програма для сервісу автоматичної верифікації новин спрямована на підвищення довіри до продукту серед користувачів і просування його ключових переваг на ринку. Це дозволить забезпечити попит на сервіс серед цільової аудиторії та відрізнити його від конкурентів.

Таблиця 4.3

Маркетингова програма

Потреба	Вигода від товару	Ключові переваги над конкурентами
1	2	3
Перевірка достовірності новин	Забезпечення надійності та точності інформації	Використання штучного інтелекту для точного аналізу тексту
Оцінка маніпуляцій та фейкових новин	Виявлення маніпуляцій і спростування фейків	Підвищена чутливість до маніпулятивних зворотів та

		емоційних висловів
Персоналізовані рекомендації	Індивідуальний підхід до кожної новини	Можливість отримання детальних рекомендацій по кожній статті
Швидкість та зручність верифікації	Миттєвий доступ до результатів аналізу	Проста інтеграція з іншими онлайн платформами та пристроями

Продовження таблиці 4.3

1	2	3
Інтерфейс для користувачів	Інтуїтивно зрозумілий інтерфейс	Адаптивний дизайн для різних пристроїв та платформ
Підвищення медіаграмотності користувачів	Навчання та підтримка критичного мислення	Прозорість аналізу та доступ до всіх етапів перевірки новини
Залучення широкої аудиторії	Поширення результатів серед користувачів	Створення активної спільноти з обміну досвідом та експертними порадами

Висновки до розділу 4

Розроблений стартап "Система верифікації новин" має значний потенціал у сфері автоматичного аналізу та оцінки достовірності новинних матеріалів. Використовуючи технології штучного інтелекту та алгоритми для виявлення маніпуляцій, система забезпечує точну оцінку достовірності інформації, надаючи користувачам детальний аналіз заголовків та текстів новин. Ключовими перевагами є висока точність оцінки достовірності, можливість виявлення маніпуляцій та рекомендації щодо подальшої перевірки інформації. Проєкт має чітку бізнес-модель, що включає інтеграцію з онлайн-платформами та потенціал для впровадження в медійну індустрію. Завдяки інноваційному підходу та акценту на точність, система може стати важливим інструментом для боротьби з дезінформацією в сучасному медіапросторі.

ВИСНОВКИ

Висновки до роботи підсумовують основні досягнення, проблеми та перспективи розвитку автоматичної верифікації новин, а також можливі шляхи її вдосконалення для боротьби з фейковими новинами.

Одним із основних результатів дослідження є розробка концептуальної моделі системи автоматичної верифікації новин, яка використовує сучасні технології машинного навчання, обробки природної мови та аналізу метаданих для виявлення фальсифікацій у новинах. Розроблена система здатна швидко та ефективно обробляти великий обсяг інформації, що дозволяє значно зменшити час, необхідний для перевірки достовірності новин, а також забезпечити високу точність у процесі верифікації. Використання Python-бібліотек таких як NLTK, spaCy, scikit-learn та TensorFlow дозволяє створити потужні моделі для аналізу тексту та класифікації новин, що підвищує рівень автоматизації процесу перевірки.

Інтеграція з зовнішніми сервісами, зокрема API соціальних мереж і новинних агентств, дає змогу отримувати додаткову інформацію для перевірки джерела та аналізу контексту новини. Це дозволяє системі працювати в реальному часі та отримувати найбільш актуальні дані для верифікації. Важливим аспектом є також забезпечення безпеки та конфіденційності інформації, що обробляється в системі, що гарантує захист від потенційних загроз і маніпуляцій.

Проектування архітектури системи виявило деякі виклики, пов'язані з необхідністю оптимізації алгоритмів для забезпечення швидкої та масштабованої роботи системи. Порівняння різних методів верифікації, таких як лінгвістичні підходи, методи на основі машинного навчання та використання метаданих, показало, що комбіноване застосування кількох підходів є найбільш ефективним для досягнення високої точності та універсальності системи.

Розробка стартап-проекту автоматичної верифікації новин виявила важливість стратегічного підходу до аналізу конкурентного середовища та

визначення цільової аудиторії. Бізнес-модель проекту повинна враховувати інноваційність продукту та потреби користувачів, що забезпечить його успішне впровадження на ринку. Розробка маркетингової стратегії для просування сервісу має на меті підвищити обізнаність користувачів про важливість перевірки новин та створити стабільну базу споживачів.

