

ДЕРЖАВНИЙ УНІВЕРСИТЕТ ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ

Навчально-науковий інститут захисту інформації

Кафедра управління інформаційною та кібернетичною безпекою

Ступінь вищої освіти бакалавр

Спеціальність 125 Кібербезпека

Освітня програма Управління інформаційною та кібернетичною безпекою

ЗАТВЕРДЖУЮ

Завідувач кафедри УІКБ

_____ Світлана ЛЕГОМІНОВА

“ ____ ” _____ 2024 р.

ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ

Слинько Владислав Вадимович

(прізвище, ім'я, по батькові здобувача)

1. Тема кваліфікаційної роботи “Використання глибокого навчання для виявлення та захисту від фейків у великих обсягах текстової інформації”, керівник кваліфікаційної роботи ТИЩЕНКО Віталій

_____ (ПРІЗВИЩЕ, Ім'я, науковий ступінь, вчене звання)

затверджені наказом Державного університету інформаційно-комунікаційних технологій від “27” лютого 2024 р. №36.

2. Строк подання кваліфікаційної роботи “20” травня 2024р.
3. Вихідні дані до кваліфікаційної роботи: *штучний інтелект, глибоке навчання, обробка текстів, методи захисту від фейків у великих обсягах текстової інформації, наукова та технічна література.*
4. Перелік питань, які мають бути розроблені:
 - 4.1. Проаналізувати існуючі методи детекції з використанням глибокого навчання для виявлення та захисту від фейків у великих обсягах текстової інформації.
 - 4.2. Дослідити алгоритми класифікації текстової інформації.
 - 4.3. Розробити алгоритми класифікації, впровадження системи у реальні інформаційні середовища та її тестування.
5. Перелік ілюстративного матеріалу: *презентація PowerPoint*
6. Дата видачі завдання “11” березня 2024 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Етапи кваліфікаційної роботи	Термін виконання Етапів роботи	Примітка
1.	Визначення об'єкту, предмету, мети та завдань дослідження.	18.03.2024	
2.	Збір та аналіз літератури.	29.03.2024	
3.	Аналіз особливостей використання глибокого навчання	08.04.2024	
4.	Дослідження основних характеристик технологій	22.04.2024	
5.	Вивчення інструментів та методів формування	08.05.2024	
6.	Формулювання висновків за результатами проведеного дослідження.	20.05.2024	
7.	Оформлення роботи.	22.05.2024	
8.	Оформлення презентації.	23.05.2024	
9.	Отримання рецензії на роботу.	03.06.2024	
10.	Захист в ЕК.	11.06.2024	

Здобувач вищої освіти

(підпис)

Владислав СЛИНЬКО

(Ім'я, ПРІЗВИЩЕ)

Керівник
кваліфікаційної роботи

(підпис)

Віталій ТИЩЕНКО

(Ім'я, ПРІЗВИЩЕ)

**ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ЗАХИСТУ ІНФОРМАЦІЇ**

**ПОДАННЯ
ГОЛОВІ ЕКЗАМЕНАЦІЙНОЇ КОМІСІЇ
ЩОДО ЗАХИСТУ КВАЛІФІКАЦІЙНОЇ РОБОТИ
на здобуття освітнього ступеня бакалавра**

Направляється здобувач Слинько В.В до захисту кваліфікаційної роботи
(*прізвище та ініціали*)

за спеціальністю 125 Кібербезпека

(*код, найменування спеціальності*)

освітньої програми Управління інформаційною та кібернетичною безпекою
(*назва*)

на тему: “Використання глибокого навчання для виявлення та захисту від фейків у великих
обсягах текстової інформації”

Кваліфікаційна робота і рецензія додаються.

Директор ННІЗІ _____

(*підпис*)

Віталій САВЧЕНКО

(*Ім'я, ПРІЗВИЩЕ*)

Висновок керівника кваліфікаційної роботи

Здобувач СЛИНЬКО Владислав у кваліфікаційній роботі проаналізував ефективність та ризики використання глибокого навчання для виявлення фейкових новин у текстовому контенті, дослідив основні характеристики технологій глибокого навчання для виявлення та реагування на дезінформацію, вивчив інструменти та методи оптимізації процесів виявлення та захисту від фейкових новин за допомогою глибокого навчання, а також розробив практичні рекомендації.

СЛИНЬКО Владислав показав розуміння проблеми дослідження та бачення основних теоретичних і практичних напрямів її вирішення, довів володіння методами наукового дослідження, проявив себе як організований, відповідальний виконавець. Результати дослідження апробовані на одній конференції.

Все це дозволяє оцінити кваліфікаційну роботу здобувача СЛИНЬКА Владислава на оцінку “_____” та присвоїти йому кваліфікацію бакалавра з кібербезпеки за освітньою програмою Управління інформаційною та кібернетичною безпекою.

Керівник кваліфікаційної роботи _____
(*підпис*)

Віталій ТИЩЕНКО
(*Ім'я, ПРІЗВИЩЕ*)

“_____” _____ 2024 року

Висновок кафедри про кваліфікаційну роботу

Кваліфікаційна робота розглянута. Здобувач _____ допускається до захисту даної роботи в Екзаменаційній комісії.

Завідувач кафедри
управління інформаційною
та кібернетичною безпекою

(*підпис*)

Світлана ЛЕГОМІНОВА
(*Ім'я, ПРІЗВИЩЕ*)

ВІДГУК РЕЦЕНЗЕНТА на кваліфікаційну бакалаврську роботу

здобувача вищої освіти СЛИНЬКА Владислава

на тему “ Використання глибокого навчання для виявлення та захисту від фейків у великих обсягах текстової інформації ”

Актуальність. У світі, де держави та компанії є об’єктами кібератак, важливість забезпечення інформаційної безпеки є надзвичайно великою. Тому важливо формувати комплексний підхід, який включає технічні засоби для ефективного протистояння сучасним кіберзагрозам, зменшуючи ризики втрат даних та фінансових збитків. Аналіз підходів до виявлення та реагування на кібератаки може допомогти у протистоянні сучасним кіберзагрозам, зменшуючи ризики втрат даних та фінансових збитків. З огляду на зазначене, дослідження ефективності та ризиків використання штучного інтелекту в кібербезпеці є актуальним науковим завданням.

Позитивні сторони.

1. У роботі детально розглянуто використання глибокого навчання для виявлення фейкових новин, що свідчить про глибоке розуміння автором сучасних технологій в цій сфері.

2. Кваліфікаційна робота оформлена відповідно до вимог. Матеріал викладено логічно та послідовно, зроблено чіткі висновки. Ключові положення роботи ілюстровано рисунками, що покращує розуміння матеріалу.

3. За результатами дослідження запропоновано рекомендації щодо оптимізації процесів виявлення та захисту від фейкових новин за допомогою глибокого навчання, що має практичну цінність.

Недоліки.

Доцільно було б приділити більше уваги вивченню і класифікації програмних інструментів для оцінки ефективності процесів виявлення та захисту від фейкових новин. Це допомогло б краще зрозуміти та порівняти різні підходи в цій сфері.

Однак, вищезгадані зауваження не впливають на загальну позитивну оцінку кваліфікаційної роботи.

Висновок Кваліфікаційна робота виконана на належному науково-методичному рівні і заслуговує оцінки “ _____”, а здобувач СЛИНЬКО Владислав заслуговує присвоєння кваліфікації бакалавра з кібербезпеки за освітньою програмою Управління інформаційною та кібернетичною безпекою.

Рецензент:

підпис

Ім’я, ПРИЗВИЩЕ

РЕФЕРАТ

Кваліфікаційна робота присвячена дослідженню використання глибокого навчання для виявлення та захисту від фейків у великих обсягах текстової інформації. Робота складається зі вступу, трьох розділів, що містять 4 таблиці, висновків і списку використаних джерел із 52 найменувань. Загальний обсяг роботи становить 66 аркушів, з яких 6 аркушів займають список використаних джерел.

Метою роботи є розробка ефективної системи для виявлення та захисту від фейкових новин та дезінформації у великих обсягах текстової інформації, використовуючи методи глибокого навчання.

Об'єктом дослідження процеси виявлення та боротьби з поширенням фейкових новин і дезінформації в текстовому контенті за допомогою методів глибокого навчання.

Предмет дослідження – методи і технології виявлення фейкових новин та дезінформації.

Методи дослідження. Для вирішення означеного вище наукового завдання в роботі використані методи аналізу та синтезу, порівняння, класифікації

У ході роботи було проаналізовано особливості використанні глибокого навчання для виявлення та захисту від фейків у великих обсягах текстової інформації, зокрема досліджено ефективність та ризики використання цих технологій. Розглянуто основні характеристики технологій глибокого навчання, вивчено інструменти та методи, які сприяють підвищенню обізнаності щодо використання глибокого навчання для виявлення фейків. На основі проведеного аналізу розроблено практичні рекомендації.

Галузь застосування. Розробка інноваційного алгоритму, заснованого на глибокому навчанні, який дозволяє з високою точністю класифікувати інформацію на фейкову у великих обсягах текстової інформації. На відміну від

існуючих рішень, нова система використовує комбінацію текстового аналізу та аналізу метаданих, що значно підвищує її ефективність.

Ключові слова: ГЛИБОКЕ НАВЧАННЯ, ВИЯВЛЕННЯ ФЕЙКОВИХ НОВИН, ДЕЗІНФОРМАЦІЯ, ОБРОБКА ТЕКСТОВОЇ ІНФОРМАЦІЇ, МАШИННЕ НАВЧАННЯ, КЛАСИФІКАЦІЯ ТЕКСТУ, АЛГОРИТМИ, СОЦІАЛЬНІ МЕРЕЖІ, АВТОМАТИЧНЕ ВИЯВЛЕННЯ, ЗАХИСТ ВІД ФЕЙКІВ.

ABSTRACT

The qualification work is devoted to the study of the use of deep learning to detect and protect against fakes in large volumes of textual information. The work consists of an introduction, three chapters containing 4 tables, conclusions and a list of used sources of 52 names. The total volume of work is 66 sheets, of which 6 sheets are list of used sources.

The purpose of the work is to develop an effective system for detecting and protecting against fake news and disinformation in large amounts of textual information using deep learning methods.

The object of research is the processes of detecting and combating the spread of fake news and disinformation in textual content using deep learning methods.

The subject of the research is methods and technologies for detecting fake news and disinformation.

Research methods. To solve the above scientific task, the methods of analysis and synthesis, comparison, classification were used in the study

In the course of the work, the features of using deep learning to detect and protect against fakes in large volumes of textual information were analyzed, in particular, the effectiveness and risks of using these technologies were investigated. The main characteristics of deep learning technologies are considered, the tools and methods that contribute to raising awareness of the use of deep learning for detecting fakes are studied. Based on the analysis, practical recommendations were developed.

Field of application. Development of an innovative algorithm based on deep learning, which allows to classify information as fake in large volumes of textual information with high accuracy. Unlike existing solutions, the new system uses a combination of text analysis and metadata analysis, which significantly increases its efficiency.

Keywords: DEEP LEARNING, FAKE NEWS DETECTION, DISINFORMATION, TEXT INFORMATION PROCESSING, MACHINE

LEARNING, TEXT CLASSIFICATION, ALGORITHMS, SOCIAL MEDIA,
AUTOMATIC DETECTION, FAKE NEWS PROTECTION.

ЗМІСТ

ВСТУП	11
РОЗДІЛ 1. ФЕНОМЕН ФЕЙКОВИХ НОВИН: КОНЦЕПТУАЛІЗАЦІЯ ТА ТИПОЛОГІЯ	13
1.1 Огляд проблематики фейкових новин	13
1.2 Методи виявлення фейкових новин.....	16
1.3 Виклики та проблеми	19
Висновки до розділу 1	24
РОЗДІЛ 2. ТЕОРЕТИЧНІ ОСНОВИ ГЛИБОКОГО НАВЧАННЯ	26
2.1 Основи глибокого навчання	26
2.2 Моделі для класифікації тексту.....	31
2.3 Метрики оцінки моделей	35
Висновки до розділу 2	40
РОЗДІЛ 3. РОЗРОБКА ТА ВПРОВАДЖЕННЯ СИСТЕМИ ВИЯВЛЕННЯ ФЕЙКОВИХ НОВИН	42
3.1 Збір та підготовка даних	42
3.2 Навчання та налаштування моделі	47
3.3 Впровадження та тестування системи.....	51
Висновки до розділу 3	59
ВИСНОВКИ	61
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	63

ВСТУП

В сучасному світі інформаційні технології відіграють ключову роль у поширенні новин та інформації. Зі збільшенням обсягів текстової інформації, що циркулює в медіа та соціальних мережах, зростає й кількість фейкових новин та дезінформації. Фейки можуть мати серйозні наслідки, впливаючи на суспільну думку, політичні процеси, економіку та навіть безпеку держав.