Попри досягнення, дослідження показало, що існують певні обмеження в ефективності автоматичних систем верифікації новин, особливо в умовах великої кількості нових, неперевіраних або маловідомих джерел. В майбутньому необхідно покращити алгоритми для обробки різноманітних форм медіаконтенту, таких як зображення, відео та аудіо, а також забезпечити більшу інтеграцію з міжнародними базами даних та інформаційними агентствами.

Загалом, розвиток автоматичних систем верифікації новин є важливим кроком у боротьбі з фейковими новинами, що має велике значення для збереження інформаційної безпеки та соціальної стабільності. В майбутньому такі системи можуть стати основними інструментами для забезпечення достовірності інформації в сучасному інформаційному середовищі.

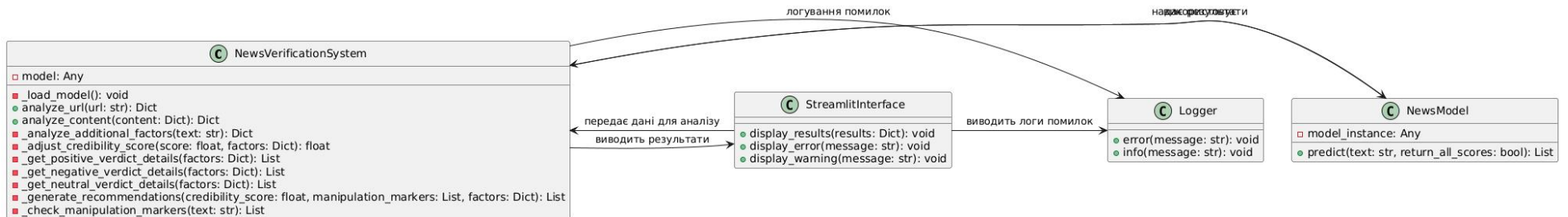
СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election.
2. Lazer, D. M., Baum, M. A., Benkler, Y., et al. (2018). The Science of Fake News. *Science*, 359(6380), 1094-1096
3. Vosoughi, S., Roy, D., & Aral, S. (2018). The Spread of True and False News Online. *Science*, 359(6380), 1146-1151
4. Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake News Detection on Social Media: A Data Mining Perspective. *ACM SIGKDD Explorations Newsletter*, 19(1), 22-36
5. Caplan, R., Hanson, L., & Donovan, J. (2018). Dead Reckoning: Navigating Content Moderation after "Fake News"
6. Conroy, N. J., Rubin, V. L., & Chen, Y. (2015). Automatic Deception Detection: Methods for Finding Fake News. *Proceedings of the Association for Information Science and Technology*, 52(1), 1-4.
7. Nguyen, T. T., Nguyen, C. M., Nguyen, D. T., Nguyen, D. D., & Nahavandi, S. (2019). Deep Learning for Deepfakes Creation and Detection: A Survey. *arXiv preprint arXiv:1909.11573*.
8. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv preprint arXiv:1810.04805*
9. Potthast, M., Kiesel, J., Reinartz, K., Bevendorff, J., & Stein, B. (2018). A Stylometric Inquiry into Hyperpartisan and Fake News. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 231-236.
10. Horne, B. D., & Adali, S. (2017). This Just In: Fake News Packs a Lot in Title, Uses Simpler, Repetitive Content in Text Body, More Similar to Satire than Real News. *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM)*.

11. Baly, R., Karadzhov, G., Alexandrov, D., Glass, J., & Nakov, P. (2018). Predicting Factuality of Reporting and Bias of News Media Sources. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 3528-3539.
12. Friggeri, A., Adamic, L. A., Eckles, D., & Cheng, J. (2014). Rumor Cascades. Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media.
13. Klein, D., & Wueller, J. R. (2017). Fake News: A Legal Perspective. Journal of Internet Law, 20(10), 1-13.
14. Pennycook, G., & Rand, D. G. (2018). The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Stories Increases Perceived Accuracy of Stories Without Warnings. Management Science, 66(11), 4944-4957.
15. Zubiaga, A., Aker, A., Bontcheva, K., Liakata, M., & Procter, R. (2018). Detection and Resolution of Rumours in Social Media: A Survey. ACM Computing Surveys, 51(2), 1-36.
16. Snopes URL: <https://en.wikipedia.org/wiki/Snopes>
17. Snopes URL: <https://www.snopes.com/>
18. factcheck.org URL: <https://en.wikipedia.org/wiki/FactCheck.org>
19. FactCheck.org URL: <https://www.factcheck.org/>
20. Reuters Fact Check URL: <https://www.reuters.com/fact-check/>
21. Agence France-Presse URL: https://en.wikipedia.org/wiki/Agence_France-Presse
22. Fact Check URL: <https://factcheck.afp.com/>
23. Reverse Image Search URL: <https://tinEye.com/>
24. Reverse Image Search Tool: Effortlessly Find Any Picture with Google! URL: <https://sites.google.com/view/reverse-images/home>
25. Computer Vision for OSINT URL: <https://vframe.io/>
26. The AI Workflow Orchestration Platform URL: <https://www.clarifai.com/>
27. Uwolnij swoją kreatywność dzięki obrazom, wideo i zdjęciom bez tantiem oraz innym zasobom URL: <https://stock.adobe.com/pl>

Додаток А

Діаграма класів



Додаток Б

Лістинг програми

```
import streamlit as st
import pandas as pd
import numpy as np
import torch
from transformers import pipeline, AutoModelForSequenceClassification, AutoTokenizer,
TrainingArguments, Trainer
from datasets import Dataset
import requests
from bs4 import BeautifulSoup
from sklearn.metrics import accuracy_score
from pathlib import Path
import logging
import warnings
import os
import re
from datetime import datetime
from kaggle.api.kaggle_api_extended import KaggleApi
from typing import Dict, List, Tuple, Optional

# Налаштування
warnings.filterwarnings('ignore')
os.environ['TF_CPP_MIN_LOG_LEVEL'] = '3'

logging.basicConfig(
    level=logging.INFO,
    format='%(asctime)s - %(name)s - %(levelname)s - %(message)s'
)
logger = logging.getLogger(__name__)

class DatasetLoader:
    """Клас для завантаження та підготовки датасетів з Kaggle"""

    def __init__(self):
        self.datasets_dir = Path("datasets")
        self.datasets_dir.mkdir(exist_ok=True)

        # Розширений список датасетів
        self.dataset_config = {
            'clmentbisaillon/fake-and-real-news-dataset': {
                'files': ['Fake.csv', 'True.csv'],
                'label_map': {'Fake.csv': 0, 'True.csv': 1}
            },
            'jainaldeed/fake-news-detection-dataset': {
                'files': ['train.csv'],
                'text_column': 'text',
                'label_column': 'label'
            },
            'bhuvanchennoj/fake-news-detection': {
                'files': ['train.csv'],
                'text_column': 'text',
                'label_column': 'label'
            },
            'amananandrai/fake-news-detection-dataset': {
                'files': ['news.csv'],
                'text_column': 'text',
                'label_column': 'label'
            },
            'hassanamin/textdb3': {
```

```

        'files': ['fake.csv', 'true.csv'],
        'label_map': {'fake.csv': 0, 'true.csv': 1}
    },
    'saurabhshahane/fake-news-classification': {
        'files': ['WELFake_Dataset.csv'],
        'text_column': 'text',
        'label_column': 'label'
    },
    'jeniaworks/fake-news': {
        'files': ['dataset.csv'],
        'text_column': 'content',
        'label_column': 'Class',
        'label_map': {'fake': 0, 'real': 1}
    },
    'raning/political-social-media-posts': {
        'files': ['train.csv'],
        'text_column': 'tweet_text',
        'label_column': 'class',
        'label_map': {'FAKE': 0, 'REAL': 1}
    },
    'algaedata/real-vs-fake-news-dataset': {
        'files': ['data.csv'],
        'text_column': 'text',
        'label_column': 'label',
        'label_map': {'fake': 0, 'real': 1}
    },
    'anaghahc/custom-fake-news-dataset': {
        'files': ['train.csv'],
        'text_column': 'text',
        'label_column': 'label'
    }
}

def download_from_kaggle(self) -> Optional[Dict[str, pd.DataFrame]]:
    """Завантаження датасетів з Kaggle"""
    try:
        api = KaggleApi()
        api.authenticate()

        datasets = {}

        for dataset_name, config in self.dataset_config.items():
            try:
                dataset_dir = self.datasets_dir / dataset_name.split('/')[1]

                # Перевіряємо чи вже є завантажений датасет
                if dataset_dir.exists() and any(dataset_dir.glob('*.csv')):
                    st.info(f"📁 Використання збереженого датасету: {dataset_name}")
                else:
                    st.info(f"📁 Завантаження {dataset_name}")
                    dataset_dir.mkdir(exist_ok=True)
                    api.dataset_download_files(
                        dataset_name,
                        path=str(dataset_dir),
                        unzip=True,
                        quiet=True
                    )

                # Зчитування даних
                for file in config['files']:
                    file_path = dataset_dir / file
                    if file_path.exists():
                        df = pd.read_csv(file_path)
                        df['text'] = df[config['text_column']]