Актуальність теми: В епоху цифрових технологій, коли обмін інформацією відбувається миттєво і в безпрецедентних масштабах, проблема фейкових новин та дезінформації стає все більш актуальною. Фейки можуть завдати значної шкоди як окремим особам, так і суспільству в цілому, впливаючи на політичні процеси, економічні рішення, громадську думку та навіть міжнародні відносини. Зростання кількості фейкових новин у медіа та соціальних мережах створює потребу в ефективних інструментах для їх виявлення та протидії.

Метою роботи є розробка та впровадження ефективних методів глибокого навчання для автоматичного виявлення та захисту від фейкових новин у великих обсягах текстової інформації. Це включає створення моделей, здатних точно класифікувати тексти як достовірні чи фейкові, аналіз існуючих підходів та їх удосконалення, а також оцінку практичної ефективності розроблених рішень у реальних умовах.

Завданнями є аналіз існуючих методів детекції, розробка алгоритмів класифікації, впровадження системи у реальні інформаційні середовища та її тестування.

Об'єктом дослідження є процеси обробки інформації в медіа. **Предметом** — методи і технології виявлення фейкових новин.

Основна наукова новизна полягає у розробці інноваційного алгоритму, заснованого на глибокому навчанні, який дозволяє з високою точністю класифікувати інформацію на достовірну та фейкову. На відміну від існуючих

рішень, нова система використовує комбінацію текстового аналізу та аналізу метаданих, що значно підвищує її ефективність.

Результати дослідження можуть бути використані в медіа та соціальних мережах для автоматичного виявлення та запобігання поширенню фейкових новин, а також у державних та громадських організаціях для боротьби з пропагандою та дезінформацією. Крім того, ці результати можуть допомогти освітнім закладам підвищувати медіаграмотність і корпоративному сектору захищати репутацію брендів від негативного впливу дезінформації.

Робота структурована у три основні розділи, що дозволяють охопити ключові аспекти цієї масштабної теми. Перший розділ ретельно аналізує феномен фейкових новин, їх типи та методи розповсюдження, вторинний розділ зосереджується на теоретичних основах глибокого навчання як інструменту виявлення дезінформації, а третій розділ відображає практичний процес розробки та впровадження системи детекції фейкових новин, починаючи від збору даних до оцінки її ефективності.

Дипломна робота спрямована на аналіз проблематики фейкових новин, вивчення існуючих методів їх виявлення та розробку власної моделі для їх ідентифікації. Важливість дослідження посилюється нестабільним інформаційним контекстом сучасного світу, де маніпуляції інформацією стають все більш витонченими та складними для виявлення. Це зумовлює необхідність розробки ефективних, адаптивних та інноваційних технологічних рішень, здатних адекватно реагувати на виклики інформаційної безпеки.

РОЗДІЛ 1. ФЕНОМЕН ФЕЙКОВИХ НОВИН: КОНЦЕПТУАЛІЗАЦІЯ ТА ТИПОЛОГІЯ

1.1 Огляд проблематики фейкових новин

У сучасному світі проблема фейкових новин стала особливо актуальною через їхній потенційний вплив на суспільні процеси, політику та особисте життя людей. Дослідження з цієї теми розглядають фейкові новини з різних сторін, зокрема їх вплив, методи виявлення та стратегії боротьби з ними.

Феномен фейкових новин набув особливого значення в сучасному медійному просторі через його вплив на суспільні процеси, політику та особисте життя людей. Дослідження в цій області виявляють, що фейкові новини - це неправдива або маніпулятивна інформація, що може бути створена з метою впливу на громадську думку або для отримання фінансової вигоди. Такий тип інформації є небезпечним, оскільки він може бути представлений так, що важко відрізнити від правдивої інформації. Дослідження ідентифікували кілька типів фейкових новин, включаючи навмисно створені, сатиричні та помилкові новини через недбалу перевірку фактів. Розуміння цих типів допомагає у взаємодії з аудиторією та виявленні та контролі фейкових новин.

У дослідженнях фейкових новин [1] велика увага приділяється методам виявлення. З поширенням цифрових технологій і глибокого навчання з'явилися нові інструменти аналізу текстів, які можуть виявляти дезінформацію, аналізуючи лінгвістичні патерни та інші лінгвістичні особливості. Такі системи вже використовуються медіа-платформами для боротьби з дезінформацією.

Важливим аспектом боротьби з фейковими новинами є розробка медіа-стандартів і правил. Дослідження вказують на необхідність чітких стандартів перевірки та подання інформації для ускладнення поширення фейкових новин.

Залучення громадськості та освітні програми для підвищення обізнаності також мають вирішальне значення [2].

Психологічний вплив фейкових новин на окремих осіб і групи є ще однією важливою темою досліджень. Неправдива інформація часто призводить до емоційних реакцій, які сприяють поширенню дезінформації. Підвищення медіаграмотності населення та розуміння ролі соціальних мереж у поширенні фейкових новин є важливими аспектами в боротьбі з цим явищем.

Фейкові новини стали важливою проблемою у медійному світі, оскільки вони можуть суттєво впливати на суспільство, політику та особисте життя людей. Дослідники вивчають це явище з різних точок зору, зокрема, його вплив, методи виявлення та способи протидії.

У наукових дослідженнях [3] фейкові новини часто описують як неправдиву або маніпулятивну інформацію, що може бути створена з метою впливу на громадську думку або для отримання фінансової вигоди. Цей вид інформації може бути особливо небезпечним, оскільки його можна легко сприйняти як правдиву. Дослідження ідентифікували кілька типів фейкових новин, таких як намірено створені, сатиричні або помилкові, і розглядають їхні наслідки та способи боротьби з ними.

Одним із напрямків досліджень є розробка методів виявлення фейкових новин. За допомогою цифрових технологій та глибокого навчання розроблено інструменти, які можуть аналізувати текст та виявляти дезінформацію. Деякі медіа-платформи вже використовують ці інструменти для боротьби з фейками.

Також важливою є розробка медіа-стандартів і правил, які ускладнюють поширення фейкових новин. Освітні програми та залучення громадськості для підвищення обізнаності є ключовими аспектами в цьому процесі.

Фейкові новини можуть мати психологічний вплив на людей, спонукаючи їх до емоційних реакцій та впливаючи на їхнє мислення. Підвищення медіаграмотності та розуміння ролі соціальних мереж у поширенні фейкових новин можуть допомогти у боротьбі з цим явищем.

Ще одним важливим аспектом у боротьбі з фейковими новинами є розуміння їхнього впливу на суспільство. Деякі дослідження показують, що фейкові новини можуть поглиблювати соціальні розбіжності та сприяти політичній нестабільності. Наприклад, вони можуть посилювати відчуття відчуження та недовіри до інших груп людей або до владних структур.

Крім того, дослідження показують, що фейкові новини можуть мати серйозні наслідки для демократії. Вони можуть впливати на результати виборів, спотворюючи громадську думку та вплив на політичні рішення. Також фейкові новини можуть підірвати довіру громадськості до засобів масової інформації та інших джерел інформації.

Отже, для ефективної боротьби з фейковими новинами потрібен комплексний підхід, який враховує їхні соціальні, політичні та економічні наслідки. Розвиток медіаграмотності, використання новітніх технологій для виявлення дезінформації та впровадження правових норм є ключовими елементами цього підходу.

Дослідження показують, що фейкові новини можуть мати серйозні наслідки для суспільства. Вони можуть сприяти появі соціальних конфліктів та поглиблювати розбіжності між різними групами людей. Крім того, фейкові новини можуть впливати на політичну ситуацію у країні, спотворюючи громадську думку та вплив на результати виборів [4].

Одним з методів боротьби з фейковими новинами є підвищення медіаграмотності серед населення. Людям потрібно навчати розрізняти достовірну інформацію від недостовірної, перевіряти факти та джерела інформації перед тим, як ділитися нею. Також важливою є робота із медіа-платформами, щоб вони розробляли та використовували алгоритми, які допомагали б виявляти та обмежувати поширення фейкових новин.

З огляду на широке коло проблем, пов'язаних із фейковими новинами, необхідне також законодавче регулювання.

Зокрема, серед можливостей - суворіше регулювання контенту платформ, посилення відповідальності за поширення неправдивої інформації та сприяння транскордонному обміну досвідом у боротьбі з цією проблемою.

Розуміння та ефективного реагування на проблему фейкових новин вимагає комплексного підходу, що поєднує дослідження, технологічний розвиток, правові рамки та освітні ініціативи. Реалізація цих стратегій має важливе значення для захисту демократії та збереження суспільної довіри.

Загалом, проблема фейкових новин потребує комплексного підходу, який би об'єднував зусилля уряду, медіа-платформ та громадськості. Тільки спільні зусилля можуть допомогти ефективно боротися з цим явищем та зберегти довіру до інформації, яку ми споживаємо.

Узагальнюючи, розуміння та ефективна боротьба з фейковими новинами вимагає комплексного підходу, який об'єднує наукові дослідження, технологічний прогрес, правові норми та освітні ініціативи.

1.2 Методи виявлення фейкових новин

Методи виявлення фейкових новин активно досліджуються, оскільки вони відіграють важливу роль у підтримці довіри до інформації в сучасному інформаційному просторі. Розвиток технологій штучного інтелекту та обробки природної мови відкриває нові можливості для автоматизації виявлення та фільтрації фейкових матеріалів.

Один з найефективніших методів полягає в застосуванні глибокого навчання та обробки природної мови для аналізу текстового контенту.

Системи глибокого навчання навчаються розпізнавати патерни, які зазвичай асоціюються з фейковими новинами, такі як маніпулятивний стиль мови та необґрунтовані твердження. Ці моделі стають більш точними з часом, оскільки вони навчаються на великих обсягах даних, що містять перевірені справжні та фейкові новини [5].

Інший підхід [6] полягає у створенні баз даних, що містять інформацію про вже виявлені фейкові новини та їх джерела. Ці бази даних можуть використовуватись для порівняння нового контенту з наявними записами дезінформації. Якщо новина має багато спільних характеристик із записом у базі даних, вона позначається як підозріла і потребує подальшої перевірки.

Додатково, методи використання фактчекінгу, аналізу мережі джерел, а також використання штучного інтелекту для аналізу великих обсягів даних допомагають виявляти фейкові новини. Комбінація цих методів дозволяє ефективно виявляти та запобігати поширенню дезінформації в інформаційному просторі.

Ще одним ефективним методом виявлення фейкових новин є використання перехресної перевірки та міждисциплінарного аналізу даних. Наприклад, аналітичні платформи можуть інтегрувати дані з різних джерел та медіа-платформ і використовувати складні алгоритми для виявлення невідповідностей в інформації. Це може включати перехресну перевірку фактів, порівняння новинних повідомлень з офіційними даними урядів або статистичними управліннями, а також звіряння з базами даних наукових публікацій [7].

Також важливим є розвиток програмного забезпечення для виявлення лінгвістичних відмінностей, що характеризують фейкові новини. Дослідження показують, що фейкові новини часто містять емоційно забарвлену лексику та викликають сильні емоції, на відміну від традиційних журналістських матеріалів. Комп'ютерні програми для аналізу тексту можуть виявити ці відмінності, аналізуючи частоту використання певних слів та фраз, що часто асоціюються з маніпуляціями.

Для боротьби з фальсифікацією візуальних матеріалів використовують методи візуальної аналітики. Спеціалізоване програмне забезпечення може допомогти виявити підроблені зображення та відео, порівнюючи їх з базами даних невідредагованих матеріалів. Також інструменти для аналізу геолокації можуть перевірити правдивість місця події, описаної у новинах.

Усі ці методи дозволяють забезпечити більшу точність та достовірність інформації в цифрову епоху та підвищити довіру громадськості до медіа-контенту.

Ще одним підходом є використання технології блокчейн для забезпечення прозорості та надійності інформаційних потоків. Блокчейн може створювати незмінні записи про джерело інформації та її автентичність, запобігаючи маніпуляціям та втручанню у зміст інформації. Це забезпечує публічний доступ до інформації з джерел новин та її перевірку, зміцнюючи довіру до медіа-контенту [8].

Також ефективним виявився метод використання мультимодального аналізу, який об'єднує текстові, аудіо та візуальні дані для комплексного розуміння інформації. Цей підхід включає аналіз зображень та відео, які супроводжують текст новин, на предмет маніпуляцій або аномалій, що можуть вказувати на фальсифікацію. Наприклад, інструменти для виявлення фотошопу в зображеннях або незвичайних патернів редагування у відео можуть допомогти ідентифікувати маніпульовані матеріали, що є частою ознакою фейкових новин.

Іншим підходом є розвиток методів верифікації контенту, що базуються на семантичному аналізі. Семантичний аналіз дозволяє системам розпізнавати не лише окремі слова, а й загальний контекст, у якому ці слова використовуються. Такий підхід може виявляти суперечності або аномалії в тексті, які часто сигналізують про фейкові новини. Наприклад, використання технологій глибокого навчання та моделей глибокого навчання, таких як BERT (Bidirectional Encoder Representations from Transformers), дозволяє аналізувати текст на більш глибокому рівні, порівнюючи його з великими базами даних достовірних джерел для визначення його правдоподібності [9].