```

```

        df['label'] = df[config['label_column']].map(config['label_map'])
        datasets[dataset_name] = df[['text', 'label']]

    except Exception as e:
        st.warning(f"⚠ Пропуск {dataset_name}")
        continue

    if datasets:
        total_samples = sum(len(df) for df in datasets.values())
        st.success(f"✅ Завантажено {len(datasets)} датасети, всього {total_samples:,} зразків")
        return datasets
    return None

    except Exception as e:
        st.error("❌ Помилка доступу до Kaggle API")
        return None

def prepare_data(self, datasets: Dict[str, pd.DataFrame]) -> Tuple[pd.DataFrame,
pd.DataFrame]:
    """Підготовка даних для навчання"""
    try:
        # Об'єднання датасетів
        combined_df = pd.concat(datasets.values(), ignore_index=True)

        # Очистка текстів
        combined_df['text'] = combined_df['text'].fillna('').apply(self._clean_text)
        combined_df = combined_df[combined_df['text'].str.len() > 100]

        # Балансування класів
        min_samples = combined_df['label'].value_counts().min()
        balanced_df = pd.concat([
            combined_df[combined_df['label'] == label].sample(n=min_samples,
random_state=42)
            for label in [0, 1]
        ])

        # Розділення на тренувальну та тестову вибірки
        train_df = balanced_df.sample(frac=0.8, random_state=42)
        test_df = balanced_df.drop(train_df.index)

        st.success(f"✅ Підготовлено {len(train_df):,} тренувальних та {len(test_df):,} тестових зразків")
        return train_df, test_df

    except Exception as e:
        st.error("❌ Помилка підготовки даних")
        return None, None

def _clean_text(self, text: str) -> str:
    """Очистка тексту"""
    if not isinstance(text, str):
        return ""
    text = re.sub(r'http\S+|www.\S+', '', text)
    text = re.sub(r'<[^>]+>', '', text)
    text = re.sub(r'^\w\s.,!?«»\--', '', text)
    text = re.sub(r'\s+', ' ', text)
    return text.strip()

class NewsAnalyzer:
    """Клас для аналізу та отримання тексту новин"""

    def __init__(self):

```

```

self.headers = {
    'User-Agent': 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36'
}

def extract_content(self, url: str) -> Optional[Dict[str, str]]:
    """Отримання контенту статті"""
    try:
        response = requests.get(url, headers=self.headers, timeout=10, verify=False)
        response.raise_for_status()
        response.encoding = response.apparent_encoding

        soup = BeautifulSoup(response.text, 'html.parser')

        # Видалення непотрібних елементів
        for tag in soup(['script', 'style', 'iframe', 'nav', 'footer', 'header',
'aside']):
            tag.decompose()

        # Пошук заголовка
        title = None
        title_selectors = [
            'h1',
            'h1.article-title',
            'h1.entry-title',
            'h1.post-title',
            '.article-title',
            '.post-title',
            '.entry-title'
        ]

        for selector in title_selectors:
            title = soup.select_one(selector)
            if title:
                break

        if not title:
            title = soup.find('title')

        title = self._clean_text(title.get_text()) if title else ""

        # Пошук основного тексту
        article_text = []

        # Спробуємо знайти основний контейнер
        content_selectors = [
            'article',
            '.article-content',
            '.post-content',
            '.entry-content',
            '.content',
            'main'
        ]

        main_content = None
        for selector in content_selectors:
            main_content = soup.select_one(selector)
            if main_content:
                break

        if not main_content:
            main_content = soup

        # Збір параграфів
        for p in main_content.find_all(['p', 'h2', 'h3', 'h4']):
            text = self._clean_text(p.get_text())

```

```

        if len(text) > 30: # Ігноруємо короткі фрагменти
            article_text.append(text)

    if not article_text:
        logger.error(f"Не знайдено текст статті для {url}")
        return None

    text = ' '.join(article_text)

    # Базова перевірка якості контенту
    if len(text) < 100 or len(title) < 5:
        logger.error(f"Замало тексту для аналізу: {len(text)} символів")
        return None

    return {
        'title': title,
        'text': text,
        'url': url,
        'date': datetime.now().strftime("%Y-%m-%d %H:%M:%S")
    }

except requests.RequestException as e:
    logger.error(f"Помилка запиту до {url}: {str(e)}")
    return None
except Exception as e:
    logger.error(f"Помилка отримання контенту з {url}: {str(e)}")
    return None

def _clean_text(self, text: str) -> str:
    """Очищення тексту"""
    if not text:
        return ""

    # Видалення HTML тегів
    text = re.sub(r'<[^>]+>', '', text)

    # Видалення множинних пробілів
    text = re.sub(r'\s+', ' ', text)

    # Видалення пробілів на початку і в кінці
    text = text.strip()

    return text

class NewsVerificationSystem:
    """Система верифікації новин"""

    def __init__(self):
        self.device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')
        self.models_dir = Path('models')
        self.models_dir.mkdir(exist_ok=True)

        # Перевіряємо чи є навчена модель
        model_path = self.models_dir / 'news_classifier'
        if model_path.exists() and list(model_path.glob('*.bin')):
            st.info("📁 Завантаження збереженої моделі...")
            try:
                self.model = pipeline(
                    'text-classification',
                    model=str(model_path),
                    device=0 if torch.cuda.is_available() else -1
                )
            except:
                pass

```

```

        st.success("✅ Збережено модель завантажено успішно")
    except Exception as e:
        st.warning(f"⚠️ Помилка завантаження збереженої моделі: {str(e)}")
        st.info("🔄 Спроба навчання нової моделі...")
        self._train_new_model()
    else:
        st.info("🔄 Перше навчання моделі...")
        self._train_new_model()

self.news_analyzer = NewsAnalyzer()

def _load_pretrained_model(self):
    """Завантаження попередньо навченої моделі"""
    try:
        st.info("📁 Завантаження базової моделі...")
        self.model = pipeline(
            'text-classification',
            model='microsoft/mdeberta-v3-base',
            device=0 if torch.cuda.is_available() else -1
        )
        st.success("✅ Базову модель завантажено")
    except Exception as e:
        st.error(f"❌ Критична помилка завантаження моделі: {str(e)}")
        raise

def _check_manipulation_markers(self, text: str) -> List[str]:
    """Перевірка маніпулятивних маркерів"""
    markers = [
        r'всі знають',
        r'очевидно',
        r'безсумнівно',
        r'звичайно ж',
        r'без сумніву',
        r'ні для кого не секрет',
        r'загальновідомо',
        r'як відомо',
        r'всім відомо',
        r'експерти вважають',
        r'більшість вважає',
        r'всі розуміють',
        r'ніхто не сумнівається',
        r'тільки дурень не розуміє',
        r'істина в тому, що'
    ]

    found_markers = []
    for marker in markers:
        if re.search(marker, text.lower()):
            found_markers.append(marker)

    return found_markers

def _train_new_model(self):
    """Навчання нової моделі"""
    try:
        # Завантаження датасетів
        with st.spinner("📁 Завантаження датасетів..."):
            dataset_loader = DatasetLoader()
            datasets = dataset_loader.download_from_kaggle()

        if not datasets:
            st.warning("⚠️ Не вдалося завантажити датасети")

```

```

        return self._load_pretrained_model()

# Підготовка даних
with st.spinner("🔄 Підготовка даних..."):
    train_df, test_df = dataset_loader.prepare_data(datasets)

if train_df is None or len(train_df) < 1000:
    st.warning("⚠ Недостатньо даних для навчання")
    return self._load_pretrained_model()

progress_text = st.empty()
progress_bar = st.progress(0)

progress_text.text("🌀 Ініціалізація моделі...")
progress_bar.progress(0.1)

# Підготовка моделі
model_name = 'microsoft/mdeberta-v3-base'
tokenizer = AutoTokenizer.from_pretrained(model_name)
model = AutoModelForSequenceClassification.from_pretrained(
    model_name,
    num_labels=2
).to(self.device)

progress_text.text("🔄 Підготовка даних для навчання...")
progress_bar.progress(0.3)

# Підготовка даних
train_encodings = tokenizer(
    train_df['text'].tolist(),
    truncation=True,
    padding=True,
    max_length=512,
    return_tensors='pt'
)

train_dataset = Dataset.from_dict({
    'input_ids': train_encodings['input_ids'],
    'attention_mask': train_encodings['attention_mask'],
    'labels': train_df['label'].tolist()
})

progress_text.text("🌀 Налаштування навчання...")
progress_bar.progress(0.4)

# Налаштування навчання
training_args = TrainingArguments(
    output_dir=str(self.models_dir / 'news_classifier'),
    num_train_epochs=3,
    per_device_train_batch_size=8,
    warmup_steps=500,
    weight_decay=0.01,
    logging_dir=str(self.models_dir / 'logs'),
    save_strategy="epoch",
    logging_steps=10
)

# Навчання
trainer = Trainer(
    model=model,
    args=training_args,
    train_dataset=train_dataset,
    tokenizer=tokenizer