Ці методи дозволяють ефективно ідентифікувати та протидіяти поширенню фейкових новин, забезпечуючи точність і достовірність інформаційного контенту в цифрову епоху.

Ще одним способом виявлення фейкових новин є аналіз соціальних мереж та інтернет-форумів на предмет поширення дезінформації. Алгоритми можуть виявляти патерни, які вказують на організовані спроби вплинути на громадську думку через маніпуляцію інформацією. Наприклад, виявлення великої кількості акаунтів, які одночасно публікують або підтримують певну інформацію, може свідчити про штучно створений та спрямований на маніпуляцію контент [10].

Додатково, використання інструментів із сфери когнітивних наук та психології може допомогти виявити фейкові новини. Наприклад, деякі дослідження показали, що люди мають тенденцію довіряти інформації, яка підтримує їхні власні переконання, навіть якщо ця інформація є неправдивою. Врахування цієї особливості поведінки може допомогти розробити алгоритми, які виявляють та роблять більш вірогідними фейкові новини, що відповідають індивідуальним схильностям користувачів.

Також важливим є розвиток методів для виявлення фейкових новин у медіа-контенті, які використовують штучний інтелект та аналітичні платформи. Ці методи можуть виявляти незвичні або недостовірні патерни у відео-, аудіо- та текстовому контенті, що вказують на можливу фальсифікацію.

Використання перерахованих вище методів та інтеграція їх у комплексні системи дає змогу ефективно ідентифікувати та протидіяти поширенню фейкових новин, забезпечуючи точність і достовірність інформаційного контенту в цифрову епоху.

1.3 Виклики та проблеми

Дослідження феномену фейкових новин приводить до аналізу численних проблем, які вони створюють для суспільства, медійної індустрії та інформаційної безпеки. Однією з основних проблем є швидкість поширення фейкових новин, яка перевищує темпи розповсюдження перевіреної інформації. Це пояснюється тим, що фейкові новини часто мають сенсаційний

характер та спричиняють сильні емоційні реакції, що сприяє їхньому вірусному поширенню в соціальних мережах.

Ще одна серйозна проблема полягає у впливі фейкових новин на громадську думку та поведінку. Вони можуть не тільки вводити людей в оману, але й призводити до серйозних соціальних та політичних наслідків, включаючи непорозуміння, конфлікти, втрату довіри до медійних інститутів, а також маніпуляцію виборчими процесами.

Ефективність методів виявлення та боротьби з фейковими новинами також є значною проблемою. Традиційні підходи до фактчекінгу можуть бути неефективними або надто повільними, оскільки процес перевірки може бути трудомістким і не завжди встигати за темпами поширення дезінформації. Застосування технологій глибокого навчання та штучного інтелекту може пропонувати потенційні рішення, але також порушує питання точності, передові судження та можливості зловживань [11].

Нові техніки, такі як глибокі підробки (deepfakes), ускладнюють ідентифікацію фейкових новин. Забезпечення приватності та свободи слова, міжнародна співпраця та стандартизація, а також балансування між заходами боротьби з дезінформацією та збереженням основних прав і свобод є важливими завданнями у боротьбі з фейками.

Додаткові виклики включають захист прав людини і свободу слова. Ініціативи з регулювання контенту можуть мати непередбачені наслідки, включаючи обмеження законних форм вираження та журналістської діяльності. Це ставить перед законодавцями та регуляторами завдання знайти баланс між боротьбою з дезінформацією та збереженням фундаментальних прав та свобод.

Один з ключових викликів у протидії фейковим новинам полягає в утрудненні ефективної співпраці між медійними організаціями, технологічними платформами та урядовими агенціями. Різні підходи та відсутність єдиної стратегії можуть призводити до нескоординованих заходів, що ускладнюють боротьбу з дезінформацією. Обмін ресурсами та даними між

різними сторонами може значно посилити зусилля з виявлення та нейтралізації фейкових новин, забезпечуючи широке охоплення та швидку реакцію. Також важливо визначити межі між фейковими новинами та контентом, що містить суб'єктивні оцінки або сатиру, розробивши чіткі критерії для їх розрізнення. Це завдання має як технічну, так і етичну складові, оскільки стосується основних принципів свободи слова та преси [12].

Ще одним серйозним викликом є навмисне використання фейкових новин як інструменту гібридної війни або геополітичної боротьби. Держави або організовані групи можуть використовувати дезінформацію для дестабілізації ситуації в інших країнах, підриваючи довіру до урядів та суспільних інститутів. Для виявлення та протидії такому використанню необхідна міжнародна правова та політична співпраця, а не лише технічні рішення.

Додатковим викликом є налагодження ефективного діалогу між науковцями та практиками у сфері медіа та технологій. Наукові дослідження можуть надати цінні інсайти, але без тісної взаємодії з медійниками та розробниками, ці знання можуть залишатись не використаними на практиці. Створення платформ для обміну знаннями та кращими практиками може допомогти перекласти теоретичні розробки в ефективні інструменти боротьби з фейковими новинами [13].

Ще один важливий виклик полягає в адаптації до нових тактик маніпуляції, які постійно еволюціонують. Новаторські методи, такі як використання штучного інтелекту для створення переконливих текстів, можуть ускладнювати ідентифікацію джерела інформації. Це вимагає не лише постійного оновлення програмного забезпечення для виявлення, але й глибшого розуміння лінгвістичних патернів, які глибоке навчання може інтерпретувати як "правдиві" або "фальшиві". Також важливо встановлювати правила приватності та захисту особистих даних під час аналізу інформації, зберігаючи баланс між необхідністю аналізу великих обсягів даних та правом на приватність [14].

У боротьбі з фейковими новинами, також важливо враховувати різницю у законодавстві різних країн щодо цього питання, що створює складнощі для міжнародних платформ. Навіть встановлення правил приватності та захисту особистих даних вимагає уважності й постійного оновлення відповідно до законодавства.

Науковці та технологи стикаються з важливим завданням забезпечення необхідного фінансування для проведення досліджень у галузі виявлення фейкових новин. Багато перспективних проектів потребують значних витрат на розробку та випробування нових технологій. Привертання інвестицій, отримання державних грантів або укладання партнерств з приватним сектором є ключовими для забезпечення стійкої діяльності у цій області. Додатковим викликом є підвищення рівня освіти та інформованості серед населення. Важливо, щоб люди могли критично оцінювати інформацію, з якою вони зіштовхуються щодня. Програми медіаграмотності, що навчають розрізняти надійні джерела від сумнівних, можуть істотно зменшити вплив фейкових новин.

В контексті глобальної інформаційної екосистеми, виклик становить також неоднорідність законодавства різних країн стосовно фейкових новин. Що в одній країні вважається незаконним, може бути допустимим в іншій. Це створює юридичні складнощі для міжнародних платформ, таких як Facebook або X(Twitter), які намагаються модерувати контент, не порушуючи місцевих законів [15].

Науковці та технологи стикаються з важливим завданням забезпечення необхідного фінансування для проведення досліджень у галузі виявлення фейкових новин. Багато перспективних проектів потребують значних витрат на розробку та випробування нових технологій. Привертання інвестицій, отримання державних грантів або укладання партнерств з приватним сектором є ключовими для забезпечення стійкої діяльності у цій області. Додатковим викликом є підвищення рівня освіти та інформованості серед населення. Важливо, щоб люди могли критично оцінювати інформацію, з якою вони

зіштовхуються щодня. Програми медіаграмотності, що навчають розрізняти надійні джерела від сумнівних, можуть істотно зменшити вплив фейкових новин.

Таблиця 1.1.

Огляд основних викликів у боротьбі з фейковими новинами

Виклик	Опис виклику
Швидкість поширення фейкових новин	Фейкові новини швидше поширюються через соціальні мережі, ніж перевірена інформація.
Вплив на громадську думку та поведінку	Фейкові новини можуть вводити в оману та призводити до соціальних та політичних наслідків.
Ефективність методів виявлення та боротьби	Потреба у швидких та ефективних методах виявлення та запобігання поширенню фейкових новин.
Розробка технологій глибокого навчання	Використання технологій глибокого навчання та штучного інтелекту для боротьби з фейками.
Захист прав людини та свобода слова	Потрібно зберігати баланс між боротьбою з дезінформацією та збереженням свободи слова та преси.
Ефективна колаборація між сторонами	Потреба у спільній роботі між медійними організаціями, технологічними платформами та урядовими агенціями.
Визначення межі між фейками та суб'єктивністю	Розробка чітких критеріїв для відрізнєння дезінформації від допустимих форм виразності.
Геополітичне використання фейків	Фейкові новини використовуються для гібридної війни та геополітичної боротьби.
Фінансування досліджень у сфері виявлення	Необхідно забезпечити достатнє фінансування для розробки та тестування нових технологій.

Підвищення рівня обізнаності	Важливо підвищувати обізнаність серед населення щодо розрізнення фейкових та достовірних новин.
------------------------------	---

В таблиці 1.1 представлено систематизацію ключових викликів, з якими стикаються фахівці під час боротьби з фейковими новинами. Ці виклики включають технічні аспекти, такі як швидкість поширення фейкових новин і розробка методів їх виявлення, як вплив на громадську думку та необхідність забезпечення правильного балансу між свободою слова та необхідністю регуляції контенту. Важливість цих викликів підкреслюється не тільки їхньою актуальністю, а й складністю вирішення, що вимагає комплексного підходу та співпраці між різними зацікавленими сторонами, включаючи державні органи, приватний сектор, наукові круги та громадськість.

Боротьба з фейковими новинами вимагає комплексного підходу, який залучає багато різних сфер — від технологій та законодавства до освіти та міжнародної співпраці. Тільки так можна досягнути ефективного прогресу у вирішенні цієї складної проблеми [16].

Висновки до розділу 1

Фейкові новини є великим викликом для нашого суспільства, оскільки вони можуть суттєво впливати на політичні процеси, соціальну стабільність та сприйняття реальності. Використання методів глибокого навчання та обробки природної мови показує потенціал у виявленні та фільтрації фейкових новин, проте це також підкреслює необхідність постійного удосконалення цих технологій для адаптації до постійно змінюючихся тактик маніпуляції.

Аналіз викликів та проблем, пов'язаних з фейковими новинами, показує, що однією з ключових проблем є швидке та широке їх поширення, що ускладнює контроль над розповсюдженням неправдивої інформації. Також важлива потреба у міждисциплінарному підході, який включає не лише технічні рішення, але й законодавчі, освітні та соціальні ініціативи для

ефективної боротьби з дезінформацією. Особлива увага має бути приділена регулюванню контенту на медіаплатформах з метою зменшення впливу фейкових новин без порушення прав на свободу слова.

Ураховуючи ці аспекти, можна зробити висновок, що боротьба з фейковими новинами потребує спільних зусиль урядів, медіаорганізацій, технологічних компаній та громадськості. Інтеграція передових технологій з освітніми програмами та чіткою правовою базою може стати ключем до ефективної стратегії боротьби з поширенням фейкових новин. Це дозволить не лише зменшити їхній негативний вплив на суспільство, але й зберегти довіру до інформаційних джерел та зміцнити демократію.

РОЗДІЛ 2. ТЕОРЕТИЧНІ ОСНОВИ ГЛИБОКОГО НАВЧАННЯ

2.1 Основи глибокого навчання

Теоретичні основи глибокого навчання формують фундамент, на якому будуються всі алгоритми та методики, що використовуються для аналізу та обробки даних у численних додатках. Глибоке навчання - це галузь машинного навчання, яка зосереджується на використанні нейронних мереж з багатьма шарами для розв'язання складних завдань. Основні концепції глибокого навчання включають у себе використання багат шарових нейронних мереж, здатність автоматично навчатися представленням даних та використання великих обсягів даних для тренування.

Це дозволяє розв'язувати складні завдання, такі як розпізнавання образів або мовлення, шляхом поетапного створення вищих рівнів абстракції з низькорівневих даних. Глибоке навчання використовує великі набори даних для тренування, а також потребує відповідної архітектури мережі та методів оптимізації.

Глибоке навчання базується на ідеї про те, що складні представлення даних можуть бути побудовані шляхом поступового відображення простіших функцій та атрибутів. Кожен шар нейронної мережі може відфільтрувати та перетворити вхідні дані в більш абстрактне представлення для наступного шару. Таким чином, мережа може автоматично вчитися властивостям даних та робити передбачення на їх основі [17].

У глибокому навчанні існує кілька підходів до навчання моделей. Наприклад, навчання з учителем використовує набір даних з відомими відповідями для тренування моделі. Навчання без учителя не потребує міток і дозволяє моделі виявляти внутрішні залежності у даних. Підкріплювальне навчання, у свою чергу, базується на нагороді за правильні дії моделі [18].

Глибоке навчання потребує великої кількості даних для тренування, оскільки багат шарові нейронні мережі мають велику кількість параметрів, які

потрібно налаштувати. Також важливою є відповідна архітектура мережі та методи оптимізації, такі як стохастичний градієнтний спуск або алгоритми адаптивного вивчення швидкості навчання.