```

```

    )

    trainer.train()

    progress_text.text("📄 Збереження моделі...")
    progress_bar.progress(0.95)

    # Збереження моделі
    model_path = self.models_dir / 'news_classifier'
    model.save_pretrained(str(model_path))
    tokenizer.save_pretrained(str(model_path))

    # Створення пайплайну
    self.model = pipeline(
        'text-classification',
        model=str(model_path),
        device=0 if torch.cuda.is_available() else -1
    )

    progress_text.text("✅ Навчання завершено успішно!")
    progress_bar.progress(1.0)

except Exception as e:
    st.error(f"❌ Помилка навчання: {str(e)}")
    logger.error(f"Training error: {str(e)}")
    return self._load_pretrained_model()

def analyze_url(self, url: str) -> Optional[Dict]:
    """Аналіз статті за URL"""
    content = self.news_analyzer.extract_content(url)
    if not content:
        return None
    return self.analyze_content(content)

def analyze_content(self, content: Dict) -> Dict:
    """Аналіз контенту"""
    try:
        # Підготовка тексту
        title = content['title'].strip()
        text = content['text'].strip()

        # Аналізуємо заголовок і текст окремо
        title_result = self.model(title, return_all_scores=True)[0]
        text_result = self.model(text[:512], return_all_scores=True)[0]

        # Отримуємо оцінки для кожного класу
        title_scores = {score['label']: score['score'] for score in title_result}
        text_scores = {score['label']: score['score'] for score in text_result}

        # Розрахунок фінальної оцінки з різними вагами
        title_weight = 0.3
        text_weight = 0.7

        credibility_score = (
            title_weight * title_scores.get('LABEL_1', 0.5) +
            text_weight * text_scores.get('LABEL_1', 0.5)
        )

        # Додаткові фактори для оцінки достовірності
        factors = self._analyze_additional_factors(text)

        # Коригування оцінки на основі додаткових факторів
        credibility_score = self._adjust_credibility_score(
            credibility_score,

```

```

        factors
    )

    # Аналіз маніпуляцій
    manipulation_markers = self._check_manipulation_markers(text)
    manipulation_score = len(manipulation_markers) / 10

    # Зменшення достовірності при наявності маніпуляцій
    credibility_score -= manipulation_score * 0.2
    credibility_score = max(0.1, min(0.9, credibility_score))

    # Визначення вердикту
    if credibility_score > 0.7:
        verdict = "🟢 Ймовірно правда"
        verdict_details = self._get_positive_verdict_details(factors)
    elif credibility_score < 0.4:
        verdict = "🔴 Ймовірно фейк"
        verdict_details = self._get_negative_verdict_details(factors)
    else:
        verdict = "🟡 Сумнівна достовірність"
        verdict_details = self._get_neutral_verdict_details(factors)

    return {
        'content': content,
        'analysis': {
            'verdict': verdict,
            'verdict_details': verdict_details,
            'metrics': {
                'credibility': credibility_score,
                'fake': 1 - credibility_score,
                'manipulation': manipulation_score,
                'factors': factors
            },
            'title_analysis': {
                'credibility': title_scores.get('LABEL_1', 0.5),
                'fake': title_scores.get('LABEL_0', 0.5)
            },
            'text_analysis': {
                'credibility': text_scores.get('LABEL_1', 0.5),
                'fake': text_scores.get('LABEL_0', 0.5)
            },
            'manipulation_markers': manipulation_markers,
            'recommendations': self._generate_recommendations(
                credibility_score,
                manipulation_markers,
                factors
            )
        }
    }

except Exception as e:
    logger.error(f"Помилка аналізу: {str(e)}")
    return None

def _analyze_additional_factors(self, text: str) -> Dict[str, float]:
    """Аналіз додаткових факторів достовірності"""
    factors = {
        'has_dates': 0.0,
        'has_numbers': 0.0,
        'has_quotes': 0.0,
        'has_sources': 0.0,
        'emotional_level': 0.0
    }