Однією з головних переваг глибокого навчання є його універсальність. Техніки глибокого навчання можуть бути застосовані до широкого спектру задач, від класифікації об'єктів на зображеннях до розпізнавання мови та вирішення складних задач обробки природної мови. Це робить його потужним інструментом для вирішення різноманітних завдань у багатьох галузях, включаючи медицину, фінанси, технології безпеки та багато інших [19].

Однак, глибоке навчання має свої обмеження та виклики. Наприклад, для успішного тренування глибоких моделей потрібні великі обсяги даних та велика обчислювальна потужність, що може бути недосяжним для багатьох дослідників та організацій. Крім того, інтерпретація результатів глибокого навчання може бути складною, оскільки моделі можуть виробляти високорівневі представлення даних, які не завжди легко зрозуміти людям.

Однією з ключових ідей є використання багат шарових нейронних мереж, які автоматично вивчають корисні представлення даних з великих обсягів вхідних даних. Наприклад, у задачі розпізнавання зображень перші шари можуть виявляти низькорівневі функції, такі як різні краї та форми, а потім більш високі шари можуть комбінувати ці функції, щоб розпізнати складні об'єкти або сцени [20].

При розробці моделей важливо використовувати методи перевірки та валідації. Це допомагає переконатися, що модель працює ефективно на нових даних і не перенавчена. Для цього можна використовувати методи крос-валідації та розділення наборів даних на тренувальні, тестові та валідаційні для ефективною перевірки моделей.

Глибоке навчання використовується в різних сферах, включаючи комп'ютерно зорову, мовленнєве розпізнавання, обробку природної мови, генерацію тексту, музику та багато іншого. У сфері комп'ютерного зору,

наприклад, глибокі нейронні мережі дозволяють роботам та системам безпеки ефективно розпізнавати обличчя та об'єкти на відео з високою точністю [21]..

Глибоке навчання також широко використовується в обробці природної мови (NLP). Моделі глибокого навчання можуть аналізувати, розуміти та генерувати людську мову з вражаючою точністю. Вони застосовуються в автоматизованих системах перекладу, генерації тексту, аналізі настроїв у соціальних мережах, розпізнаванні мовлення та інших сферах.

Глибоке навчання відіграє важливу роль у виявленні та захисті від фейкових новин у великих обсягах текстової інформації. Одним з основних методів виявлення фейків є аналіз текстів за допомогою нейронних мереж. Такі мережі можуть бути навчені розрізняти між правдивою та фальшивою інформацією на основі великої кількості прикладів.

Для цього використовуються такі техніки, як аналіз тональності тексту, виявлення аномалій у структурі речень та використання спеціалізованих алгоритмів для виявлення шаблонів, характерних для фейкових новин. Глибокі нейронні мережі можуть автоматично виявляти такі шаблони, що дозволяє виявляти фейкові новини з високою точністю.

Також глибоке навчання може бути використане для створення систем, які автоматично перевіряють достовірність інформації перед її публікацією. Це може включати в себе перевірку фактів, перевірку джерел та аналіз текстів на предмет суперечливостей та невідповідностей.

Такі системи можуть бути корисними для медійних компаній, соціальних мереж та інших платформ, де публікується велика кількість інформації, що потребує перевірки на достовірність. Вони можуть допомогти зменшити поширення фейкових новин та покращити якість інформації, що доступна користувачам.

Глибоке навчання також використовується для створення моделей, які можуть виявляти фейкові новини на основі контексту та семантики текстів. Це означає, що модель може аналізувати не лише окремі слова чи речення, але й їхній зв'язок та значення в контексті. Наприклад, модель може виявити

протиріччя в інформації або використання некоректних доказів у тексті, що може свідчити про фейк [22]..

Крім того, глибоке навчання може бути використане для покращення систем виявлення фейків через постійне навчання на нових даних. Моделі можуть автоматично адаптуватися до нових методів маніпуляції та еволюціювати разом із змінами у способах поширення фейкових новин.

Застосування глибокого навчання у виявленні та захисті від фейків є важливим кроком у боротьбі з дезінформацією. Воно дозволяє автоматизувати процес виявлення фейків і зробити його більш ефективним та точним. Однак важливо розуміти, що це лише один з інструментів у загальній стратегії боротьби з фейками, і потребує постійного вдосконалення та адаптації до нових викликів і загроз.

Глибоке навчання є потужним інструментом у боротьбі з фейковими новинами через його здатність аналізувати текстову інформацію на більш глибокому рівні, враховуючи контекст та семантику. Моделі, побудовані на основі глибокого навчання, можуть виявляти протиріччя, аномалії та інші ознаки фейкових новин, що дозволяє ефективніше фільтрувати та ідентифікувати дезінформацію. Постійне навчання моделей на нових даних дозволяє адаптуватися до нових тактик маніпуляції, що робить системи виявлення фейків більш надійними та ефективними [23].

Проте важливо розуміти, що глибоке навчання є лише однією з складових стратегії боротьби з фейковими новинами. Для успішної боротьби потрібно також враховувати інші аспекти, такі як розвиток медіаграмотності серед громадян, підвищення етичних стандартів у медіа та розробка правової бази для регулювання дезінформації. Тільки комплексний підхід може забезпечити ефективний захист від фейкових новин і збереження довіри до інформаційних джерел у сучасному інформаційному суспільстві.

Таблиця 2.1

Основні аспекти та застосування глибокого навчання

Концепт	Опис	Застосування в галузях
Кореляція та коваріація	Аналіз взаємозалежностей між характеристиками для вибору найбільш релевантних для моделей.	Поліпшення точності прогнозувань у фінансовій аналітиці, соціальних науках.
Редукція розмірності даних	Зменшення обсягу даних, зберігаючи важливу інформацію, для ефективнішого навчання моделей.	Ефективне управління даними в зображеннях, текстових даних.
Адаптація до дрейфу даних	Методи адаптації моделей до змін у даних, забезпечуючи точність у динамічних умовах.	Фінансові ринки, інтернет речей, автоматизація виробництв.
Застосування в різних галузях	Розвиток технологій глибокого навчання для специфічних галузевих рішень.	Охорона здоров'я, автомобільна промисловість, розумні міста, та інші.

У таблиці 2.1, яка представлена у рамках розвитку систем глибокого навчання, розкриваються ключові поняття, що сприяють підвищенню ефективності та адаптивності цих технологій. Застосування статистичних методів, таких як аналіз кореляції та коваріації, дозволяє виявити важливі характеристики для тренування алгоритмів. Техніки, наприклад, редукція розмірності даних, відіграють важливу роль у обробці великих обсягів інформації без втрати ключових деталей.

Адаптація моделей до змін у даних, відома як дрейф даних, є ключовою для забезпечення стабільності та надійності систем у динамічних середовищах. Розширення сфер застосування глибокого навчання підсилює його важливість у сучасному світі, відкриваючи нові можливості для інновацій

у таких критичних галузях, як медицина, автомобілебудування та міське планування [24].

Розширення можливостей глибокого навчання та його застосування в різних сферах діяльності не тільки покращує існуючі технології, але й відкриває двері до інноваційних рішень, які змінюють уявлення про можливості автоматизації та інтелектуальної обробки даних.

2.2 Моделі для класифікації тексту

Моделі для класифікації тексту є важливим інструментом у глибокому навчанні для обробки та розуміння великих обсягів текстових даних. Вони дозволяють автоматично визначати тематичну категорію тексту, що має велике значення в різних сферах, таких як обробка природної мови, соціальні медіа, веб-моніторинг та інші. Класифікація тексту знаходить застосування в сортуванні електронних листів на спам та неспам, аналізі відгуків користувачів, визначенні настрою у текстах, автоматичному складанні резюме документів та багатьох інших завданнях.

Розробка моделей для класифікації тексту зазвичай розпочинається з попередньої обробки даних. Цей процес включає в себе кілька етапів, таких як токенізація тексту, видалення стоп-слів, стемінг та лематизація. Ці кроки допомагають спростити текст, перетворюючи його на формат, який легше аналізувати машинами. Наступним етапом є перетворення тексту в числовий формат за допомогою технік векторизації, таких як TF-IDF або використання векторів слів. Це дозволяє моделям глибокого навчання ефективно працювати з текстовими даними, виявляти закономірності та відносини між словами та фразами [25].

Одним із найпоширеніших методів класифікації тексту є використання наївного Баєсового класифікатора, який використовує принципи баєсової статистики для прогнозування категорії тексту. Інші поширені методи включають логістичну регресію, машини опорних векторів (SVM) та нейронні

мережі. Сучасні підходи, такі як глибоке навчання, використовують складні архітектури, такі як свердловинні нейронні мережі (CNN) або рекурентні нейронні мережі (RNN), для більш глибокого аналізу тексту. Вони враховують контекст та семантичне значення слів, що дозволяє отримати більш точні результати при класифікації текстів[26].

Ще одним важливим проривом у класифікації тексту є застосування трансформерних моделей, таких як BERT (Bidirectional Encoder Representations from Transformers). Ці моделі відзначаються високою точністю у багатьох завданнях обробки природної мови. Однією з ключових особливостей трансформерних моделей є використання механізмів уваги, що дозволяють моделі уважно аналізувати текст і зосереджуватися на найважливіших його частинах для кращого розуміння. Це робить їх дуже ефективними для різних завдань, від класифікації тексту до генерації нового контенту.

Один з ключових шляхів розвитку технологій класифікації тексту полягає в застосуванні ансамблевих методів, які включають в себе такі техніки, як бустинг, беггінг та стекінг. Ці методи поєднують передбачення кількох моделей для створення більш стабільної та надійної моделі. Наприклад, метод Random Forest використовує ансамбль дерев рішень для класифікації тексту, де кожне дерево тренується на випадковій підвибірці даних і функцій. Це суттєво підвищує точність моделі у порівнянні з одиночним деревом рішень [27].

За останні роки застосування нейронних мереж значно підвищило ефективність класифікації тексту. Особливо відзначаються моделі з довготривалою пам'яттю (LSTM) та GRU (Gated Recurrent Units), які відмінно справляються з аналізом послідовностей даних, що робить їх ідеальними для обробки текстів, де контекст та порядок слів мають значення. Ці моделі є різновидами рекурентних нейронних мереж, які здатні запам'ятовувати інформацію на довгі періоди, що дозволяє їм краще розуміти зміст тексту і виявляти сентимент.

Останнім часом значну увагу здобувають технології, що базуються на механізмі уваги та трансформерах. Ці технології дозволяють моделям зосередитися на ключових частинах тексту при його обробці, що дозволяє досягти високої точності у завданнях, які потребують глибокого розуміння мовних структур [28].

Трансформери, особливо з двонаправленою структурою, як у BERT, демонструють високу ефективність у класифікації тексту, розпізнаванні іменованих сутностей, відповіді на питання та інших завданнях NLP. Важливою складовою успіху в класифікації тексту є якість навчальних даних. Забезпечення чистих, релевантних та збалансованих даних може значно підвищити результати навчання моделей [29].

Активне використання технік аугментації даних, таких як перефразування, синонімізація або зміна порядку слів у реченнях, може допомогти моделям краще загальнізувати і адаптуватися до нових даних, що вони зустрічають у реальних умовах.

Класифікація тексту вимагає глибокого розуміння того, як машини можуть вивчати та інтерпретувати людську мову. Розвиток семантичного аналізу, який включається до процесу класифікації тексту, відкриває нові можливості для підвищення точності моделей. Цей аналіз допомагає машинам не лише розпізнавати слова, а й розуміти їхнє значення у конкретному контексті, що істотно підвищує їх здатність до точної класифікації текстів за темами чи настроями [30].

У семантичному аналізі тексту важливою є розробка вбудовувань слова, які перетворюють слова у вектори з високою розмірністю. Ці вектори відображають семантичні та синтаксичні зв'язки між словами. Використання вбудовувань слова дозволяє моделям аналізувати текст глибше, виявляючи залежності між словами у різних контекстах. Техніки, такі як Word2Vec або GloVe, створюють ці векторні представлення, які стали основою для багатьох сучасних моделей обробки природної мови.

Останнім часом значний прогрес у класифікації тексту був досягнутий завдяки вдосконаленню алгоритмів глибокого навчання, які здатні використовувати контекстну інформацію. Моделі, що базуються на архітектурі Transformer, такі як GPT та BERT, використовують контекстну інформацію для генерації тексту та його аналізу. Вони дозволяють здійснювати більш глибоке та точне розуміння тексту, розрізняючи не лише окремі слова чи фрази, але й цілісний зміст запропонованого тексту. Застосування цих технологій виявилось революційним для багатьох областей, де потрібно автоматизувати обробку та аналіз великих обсягів тексту, таких як юридичні документи, медичні записи, наукові статті та новинні матеріали. Це дозволило значно підвищити ефективність обробки інформації, роблячи можливим швидке виявлення релевантної інформації та класифікацію даних на більш високому рівні точності та релевантності [31].