```

```

# Пошук дат
date_patterns = [
    r'\d{1,2}\.\d{1,2}\.\d{4}',
    r'\d{1,2}\s(?:січня|лютого|березня|квітня|травня|червня|липня|серпня|вересня|жовтня|листопад
а|грудня)\s?\d{4}',
    r'(?:вчора|сьогодні|минулого тижня|цього місяця)'
]
for pattern in date_patterns:
    if re.search(pattern, text, re.IGNORECASE):
        factors['has_dates'] += 0.3

# Пошук чисел і статистики
number_patterns = [
    r'\d+(?:,\d+)?%',
    r'\d+(?:,\d+)?\s?(?:гривень|грн|доларів|\$|євро|€)',
    r'\d+\s?(?:осіб|людей|чоловік)'
]
for pattern in number_patterns:
    if re.search(pattern, text):
        factors['has_numbers'] += 0.2

# Пошук цитат
quote_patterns = [
    r'«[^»]+»',
    r'"[^"]+"',
    r'"s*"[^"]+s*'
]
for pattern in quote_patterns:
    if re.search(pattern, text):
        factors['has_quotes'] += 0.25

# Пошук джерел
source_patterns = [
    r'(?:за даними|за словами|як повідомляє|джерело|згідно з|відповідно до)',
    r'(?:прес-служба|міністерство|відомство|організація)',
    r'(?:заявив|повідомив|зазначив|підкреслив|наголосив)'
]
for pattern in source_patterns:
    if re.search(pattern, text, re.IGNORECASE):
        factors['has_sources'] += 0.25

# Оцінка емоційності
emotional_words = [
    r'(?:жахливий|страшний|небезпечний|критичний|катастрофічний)',
    r'(?:неймовірний|дивовижний|приголомшливий|сенсаційний)',
    r'(?:терміново|негайно|виключно|тільки|єдиний)'
]
for pattern in emotional_words:
    if re.search(pattern, text, re.IGNORECASE):
        factors['emotional_level'] += 0.2

# Нормалізація значень
for key in factors:
    factors[key] = min(1.0, factors[key])

return factors

def _adjust_credibility_score(self, score: float, factors: Dict[str, float]) -> float:
    """Коригування оцінки достовірності на основі додаткових факторів"""

    # Позитивні фактори підвищують достовірність
    positive_adjustment = (
        factors['has_dates'] * 0.1 +

```

```

        factors['has_numbers'] * 0.1 +
        factors['has_quotes'] * 0.15 +
        factors['has_sources'] * 0.15
    )

    # Емоційність знижує достовірність
    negative_adjustment = factors['emotional_level'] * 0.2

    # Застосування коригувань
    adjusted_score = score + positive_adjustment - negative_adjustment

    # Обмеження значення
    return max(0.1, min(0.9, adjusted_score))

def _get_positive_verdict_details(self, factors: Dict[str, float]) -> List[str]:
    """Деталі для позитивного вердикту"""
    details = []
    if factors['has_dates'] > 0:
        details.append("✓ Наявні конкретні дати")
    if factors['has_numbers'] > 0:
        details.append("✓ Використано точні числові дані")
    if factors['has_quotes'] > 0:
        details.append("✓ Є прямі цитати")
    if factors['has_sources'] > 0:
        details.append("✓ Вказані джерела інформації")
    return details

def _get_negative_verdict_details(self, factors: Dict[str, float]) -> List[str]:
    """Деталі для негативного вердикту"""
    details = []
    if factors['emotional_level'] > 0.5:
        details.append("⚠ Високий рівень емоційності")
    if factors['has_dates'] < 0.2:
        details.append("⚠ Відсутні конкретні дати")
    if factors['has_sources'] < 0.2:
        details.append("⚠ Не вказані джерела інформації")
    return details

def _get_neutral_verdict_details(self, factors: Dict[str, float]) -> List[str]:
    """Деталі для нейтрального вердикту"""
    details = []
    if factors['has_quotes'] > 0:
        details.append("i Наявні цитати потребують перевірки")
    if factors['has_sources'] > 0:
        details.append("i Варто перевірити вказані джерела")
    if factors['emotional_level'] > 0.3:
        details.append("i Присутні емоційні висловлювання")
    return details

def _generate_recommendations(
    self,
    credibility_score: float,
    manipulation_markers: List[str],
    factors: Dict[str, float]
) -> List[str]:
    """Генерація рекомендацій"""
    recommendations = []

    if credibility_score < 0.4:
        recommendations.extend([
            "⚠ Ця новина має ознаки недостовірної інформації",
            "🔍 Перевірте інформацію в офіційних джерелах",
            "⚡ Утримайтесь від поширення без перевірки"
        ])