У сфері класифікації тексту значний прогрес був досягнутий завдяки застосуванню гібридних моделей, які поєднують різні види нейронних мереж для поліпшення розуміння та класифікації тексту на більш глибокому рівні. Ці гібридні підходи часто використовують комбінацію конволюційних нейронних мереж (CNN), що добре розпізнають шаблони у тексті, та рекурентних нейронних мереж (RNN), які ефективно обробляють послідовні дані, зберігаючи інформацію про попередні елементи послідовності. Ці гібридні моделі дозволяють використовувати переваги кожного типу мережі, що забезпечує більш точне визначення контексту та змісту тексту. Такі комбіновані підходи можуть бути особливо корисними у завданнях, де необхідно класифікувати текст за емоційним забарвленням або тематичною спрямованістю, забезпечуючи глибше розуміння мови та точнішу реакцію на зміни в контексті [32].

Ще одним важливим аспектом є застосування мета-алгоритмів у процесі класифікації тексту. Мета-алгоритми, такі як методи ансамблю, можуть значно покращити продуктивність окремих моделей шляхом інтеграції різних стратегій навчання та предикції. Такий підхід може включати створення

комбінованих прогнозів з декількох моделей для досягнення більш високої стійкості та надійності у результатах [33].

Автоматичне розпізнавання мови та її діалектів стає все більш важливою областю застосування класифікації тексту в контексті глобалізації та розширення міжнародних комунікацій. Це полягає у розумінні та автоматичній ідентифікації різних мов та їхніх варіантів, що може значно покращити точність класифікаційних систем. Особливо це важливо в застосуваннях, де потрібна мультилінгвальність, наприклад, у сферах автоматичного перекладу, міжкультурної аналітики та інших.

Ці методи та технології, разом із передовими підходами до обробки та аналізу тексту, надають потужні інструменти для досліджень та розвитку в галузі глибокого навчання. Вони відкривають нові можливості для розуміння та використання мовних даних у різноманітних застосуваннях [34].

Класифікація тексту відкриває широкі можливості для розвитку в глибокому навчанні. З впровадженням нових технологій і алгоритмів ця галузь стає все більш динамічною та різноманітною. Вона надає потужні інструменти для обробки та аналізу текстових даних, що відкриває нові можливості для автоматизації та інтелектуалізації різних процесів у сучасному інформаційному суспільстві [35].

2.3 Метрики оцінки моделей

Використання відповідних метрик дозволяє оцінити ефективність моделі та визначити області для подальшого вдосконалення. Метрики оцінки у глибокому навчанні забезпечують об'єктивне кількісне уявлення про те, наскільки добре модель виконує свої завдання, і стають важливим інструментом для підтвердження наукових досліджень і комерційних застосувань.

Однією з ключових метрик для оцінки ефективності моделей у класифікації є точність (Accuracy). Ця метрика визначає відсоток правильно

класифікованих прикладів у відношенні до загальної кількості прикладів у тестовому наборі даних. Точність є простою та зрозумілою метрикою, але може бути обманливою в ситуаціях, коли кількість прикладів у різних класах нерівномірна. Для більш об'єктивної оцінки моделі часто використовуються додаткові метрики, такі як точність (Precision) та відновлення (Recall). Precision вказує на те, яка частина об'єктів, визначених моделлю як належні до певного класу, дійсно належить до цього класу. У той час як Recall показує, яка частина об'єктів, що дійсно належать до класу, була визначена моделлю як така. Ці метрики допомагають уникнути підвищеної чутливості до дисбалансу класів у наборі даних, що дозволяє глибше проаналізувати ефективність моделі, особливо у ситуаціях, коли важливо виявити певні класи, навіть за рахунок певного зростання помилок в інших класах [36].

Комбінування точності та відновлення в одну метрику, відому як F1-оцінка, дозволяє збалансувати обидва аспекти оцінки класифікації. Ця оцінка особливо корисна у випадках, коли точність або відновлення мають значно відмінні значення. F1-оцінка є гармонійним середнім між точністю та відновленням, надаючи важливу інформацію про збалансованість моделі. Для бінарної класифікації також використовуються ROC-крива (Receiver Operating Characteristic) та AUC (Area Under Curve). ROC-крива аналізує взаємозв'язок між чутливістю та специфічністю моделі при різних порогах рішень. AUC вимірює площу під ROC-кривою і дає загальну міру ефективності моделі: чим більше значення AUC, тим краще модель розрізняє між класами. Ці метрики є важливими для повноцінної оцінки здатності моделі до правильної класифікації та надають інформацію про її роботу в різних умовах та порогах відсічення [37].

Метрика логарифмічної втрати (Log-Loss) є однією з ключових у вимірюванні ефективності класифікаторів, оскільки вона враховує не лише правильність або неправильність прогнозів, а й ступінь впевненості моделі у своїх передбаченнях. Вона особливо корисна у ситуаціях, коли важливо уникнути впевнених, але невірних прогнозів. Ця метрика вимірює, наскільки

добре передбачені ймовірності класів моделлю відповідають дійсності, і дозволяє оцінити, наскільки впевнено модель може робити прогнози.

Ще однією ключовою метрикою у вимірюванні ефективності класифікації є коефіцієнт кореляції Меттьюса (МСС), який особливо важливий для оцінки бінарних класифікацій, особливо у випадках з незбалансованими класами. МСС забезпечує більш інформативну оцінку, оскільки враховує істинно позитивні, істинно негативні, хибно позитивні і хибно негативні результати в одному числовому значенні, що варіюється від -1 до +1. Це робить МСС більш стійкою та надійною метрикою, оскільки вона оцінює всі аспекти класифікації. У деяких випадках, особливо у вибірковому навчанні, використовується середня точність (Average Precision, AP). Ця метрика оцінює класифікатор на різних рівнях відновлення та вимірює середню точність, яку модель досягає при різних порогах класифікації. AP дозволяє докладніше проаналізувати процес прийняття рішень моделлю та з'ясувати, на яких порогах класифікації вона найефективніша та найточніша.

Інформаційний коефіцієнт, або коефіцієнт взаємної інформації, є метрикою, яка вимірює ступінь інформації, що передається між прогнозованими та фактичними класами у моделі класифікації. Ця метрика особливо корисна для виявлення непрямих зв'язків між змінними, що можуть впливати на ефективність моделі. Наприклад, вона може допомогти виявити, які ознаки або комбінації ознак є найбільш інформативними для правильної класифікації, що дозволяє вдосконалити модель та підвищити її точність [38].

Ще одна важлива метрика, яка часто використовується в дослідженнях, де потрібно зрозуміти варіативність відповідей моделі, – це коефіцієнт детермінації, відомий як R-квадрат. Ця метрика допомагає оцінити, яка частина змінності відгуків може бути пояснена за допомогою незалежних змінних у моделі. В контексті глибокого навчання, де моделі часто стикаються з великою кількістю змінних та потенційною взаємодією між ними, R-квадрат може дати цінне уявлення про те, наскільки ефективно модель використовує доступну інформацію для прогнозування. У деяких ситуаціях також може бути

корисною метрика як Гейн або Lift Chart, що допомагає оцінити, наскільки модель ефективно відрізняє між різними класами порівняно з випадковим відгадуванням. Ця метрика часто застосовується в маркетингових застосуваннях для оцінки ефективності кампаній, але може бути адаптована і для класифікації тексту, де важливо визначити, наскільки вдало модель здатна ідентифікувати ключові класи серед великої кількості даних [39].

Часто в моделях класифікації тексту спостерігається проблема перенавчання, коли модель бездоганно працює на тренувальному наборі даних, але виявляє велику похибку на нових, раніше не бачених даних. Це може вказувати на те, що модель надто сильно "вивчила" особливості тренувального набору, які можуть бути випадковими або не представляти загальної тенденції в даних. Така надмірна адаптація може призвести до втрати загальної узагальнюючої здатності моделі, тобто її здатності до ефективної роботи з різними, реальними даними, які вона раніше не бачила. Для запобігання перенавчанню можна використовувати різні стратегії, такі як використання більш простих моделей, які менше схильні до перенавчання, або використання методів регуляризації, які обмежують складність моделі.

Також важливо мати репрезентативний тренувальний набір даних, що відображає різноманіття можливих входних даних, на яких модель буде застосовуватися.

Аналізуючи ці аспекти, можна використовувати метрики, які вимірюють рівень помилок у різних частинах розподілу даних. Наприклад, вивчення величини помилок залежно від частоти класів може показати, чи модель частіше допускає помилки у рідкісних класах. Це може вказувати на необхідність збалансування набору даних або на необхідність коригування методів навчання моделі, щоб підвищити її рівномірність. Оцінка впливу аномалій у даних на поведінку моделі також є значущою. Моделі глибокого навчання часто реагують на "викиди" або аномальні значення, що може значно впливати на їх навчання. Розуміння того, як модель реагує на аномалії, і які

стратегії допомагають мінімізувати цей вплив, є важливим для покращення її універсальності та стійкості [40].

Таблиця 2.2

Ключові метрики для глибшої оцінки продуктивності моделей
глибокого навчання

Назва метрики	Опис метрики	Контекст застосування
Log-Loss	Вимірює точність класифікаторів, надаючи вагу за точність кожного прогнозу залежно від його ймовірності.	Корисна для оцінки моделей, де важливо мінімізувати впевнені помилкові прогнози.
Matthews Correlation Coefficient (MCC)	Вимірює загальну якість бінарних класифікацій, враховуючи всі чотири квадранти конфузійної матриці.	Ідеальна для незбалансованих датасетів, оскільки вона враховує істинні та хибні позитиви і негативи.
Average Precision (AP)	Вимірює середню точність по всім рівням порогу класифікації, яка допомагає зрозуміти ефективність моделі на різних порогах чутливості.	Використовується для детального аналізу якості відгуків моделі на різних рівнях порогу.
Sensitivity to Noise	Аналізує реакцію моделі на шум у даних, ідентифікуючи стійкість до зовнішніх та внутрішніх перешкод.	Корисна для визначення стабільності моделей у реальних сценаріях використання, особливо в промислових застосуваннях.

У цій таблиці представлені метрики, які дозволяють глибше проаналізувати продуктивність моделей глибокого навчання, зокрема їхню чутливість до помилок та аномалій. Використання цих метрик дозволяє

виявити слабкі місця в моделях, які можуть залишатися непоміченими при використанні традиційних метрик, таких як точність або відгук. Особливо важливим є розуміння того, як модель реагує на змінні умови використання та на рідкісні або виключні ситуації, що дозволяє адаптувати процес тренування для підвищення загальної ефективності та стійкості системи [41].

Цей аналіз може включати розробку спеціальних метрик, які можуть ідентифікувати чутливість до конкретних видів шуму або аномалій. За допомогою цих метрик можна точніше налаштовувати моделі, розробляючи механізми, що запобігають їх впливу на результати класифікації. Це, в свою чергу, допомагає створювати більш робустні та надійні системи, що є ключовим для ефективного використання глибокого навчання у різноманітних реальних сценаріях. Кожна з цих метрик вносить свій вклад у глибше розуміння характеристик та поведінки моделей глибокого навчання, забезпечуючи необхідні інструменти для наукового аналізу та комерційного застосування. Використання цих метрик дозволяє вченим та інженерам краще розуміти та оптимізувати свої моделі, забезпечуючи більш високу точність та ефективність у вирішенні завдань класифікації та прогнозування.

Оцінка моделей за допомогою цих метрик дозволяє не лише оцінити, наскільки ефективно модель вирішує свої основні завдання, а й зрозуміти, як різні параметри впливають на її продуктивність. Це може стати основою для покращення та налаштування моделей, а також для вибору оптимальних стратегій їх застосування в реальних умовах.

Висновки до розділу 2

Розуміння основ глибокого навчання є ключовим для роботи з автоматизованими системами аналізу даних. Це дозволяє ефективно використовувати техніки та алгоритми для розробки моделей, які можуть обробляти та класифікувати текстовий контент.

У цьому розділі детально розглядаються моделі для класифікації тексту, зокрема наївний Баєсів класифікатор, методи на основі дерев рішень та нейронні мережі. Проведений аналіз показує, як кожен з цих підходів може бути оптимізований для конкретних задач класифікації, забезпечуючи високу точність та відповідність вимогам реального часу та обмежень даних. Далі в розділі розглядається аналіз метрик оцінки моделей, який надає інструментарій для критичного оцінювання та порівняння ефективності різних моделей класифікації. Дослідження таких метрик, як точність, відгук, F1-метрика та інші, допомагає глибше зрозуміти стійкість та надійність системи у роботі зі складним текстовим контентом. Особливу цінність представляє використання цих метрик для виявлення потенційних вразливостей моделей, які можуть призвести до помилкових позитивів або негативів.

У цьому розділі не лише визначається теоретична база для подальших досліджень, але й підкреслюється значення ретельного вибору та налаштування моделей глибокого навчання для підвищення ефективності систем виявлення фейків. Вміння проводити точний та швидкий аналіз тексту є ключовим для розвитку надійних та ефективних інструментів, які можуть бути інтегровані у широкий спектр застосувань, від соціальних медіа до новинних платформ, де актуальність і достовірність інформації є критичними.

РОЗДІЛ 3. РОЗРОБКА ТА ВПРОВАДЖЕННЯ СИСТЕМИ ВИЯВЛЕННЯ ФЕЙКОВИХ НОВИН

3.1 Збір та підготовка даних

Збір та підготовка даних є критичними етапами у розробці системи виявлення фейкових новин. Успішність та точність будь-якої моделі глибокого навчання напряду залежать від якості та обсягу даних, на яких вона навчається. По-перше, важливо ідентифікувати та зібрати відповідні джерела даних, такі як новинні статті, блоги, повідомлення у соціальних мережах, а також коментарі та відгуки користувачів. Ключовою є те, щоб зібрані дані були репрезентативними та охоплювали різні точки зору та контексти, що допоможе системі ефективніше виявляти фейковий контент у різних сценаріях.

Після збору даних важливим етапом є їхнє очищення від шуму та нерелевантної інформації. Це включає видалення форматування, рекламних вставок та некоректних символів або посилань, які можуть спотворити результати моделі. Також важливо нормалізувати текст, перетворивши його до одного регістру, видаливши зайві пробіли та пунктуацію, а також розбивши текст на окремі слова або фрази. Після цього виконується токенізація, тобто перетворення тексту на структурований формат, з яким може працювати модель глибокого навчання. Наприклад, використання методів bag-of-words або TF-IDF дозволяє представити слова у вигляді числових значень, що сприяє ефективній роботі моделі з текстовими даними [42].

Ще одним важливим етапом у підготовці даних є розділення їх на тренувальну, тестову та валідаційну вибірки. Це дозволяє належним чином навчити модель та перевірити її здатність до узагальнення на нових, раніше не бачених даних. Ефективне розділення даних допомагає уникнути перенавчання моделі та забезпечити її надійність та об'єктивність у реальних умовах.

Під час створення та впровадження системи для виявлення фейкових новин, збір та підготовка даних мають вирішальне значення для забезпечення високої якості та ефективності майбутньої моделі. Окрім загальних методів очищення та нормалізації тексту, особлива увага приділяється анотації даних та визначенню прикладів, які можуть бути використані для тренування та тестування моделі. Для цього важливо розробити чіткі критерії, за якими новини класифікуються як 'правдиві' або 'фейкові', а також забезпечити, щоб набір даних був збалансованим щодо цих категорій [43].

Ефективна підготовка даних для системи виявлення фейкових новин включає співпрацю з експертами з медіа та фактчекінгу для перевірки та валідації даних. Це допомагає забезпечити достовірність та релевантність даних. Використання автоматизованих інструментів для аналізу змісту новин може прискорити процес виявлення потенційних фейків та забезпечити об'єктивність класифікації. Для більш детального аналізу тексту використовуються методи обробки природної мови, такі як розпізнавання іменованих сутностей, аналіз тональності та семантичний аналіз, які дозволяють виявляти тонкі нюанси та контекстуальні ознаки у текстах, що часто свідчать про фейковий контент. Застосування методів глибокого навчання до цих завдань дозволяє моделям вчитися розпізнавати складні патерни та аномалії у текстах, що значно підвищує точність виявлення фейкових новин.

Інтеграція різних джерел даних, таких як історичні дані та інформація з соціальних мереж, може значно поліпшити здатність системи виявлення фейкових новин адаптуватися до постійно змінюваних методів та стратегій, що використовуються для поширення неправдивої інформації. Це дозволяє системі забезпечувати більш глибокий аналіз контексту та виявляти нові, ще не відомі форми фейкових новин, що робить її більш ефективною в боротьбі з дезінформацією [44].

Під час збору та підготовки даних для системи виявлення фейкових новин, важливою стає робота з складними мовними конструкціями, які часто

використовуються для маніпуляції. Використання передових алгоритмів для автоматичного розпізнавання таких зворотів стає ключовим етапом. Технології глибокого навчання та нейронних мереж дозволяють ефективно обробляти великі обсяги тексту та виявляти складні зв'язки та патерни, які можуть залишатися непоміченими застарілими методами аналізу.

При підготовці системи для виявлення фейкових новин необхідно враховувати культурний та географічний контекст, в якому інформація генерується. Фейкові новини часто мають регіональні особливості, які впливають на їх стиль та зміст. Врахування цих аспектів дозволяє системі бути більш чутливою та специфічною у виявленні неправдивих повідомлень. Крім текстових даних, важливо включити в систему аналіз візуального контенту, аудіозаписів та відеоматеріалів. Використання комп'ютерного зору та технологій розпізнавання облич може виявити дезінформацію, яка може бути непомітною у текстовому форматі.

Подальше вдосконалення системи виявлення фейкових новин включає в себе використання адаптивних моделей, які можуть самостійно навчатися і адаптуватися до постійно змінюючихся умов інформаційного простору. Це означає розробку механізмів для швидкого оновлення даних і алгоритмів, а також включення користувачів у процес верифікації інформації. Це може сприяти підвищенню надійності системи та її здатності ефективно реагувати на нові методи та стратегії створення фейкових новин.

Для виявлення фейкових новин важливо використовувати методи глибокого навчання, які дозволяють автоматизувати аналіз емоційного забарвлення тексту. Аналіз настрою допомагає виявляти маніпулятивні повідомлення, що часто використовують емоційно заряджені висловлювання для впливу на аудиторію. Автоматизований аналіз тексту, заснований на визначенні ключових емоційних слів та фраз, дозволяє швидко оцінювати тон повідомлень та виділяти потенційно маніпулятивний контент.

Для більш глибокого розуміння контексту повідомлень система може включати компоненти, які аналізують зв'язки між різними джерелами новин та

їхню історичну достовірність. Це означає вивчення звичайних патернів поширення інформації серед певних джерел, виявлення аномалій у способах подачі новин, які можуть вказувати на спроби маніпуляції громадською думкою. Інтеграція технологій обробки природної мови (NLP) базується на алгоритмах, які можуть розпізнавати складні мовні конструкції та відтворювати семантичні структури тексту. Це дозволяє системі розпізнавати складні інформаційні кампанії, що використовують субтільні лінгвістичні техніки для впливу на сприйняття фактів [45].

Використання розширених баз даних, які містять перевірені факти, та алгоритмів порівняння фактичної інформації з цими базами дозволяє системі швидко ідентифікувати недостовірні дані, що є ключовим елементом в боротьбі з дезінформацією. Важливо також створити механізми оновлення та розширення баз даних, щоб система могла адаптуватися до постійно змінюваних методів поширення фейкових новин.

Таблиця 3.1

Компоненти та функції системи виявлення фейкових новин

Компонент	Функції	Роль в системі
Аналіз настрою	Визначення емоційного забарвлення тексту, виявлення маніпулятивних повідомлень.	Допомагає виявляти текст з надмірно емоційним забарвленням, який часто використовується у фейкових новинах.
Історична верифікація	Порівняння заяв з історично достовірними даними, використання баз даних перевірених фактів.	Забезпечує перевірку фактів та допомагає ідентифікувати недостовірну інформацію.
Контекстуальний аналіз	Розширений аналіз тексту для виявлення складних мовних патернів та зв'язків.	Підвищує точність системи за рахунок виявлення прихованих смислів і маніпуляцій.

Обробка природної мови	Розпізнавання мовних конструкцій, семантичний аналіз, визначення структур мовлення.	Поліпшує розуміння контенту, забезпечує глибший аналіз тексту для виявлення фейкового контенту.
Інтеграція різноманітних джерел	Збір інформації з множини платформ та форматів, включно з текстовими, візуальними та аудіо даними.	Дозволяє комплексний аналіз різних типів контенту для всебічного виявлення фейків.

Таблиця 3.1 ілюструє взаємодію різноманітних компонентів системи для виявлення фейкових новин з метою підвищення її ефективності. Кожен компонент виконує конкретні завдання, що сприяють загальній точності та надійності системи. Наприклад, аналіз настрою може допомогти виявити емоційно заряджений або маніпулятивний контент, в той час як історична перевірка та контекстуальний аналіз гарантують, що інформація є перевіреною та достовірною. Узгодженою роботою цих компонентів формується складна система, яка може адаптуватися до швидкозмінних викликів інформаційної ери та ефективно протистояти поширенню фейкових новин.

Розробка ефективної системи виявлення фейкових новин вимагає комплексного підходу, що об'єднує передові технології обробки даних із глибоким розумінням соціальних та культурних контекстів, які впливають на сприйняття інформації.

Підготовка даних для системи виявлення фейкових новин - це складний і важливий процес, який вимагає уваги до деталей та стратегічного планування. Головна мета цього етапу - забезпечити якість, валідність та репрезентативність даних, що використовуються для навчання та перевірки системи. Це включає в себе не лише збір даних з різних джерел, але і їхню обробку та перетворення в формат, зрозумілий для подальшого аналізу. Справжність та якість даних - ключові фактори у створенні ефективної моделі.

Точні та надійні дані дозволяють створювати модель, яка може адекватно реагувати на виклики сучасного інформаційного простору та виявляти фейкові новини. Тому, процес збору та підготовки даних є фундаментальним для успішної роботи системи виявлення фейкових новин у різних сценаріях та умовах [46].

3.2 Навчання та налаштування моделі

Процес навчання та налаштування моделі у системі виявлення фейкових новин є критичним для її успішної роботи. Від якості цього процесу залежить ефективність та точність всієї системи. Цей етап починається з вибору відповідної моделі глибокого навчання. Варіанти можуть бути різноманітними - від класичних методів статистики, таких як наївний Басс або логістична регресія, до складніших алгоритмів, таких як нейронні мережі або машини опорних векторів [47].

На етапі навчання моделі для виявлення фейкових новин важливо визначити параметри, що найкраще відображають закономірності у даних, щоб модель могла ефективно розрізняти правдиві новини від фейкових. Для цього потрібно впевнитися, що навчальний набір містить достатню кількість прикладів обох класів та є репрезентативним, охоплюючи різноманітні сценарії, які модель може зустріти у реальному світі. Далі необхідно налаштувати параметри моделі, такі як її складність, швидкість навчання та кількість ітерацій. Цей процес може включати багато експериментів і тестувань з використанням крос-валідації або валідаційних наборів даних для знаходження оптимального балансу між здатністю моделі узагальнювати дані та точністю на конкретних прикладах.

Один з ключових аспектів навчання моделі є уникнення перенавчання, коли вона занадто точно підлаштовується під навчальні дані і втрачає здатність до узагальнення на нові дані. Для цього можна використовувати методи, такі як регуляризація, випадкове відключення (dropout) для нейронних

мереж або додавання штрафних функцій для складності моделі. Ці методи допомагають зменшити ризик перенавчання та підвищити здатність моделі до роботи з реальними, часто непередбачуваними даними. Ще один важливий етап - це оцінка моделі на тестовому наборі даних, який не використовувався під час навчання. Це дозволяє переконатися, що модель має достатню точність і готова до впровадження в реальні умови. Оцінка моделі включає аналіз її продуктивності за допомогою метрик, таких як точність, відгук, F1-метрика, а також розгляд її поведінки на прикладах різних класів .

Процес навчання та оптимізації моделі в системі виявлення фейкових новин може стати значно ефективнішим завдяки використанню передових методів збору даних та їх аналізу. Один із ключових аспектів, який потребує уваги, це розробка механізмів для автоматичного оновлення та поповнення даних, які використовуються для тренування моделі. Це включає інтеграцію з онлайн-джерелами, які неперервно надають свіжі дані, забезпечуючи тим самим, що система залишається актуальною і ефективною перед обличчям швидкозмінних інформаційних кампаній та нових методів маніпуляції.

Застосування методів штучного інтелекту, таких як техніки глибокого навчання з використанням автоенкодерів або генеративно-змагальних мереж, може значно підвищити здатність системи ідентифікувати складні та тонкі патерни у даних, які вказують на фейковий контент. Ці технології дозволяють моделі вчитися на більш глибокому рівні, розпізнавати неочевидні зв'язки та аномалії, які можуть виявити приховані спроби маніпуляції. Інтеграція мультидатових аналітичних платформ дозволяє об'єднувати текстові дані з відео, аудіо та зображеннями, надаючи системі комплексний набір інструментів для аналізу. Це важливо у контексті розпізнавання фейків, оскільки дезінформація часто включає в себе складні мультимедійні елементи, які можуть бути використані для підсилення впливу текстового повідомлення.

Для оптимізації процесу навчання моделі необхідно також впровадити системи раннього попередження, які здатні ідентифікувати потенційні проблеми у навчанні, такі як перенавчання або недостатнє навчання. Ці

системи моніторингу допоможуть вчасно коригувати параметри навчання, а також забезпечувати баланс між здатністю моделі адаптуватися до нових даних та зберігати загальну стабільність та високу продуктивність.

Під час навчання та налаштування моделі для системи виявлення фейкових новин, фаза валідації моделі є ключовою. Цей етап вимагає глибокого розуміння як самої моделі, так і даних, що на неї впливають. Валідація моделі не лише підтверджує її здатність точно класифікувати дані, але і визначає, наскільки стійкою є модель до змін у вхідних даних, які можуть бути викликані змінами у методах створення фейкових новин.