```

```

    ])
elif credibility_score < 0.7:
    recommendations.extend([
        "⚠ Потрібна додаткова перевірка",
        "🔍 Порівняйте з іншими джерелами",
        "👁 Зверніть увагу на контекст"
    ])
else:
    recommendations.extend([
        "✅ Інформація має ознаки достовірності",
        "👁 Можна перевірити додаткові джерела"
    ])

if manipulation_markers:
    recommendations.append(
        "⚠ Виявлено маніпулятивні звороти: " +
        ", ".join(manipulation_markers)
    )
    recommendations.append(
        "🌀 Будьте уважні до емоційних маніпуляцій"
    )

# Додаткові рекомендації на основі факторів
if factors['emotional_level'] > 0.5:
    recommendations.append(
        "⚠ Текст містить підвищену емоційність"
    )
if factors['has_sources'] < 0.2:
    recommendations.append(
        "🔍 Рекомендується перевірити джерела інформації"
    )

return recommendations

def main():
    st.set_page_config(
        page_title="Верифікація новин",
        page_icon="📰",
        layout="wide"
    )

    try:
        if 'system' not in st.session_state:
            with st.spinner("🌀 Ініціалізація системи..."):
                st.session_state.system = NewsVerificationSystem()

        st.title("🔍 Система верифікації новин")

        st.write("""
        ### Як це працює:
        1. Вставте URL новини для аналізу
        2. Система проаналізує текст за допомогою штучного інтелекту
        3. Ви отримаєте оцінку достовірності та рекомендації
        """)

        url = st.text_input(
            "URL новини:",
            placeholder="https://example.com/news"
        )

        if st.button("Аналізувати"):
            if url:
                with st.spinner("📊 Аналіз статті..."):
                    results = st.session_state.system.analyze_url(url)

```

```

if results:
    content = results['content']
    analysis = results['analysis']

    # Відображення результатів
    st.subheader("📄 Знайдена стаття")
    st.write(f"**Заголовок:** {content['title']}")
    st.write(f"**Дата аналізу:** {content['date']}")

    # Основний вердикт
    st.subheader("📊 Результати аналізу")
    st.markdown(f"### {analysis['verdict']}")

    # Метрики
    col1, col2, col3 = st.columns(3)
    with col1:
        st.metric(
            "Достовірність",
            f"{analysis['metrics']['credibility']:.1%}"
        )
    with col2:
        st.metric(
            "Ймовірність фейку",
            f"{analysis['metrics']['fake']:.1%}"
        )
    with col3:
        st.metric(
            "Рівень маніпуляцій",
            f"{analysis['metrics']['manipulation']:.1%}"
        )

    # Маніпулятивні маркери
    if analysis['manipulation_markers']:
        st.subheader("⚠️ Виявлені маніпулятивні звороти")
        for marker in analysis['manipulation_markers']:
            st.markdown(f"- {marker}")

    # Рекомендації
    st.subheader("💡 Рекомендації")
    for rec in analysis['recommendations']:
        st.markdown(rec)

    # Текст статті
    with st.expander("📄 Показати текст статті"):
        st.write(content['text'])
    else:
        st.error("❌ Не вдалося проаналізувати статтю. Перевірте URL.")
else:
    st.warning("⚠️ Будь ласка, введіть URL новини.")

# Сайдбар з інформацією
with st.sidebar:
    st.header("📖 Про систему")
    st.write("""
    Система використовує:
    - Штучний інтелект для аналізу тексту
    - Перевірку маніпулятивних зворотів
    - Оцінку достовірності джерел
    """)

    st.header("🔗 Технічна інформація")
    st.write(f"""

```

```

- Статус:  Готова до роботи
- Пристрій: {'GPU' if torch.cuda.is_available() else 'CPU'}
""")

if st.button(" Перенавчати модель"):
    with st.spinner(" Перенавчання моделі..."):
        if 'system' in st.session_state:
            del st.session_state['system']
        if (Path('models') / 'news_classifier').exists():
            import shutil
            shutil.rmtree(Path('models') / 'news_classifier')
        st.rerun()

if st.button(" Перезавантажити систему"):
    st.session_state.pop('system', None)
    st.rerun()

except Exception as e:
    st.error(f" Критична помилка: {str(e)}")
    logger.error(f"Critical error: {str(e)}", exc_info=True)
    st.info(" Спробуйте перезавантажити сторінку")

if __name__ == "__main__":
    main()

```