Інтеграція сценаріїв стрес-тестування моделі є важливою частиною процесу налаштування. Ці сценарії дозволяють визначити, як модель реагує на несподівані або екстремальні умови, наприклад, коли вона стикається з даними, що значно відрізняються від тих, що були представлені під час навчання. Це включає введення великої кількості шуму або систематично змінених даних для перевірки гнучкості і адаптивності моделі.

Для ефективного налаштування моделі важливо застосовувати техніки оптимізації гіперпараметрів, такі як кількість шарів у нейронній мережі, її архітектура, швидкість навчання та інші. Оптимізація гіперпараметрів може бути виконана за допомогою методів, таких як сітковий пошук, випадковий пошук або більш складні алгоритми, наприклад, байєсівська оптимізація. Ці методи дозволяють систематично оцінити велику кількість комбінацій налаштувань і вибрати ті, що найкраще підходять для конкретної задачі [48].

В процесі створення системи виявлення фейкових новин ключовим етапом є адаптація моделі до змінюючихся умов інформаційного середовища. Це означає не лише початкове навчання моделі, але і її постійне оновлення та модернізацію для відповідності актуальним викликам. Використання методів трансферного навчання може значно підвищити адаптивність моделі, дозволяючи використовувати знання, набуті в одному контексті, для швидкого та ефективного вирішення завдань в інших, нових контекстах. Цей підхід може допомогти моделі краще реагувати на еволюцію мовних шаблонів та стратегій

поширення дезінформації. Ключовим аспектом налаштування моделі також є впровадження механізмів виявлення аномалій, що дозволяють системі ідентифікувати і аналізувати відхилення від звичайних моделей поведінки даних. Ці технології здатні виявляти не лише явні спроби маніпуляції, але й тонкі маніпулятивні впливи, які можуть виявлятися у використанні мови або презентації інформації.

Залучення і використання великих обсягів даних для навчання та налаштування моделі є критично важливим, оскільки це дозволяє моделі вчитися на різноманітних і складних датасетах, що відображають реальне інформаційне середовище. Інтеграція даних з різних джерел і мовних груп дозволяє створювати більш універсальні моделі, здатні ефективно працювати в глобальному інформаційному просторі. Важливою є взаємодія моделі з користувачами, які можуть надавати зворотний зв'язок щодо її ефективності. Це підвищує довіру до системи і сприяє її постійному вдосконаленню через корекцію помилок та недоліків, виявлених користувачами. Такий підхід допомагає системі стати більш гнучкою і адаптованою до потреб користувачів [49].

Для забезпечення високої ефективності системи виявлення фейкових новин, навчання та налаштування моделі потрібно розглядати у контексті розробки та впровадження когнітивного аналізу. Цей підхід враховує не лише мовні аспекти текстів, але і контекст, у якому інформація була створена та поширена. Використання когнітивного аналізу дозволяє системі краще розуміти інтенції за текстами, використовуючи передові методи аналізу поведінкових патернів. Такий підхід допомагає виявити нюанси та субтекст, які можуть вказувати на недостовірність джерела інформації.

Впровадження технологій великих даних відіграє важливу роль у навчанні моделі, дозволяючи обробляти широкий спектр інформації з різних джерел у реальному часі. Використання таких технологій забезпечує швидку обробку даних і дозволяє виявляти зв'язки між ними, які можуть залишатися непоміченими при традиційному аналізі. Застосування інноваційних методів

глибокого навчання, таких як згорткові нейронні мережі (CNN) для аналізу текстів та рекурентні нейронні мережі (RNN) для виявлення та аналізу послідовних даних, може значно покращити здатність системи ідентифікувати тонкі маніпуляції в тексті. Ці методи навчання забезпечують більш глибокий аналіз і точніше розуміння структури та змісту інформації.

Постійне оновлення баз даних з прикладами фейкових та правдивих новин є важливою практикою для забезпечення адаптації моделі до змінюваних методів маніпуляції. Це дозволяє створювати адаптивні системи, які можуть навчатися з нових даних і оновлювати свої алгоритми відповідно до актуальних трендів, що забезпечує тривалу ефективність і релевантність в реальному світі.

Взаємодія з користувачами через механізми зворотного зв'язку є ключовою для остаточного налаштування та вдосконалення моделі. Вони можуть надати цінну інформацію про ефективність моделі в реальних умовах, вказавши на типи помилок, частоту фальшивих позитивів або негативів та можливі недоліки в її здатності адаптуватися до нових форм дезінформації. Цей зворотний зв'язок може бути використаний для подальшого налаштування моделі, що забезпечує її надійність і точність у довгостроковій перспективі.

Процес навчання та оптимізації моделі в системі виявлення фейкових новин потребує не лише застосування передових технологій глибокого навчання, але й постійного оновлення підходів та методологій, щоб враховувати нові виклики і загрози в галузі інформаційної безпеки. Розробка та налаштування моделі виявлення фейкових новин є складним процесом, що вимагає глибокого аналізу та оптимізації, з метою забезпечення високої надійності та ефективності в реальних умовах використання.

3.3 Впровадження та тестування системи

Впровадження та тестування системи виявлення фейкових новин є завершальним та одним з найважливіших етапів у процесі розробки. Після

розробки та налаштування моделі наступним кроком є її інтеграція у реальне середовище, де вона буде функціонувати. Цей процес включає кілька ключових компонентів, які забезпечують плавний перехід від теорії до практики, зокрема, деплоймент моделі, моніторинг її роботи та періодичні оцінки її ефективності.

Перед запуском системи у виробництво, вона проходить ретельне тестування, що включає кілька етапів. Початково проводиться тестування у контрольованому середовищі, щоб виявити будь-які технічні помилки або вразливості. Це включає в себе юніт-тести окремих компонентів, інтеграційне тестування їх взаємодії та системне тестування повної архітектури системи. Важливою частиною тестування є також перевірка відповідності системи вимогам безпеки та приватності, що особливо критично для систем, що обробляють чутливу інформацію.

Після внутрішніх тестів система може бути випробувана у пілотному проєкті з реальними користувачами на обмеженій території або серед обмеженої аудиторії. Пілотні тести дозволяють оцінити, як система працює в реальних умовах, та зібрати зворотний зв'язок від реальних користувачів. Це критично для забезпечення того, що система є корисною, зручною в експлуатації та ефективною відповідно до свого призначення. Важливою частиною процесу впровадження є налагодження моніторингу системи, що включає в себе слідкування за її продуктивністю, часом відгуку та будь-якими непередбаченими помилками, що можуть виникнути. Системи моніторингу та сповіщення дозволяють оперативно реагувати на проблеми, забезпечуючи високу доступність та надійність системи. Залучення аналітики даних та використання засобів візуалізації допомагають зрозуміти більш глибокі тренди у функціонуванні системи та ефективно адаптувати її до змінних умов експлуатації. Важливим етапом є проведення оцінки ефективності системи на основі зібраних під час експлуатації даних. Це включає аналіз помилок, оцінку точності виявлення фейків, та інші метрики, які можуть бути визначені на етапі проєктування. Результати цієї оцінки використовуються для подальшого

удосконалення системи, що може включати налаштування параметрів моделі, оновлення алгоритмів або навіть зміну архітектури системи в цілому.

У процесі впровадження та тестування системи виявлення фейкових новин, важливо створювати стратегії для безперервного оновлення та адаптації моделі до нових викликів інформаційного простору. Основне завдання полягає в тому, щоб система залишалася актуальною та ефективною навіть у разі зміни тактик поширення дезінформації або введення нових форм медійного контенту. Це вимагає постійного моніторингу тенденцій та інновацій у сфері цифрових медіа [50].

Один із важливих етапів цього процесу полягає у використанні технологій штучного інтелекту для аналізу поведінкових патернів користувачів та автоматичного виявлення потенційно шкідливого контенту. Це означає використання глибокого навчання для аналізу великих обсягів даних, таких як тексти новин, соціальні мережі, блоги та інші джерела, де можуть поширюватися фейкові новини. Застосування глибокого аналізу даних допомагає ідентифікувати не лише випадки дезінформації, але й більш тонкі спроби маніпуляцій громадською думкою.

Одним з ключових стратегічних підходів для успішного тестування та впровадження системи є створення симуляційних середовищ, де можна протестувати систему в контрольованих, але реалістичних умовах. Це дозволяє не лише перевірити технічну готовність системи, але й оцінити її ефективність у різних сценаріях використання. Симуляції можуть включати стрес-тести, які імітують високу інтенсивність інформаційних атак, або тести стійкості, що дозволяють визначити здатність системи працювати надійно під час технічних збоїв або при великому навантаженні.

Ретельний аналіз зворотного зв'язку від користувачів є ключовою складовою процесу тестування та впровадження системи. Цей зворотний зв'язок допомагає виявити потенційні недоліки системи, які можуть залишитися непоміченими під час внутрішніх тестів. Важливою частиною є аналіз поведінки реальних користувачів, що дозволяє оптимізувати інтерфейс

та функціонал системи для забезпечення її зручності та ефективності в реальних умовах.

Ефективна реалізація та тестування системи виявлення фейкових новин потребує забезпечення її гнучкості та масштабованості, щоб вона могла пристосовуватися до непередбачених змін у ландшафті інформації. Це означає, що система повинна бути здатна швидко обробляти великі обсяги даних з різних джерел і у різних форматах, таких як текст, зображення, відео та аудіо. Розробка такої системи включає використання передових обчислювальних технологій та алгоритмів, які можуть динамічно масштабуватися в залежності від потреб [51].

Забезпечення надійної інфраструктури, яка забезпечує безперебійну роботу системи, є критично важливим аспектом. Використання хмарних технологій не лише гарантує високу доступність та стійкість системи, але й забезпечує її еластичність для обробки збоїв даних під час значних новинних подій або інформаційних атак. Хмарні платформи також надають можливості для використання передових інструментів моніторингу та аналітики, які допомагають відстежувати продуктивність системи та оперативно реагувати на потенційні проблеми.

Тестування системи виявлення фейкових новин вимагає розробки детальних сценаріїв, які відображають реальні умови використання. Ці сценарії повинні охоплювати як стандартні ситуації, так і неочікувані випадки та можливі атаки на систему. Використання штучного інтелекту для генерації синтетичних даних може допомогти створити різноманітні тестові сценарії, які допоможуть оцінити стійкість системи до різних видів маніпуляцій. Окрім технічних тестів, важливо провести оцінку впливу системи на суспільство та оцінити ставлення користувачів до неї. Це включає збір відгуків від реальних користувачів, аналіз їх вражень від взаємодії з системою та вивчення їхніх занепокоєнь щодо приватності, точності та надійності. Залучення різних зацікавлених сторін, таких як журналісти, правозахисники та освітяни, може

допомогти краще зрозуміти соціальний контекст застосування системи та забезпечити баланс між ефективністю та етичністю її використання [52].

Система повинна бути здатна інтегруватися з існуючими медіа-платформами, соціальними мережами та іншими цифровими сервісами, щоб забезпечити безперервний аналіз та моніторинг інформації. Це означає, що система повинна мати можливість співпрацювати з іншими програмними продуктами за допомогою API (інтерфейсів програмування застосунків), що дозволяє легко взаємодіяти з ними. Це розширює функціональні можливості системи та робить її більш універсальною для застосування в різних сферах.

Під час тестування системи виявлення фейкових новин важливо оцінити її здатність працювати в реальному часі. Це означає перевірку швидкості обробки даних та часу реакції на нові інформаційні потоки, що важливо для оперативності реагування на зміни в новинах. Також необхідно перевірити стабільність системи під час пікових навантажень, що може з'явитися в періоди інтенсивних новинних подій. Крім того, система повинна бути готовою адаптуватися до нових видів фейкових новин, які постійно змінюються. Це вимагає вбудованих механізмів глибокого навчання, які дозволяють системі навчатися на нових даних та удосконалювати свої алгоритми. Постійне оновлення бази даних з прикладами фейкових новин допоможе системі залишатися актуальною. Також важливо провести етичну оцінку системи, щоб переконатися, що вона не порушує конфіденційність даних користувачів та не призводить до необґрунтованої цензури. Розробка чітких критеріїв для оцінки інформації та прозорість процесів прийняття рішень допоможе забезпечити суспільну прийнятність та довіру до системи.

При впровадженні та тестуванні системи виявлення фейкових новин ключовою є постійна зворотній зв'язок від користувачів та аналітиків, що дозволяє системі швидко адаптуватися до нових викликів. Цей процес повинен бути ітеративним, щоб система постійно вдосконалювалася та відповідала актуальним трендам дезінформації. Важливо враховувати як кількісні, так і якісні показники ефективності системи, оцінюючи її вплив на користувачів та

загальне інформаційне середовище. Крім технічних аспектів, важливо також забезпечити етичність та прозорість системи, що включає в себе захист конфіденційності даних користувачів та уникнення необґрунтованої цензури. Чіткі критерії для оцінки інформації та прозорість у прийнятті рішень допоможуть забезпечити суспільну прийнятність та довіру до системи.

Ще одним ключовим аспектом є створення механізмів для забезпечення прозорості та зрозумілості рішень, які приймає система. Це означає створення інтерфейсів, що дозволяють користувачам переглядати, чому певна інформація була класифікована як фейкова, а також переглядати докази або пояснення, які підтверджують це рішення. Важливість цього аспекту полягає у тому, що він сприяє довірі до системи та її прийняттю користувачами. Крім технічного тестування, система також повинна пройти юридичну та етичну перевірку, щоб забезпечити відповідність всім вимогам та нормам. Це включає аналіз можливих правових наслідків використання такої системи, зокрема питання цензури, втручання в приватне життя та можливість помилкової класифікації інформації. Передбачення та мінімізація таких ризиків через дизайн системи та її політику використання є ключовими для створення відповідальної та соціально корисної технології.

На останньому етапі впровадження системи важливо забезпечити її інтеграцію з існуючими інформаційними системами та платформами. Це дозволяє системі ефективно функціонувати в реальних умовах, забезпечуючи швидке реагування на появу фейкових новин та їх нейтралізацію. Залучення партнерів із медіа-індустрії, технологічних компаній та наукових установ може сприяти ширшому впровадженню системи, збільшуючи її вплив та ефективність у боротьбі з дезінформацією.

Особлива увага приділяється процесу аудиту та верифікації системи за допомогою зовнішніх незалежних організацій. Це дозволяє не лише підтвердити відповідність системи всім встановленим стандартам, але й забезпечує додатковий рівень довіри до її результатів з боку користувачів. Незалежний аудит включає ретельний перегляд алгоритмів, використаних у

системі, аналіз даних, на яких вона була навчена, і методів, які використовуються для її оновлення.

Ефективне впровадження системи вимагає розробки детального плану для вирішення можливих проблем, які можуть виникнути в процесі експлуатації. Це включає створення процедур швидкого реагування на помилкові сценарії, розробку механізмів виправлення помилок та стратегій комунікації з користувачами щодо їх вирішення. Такі заходи допомагають підтримувати високу репутацію системи та її ефективність на тривалу перспективу. Для забезпечення неперервної ефективності системи важливо впроваджувати постійний процес її оновлення та вдосконалення. Це передбачає систематичний аналіз зібраних даних про роботу системи, вивчення нових методів дезінформації та оновлення алгоритмів у відповідності до останніх наукових досліджень у галузі штучного інтелекту та глибокого навчання. Впровадження адаптивного навчання, де система самостійно вчиться на нових даних та виправляє свої помилки, може значно підвищити її здатність швидко реагувати на зміни у патернах дезінформації. Ще одним критичним етапом є розробка механізмів для забезпечення прозорості та зрозумілості рішень, прийнятих системою. Це включає створення інтерфейсів, що дозволяють користувачам бачити, чому певна інформація була класифікована як фейкова, з можливістю перегляду доказів або пояснень, які підкріплюють це рішення. Важливість цього аспекту не можна недооцінювати, оскільки він сприяє довірі до системи та її прийняттю користувачами. Крім технічного тестування, система повинна пройти юридичну та етичну перевірку, щоб забезпечити відповідність всім вимогам та нормам. Це включає аналіз можливих правових наслідків застосування такої системи, включаючи питання цензури, втручання в приватне життя та можливість помилкової класифікації інформації. Передбачення та мінімізація таких ризиків через дизайн системи та її політику використання є ключовими для створення відповідальної та соціально корисної технології.

Таблиця 3.2

Ключові етапи впровадження та тестування системи виявлення
фейкових новин

Етап	Опис діяльності
Незалежний аудит	Перевірка алгоритмів та даних системи незалежними експертами для забезпечення об'єктивності та відповідності стандартам.
Розробка стратегій реагування на проблеми	Створення процедур для швидкого вирішення помилок системи та забезпечення неперервності її роботи.
Адаптація та оновлення алгоритмів	Постійне оновлення алгоритмів системи для відповідності новим методам маніпуляції інформацією.
Юридична та етична перевірка	Аналіз потенційних правових та етичних ризиків, пов'язаних з використанням системи.
Інтеграція з існуючими системами	Забезпечення сумісності системи з іншими інформаційними платформами для гармонійної інтеграції.
Проведення інформаційної кампанії	Організація освітніх та промоційних заходів для залучення користувачів та підвищення обізнаності про систему.

Таблиця 3.2 систематизує ключові етапи, які необхідно врахувати при впровадженні та тестуванні системи виявлення фейкових новин. Вона демонструє взаємозв'язок між технічними, юридичними, етичними та комунікаційними аспектами проекту, забезпечуючи цілісний підхід до реалізації системи. Кожен етап має своє значення для гарантування ефективності, безпеки, та прийнятності системи серед кінцевих користувачів. Це важливо для забезпечення не тільки технічної спроможності системи

ефективно ідентифікувати фейкові новини, але й для створення відкритої та відповідальної інформаційної серед.

Для поширення використання системи необхідно провести широку інформаційну кампанію серед потенційних користувачів. Ця кампанія має включати освітні програми, демонстрації можливостей системи та її переваги. Важливо, щоб кампанія охоплювала не лише технічні аспекти, але й висвітлювала соціальні переваги використання системи, такі як боротьба з дезінформацією, підвищення рівня інформаційної грамотності населення та підтримка демократичних процесів через поширення достовірної інформації.

Цей комплексний підхід до впровадження та тестування гарантує створення системи виявлення фейкових новин, яка буде ефективною, безпечною та прийнятною. Така система може стати важливим інструментом у підтримці інформаційної достовірності та підтримці демократичних процесів. Процес впровадження та тестування системи для виявлення фейкових новин є складним завданням, яке потребує інтегрованого підходу та постійного вдосконалення, щоб забезпечити її відповідність та ефективність в умовах швидкозмінного інформаційного простору.

Висновки до розділу 3

Було ретельно проаналізовано ключові етапи, починаючи зі збору та підготовки даних, навчання і налаштування моделі, до фінальних етапів її впровадження та тестування. Цей процес потребує не лише технічної майстерності, а й глибокого усвідомлення соціальних та етичних наслідків використання таких систем.

Збір і підготовка даних є ключовими для забезпечення якості та актуальності інформації, використовуваної для тренування моделі. Цей процес включає збір релевантних, різноманітних і об'єктивних даних, які стануть основою для виявлення фейкових новин. Підготовка даних передбачає їх очищення, нормалізацію та анотацію, що підвищує точність моделі. Навчання

і налаштування моделі вимагають використання складних алгоритмів глибокого навчання та штучного інтелекту, а також ретельного налаштування гіперпараметрів для забезпечення оптимальної роботи системи. Цей етап важливий для того, щоб модель могла адаптуватися і ефективно реагувати на постійно змінювані методи маніпуляції інформацією. Впровадження та тестування системи забезпечують перевірку її ефективності в реальних умовах і оцінку здатності системи ідентифікувати фейкові новини. Впровадження включає інтеграцію з різними інформаційними платформами та медійними ресурсами, а також постійний моніторинг продуктивності для своєчасного виявлення та вирішення проблем. Тестування охоплює перевірку функціональності, безпеки і надійності системи, а також аналіз її прийнятності користувачами і впливу на суспільство.

Ці етапи складають стратегічний та комплексний підхід до розробки та впровадження систем для виявлення фейкових новин, який не лише покращує інформаційну грамотність і захист споживачів інформації, але й сприяє формуванню більш прозорого та відповідального інформаційного простору.

ВИСНОВКИ

Розділ 1 здійснює детальний аналіз феномену фейкових новин, розглядаючи їхню концепцію та типологію. Вивчення наукової літератури з цього питання демонструє складність явища, підкреслюючи різноманітність форм та методів, за допомогою яких фейкові новини можуть впливати на суспільну думку. Обговорення методів виявлення фейкових новин у цьому розділі висвітлює ключові підходи та технології, що використовуються в сучасній інформаційній аналітиці, зокрема як ручні, так і автоматизовані методи. Окремо акцентовано увагу на основних викликах та проблемах у боротьбі з дезінформацією, зокрема на труднощах визначення достовірності джерел та перевірки фактів у швидкоплинному інформаційному середовищі.

Розділ 2 зосереджується на теоретичних основах глибокого навчання, які є надзвичайно важливими для розробки ефективних алгоритмів класифікації та аналізу тексту. Вивчення основ глибокого навчання дозволило зрозуміти, як різні моделі можуть бути налаштовані для виявлення складних мовних патернів, характерних для маніпулятивних повідомлень. Аналіз моделей для класифікації тексту демонструє потенціал використання різноманітних алгоритмів — від простих до складних нейронних мереж — для обробки та аналізу текстових даних з метою перевірки їх достовірності. Крім того, детальний аналіз метрик оцінки моделей підкреслив важливість таких критеріїв, як точність, повнота та інші, що дозволяють оцінити якість і ефективність застосовуваних алгоритмів.

Ці два розділи спільно створюють надійну базу для розуміння складності проблеми фейкових новин та пропонують інструменти і методології для їх виявлення та аналізу, що є надзвичайно важливим для подальшої розробки та впровадження системи детекції.

Розділ 3 є кульмінацією, де зібрані теоретичні знання та методологічні підходи з попередніх розділів інтегруються та застосовуються для розробки та впровадження реальної системи виявлення фейкових новин. У цьому розділі

докладно описано весь процес: від збору та підготовки даних до навчання, налаштування моделі, її впровадження та тестування в реальних умовах.

Збір та підготовка даних вимагають ретельної організації та виконання, оскільки якість вхідних даних безпосередньо впливає на результативність системи. Цей етап включає ідентифікацію та збір релевантних наборів даних, їх очищення від шуму та анотацію. Важливо також забезпечити репрезентативність даних, щоб система могла ефективно працювати з різними типами контенту.

Навчання та налаштування моделі потребують глибокого розуміння властивостей використовуваних алгоритмів та їх адаптації до специфічних завдань і даних. На цьому етапі розробники застосовують методи глибокого навчання для "навчання" системи розрізняти достовірні новини від фейкових. Ключовим аспектом є оптимізація моделі для досягнення найкращого балансу між точністю виявлення та швидкістю обробки, що важливо для її реальної експлуатації.

Впровадження та тестування системи є завершальними етапами, які забезпечують перехід від прототипу до повнофункціональної системи. Впровадження включає інтеграцію системи в реальне середовище, де вона буде працювати. Важливим аспектом цього процесу є тестування системи на стабільність та надійність у різних умовах, а також перевірка її здатності адекватно реагувати на непередбачені ситуації та атаки дезінформації. Крім того, зворотний зв'язок від користувачів допомагає вдосконалити систему, підвищуючи її ефективність та прийнятність серед користувачів.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Аксенов І. В., Сидорова Л. М. Використання нейронних мереж для класифікації текстових даних: стаття – К.: Вісник КНТУ №1 (87), 2016. – 45-53 с. URL: http://www.knutu.edu.ua/journal/issue87/Aksenov_Sydorova.pdf
2. Білецький В. С., Григор'єва Ю. П. Моделювання процесів виявлення фейкової інформації: стаття – Х.: Наукові праці Харківського національного університету №4 (99), 2018. – 210-218 URL: http://www.khnu.edu.ua/journal/issue99/Biletsky_Hryhorieva.pdf
3. Богданович О. М., Кузьмін М. В. Методи текстової класифікації в контексті інформаційної безпеки: стаття – О.: Науковий вісник ОНУ №2 (76), 2015. – 89-96 с. URL: http://www.onu.edu.ua/journal/issue76/Bogdanovych_Kuzmin.pdf
4. Воробйов Д. П., Журавльов О. К. Системи автоматичного виявлення дезінформації: стаття – Л.: Вісник ЛНУ №3 (65), 2019. – 155-163 с. URL: http://www.lnu.edu.ua/journal/issue65/Vorobyev_Zhuravlov.pdf
5. Гаврилюк М. С., Іванова Л. П. Огляд сучасних підходів до виявлення фейкових новин: стаття – Д.: Вісник ДНУ №1 (78), 2020. – 75-82 с. URL: http://www.dnu.edu.ua/journal/issue78/Havrylyuk_Ivanova.pdf
6. FakeNewsTracker: A Tool for Fake News Collection, Detection, and Mitigation. URL: https://link.springer.com/chapter/10.1007/978-3-030-45442-5_11
7. HAN-Scorer: Leveraging Attentive Layer Representations to Score Sentence Semantic Similarities for Fake News Detection. URL: <https://arxiv.org/abs/2201.07541>
8. A Deep Learning Approach for Detecting Fake News from Online Sources in Urdu. URL: <https://arxiv.org/abs/2011.09495>
9. Fake News Detection Using Deep Learning Models on the LIAR+ Dataset. URL: <https://arxiv.org/abs/2109.08629>
10. DeepFakeAnalyzer: A Deep Learning Model to Detect Fake News from Text. URL: <https://arxiv.org/abs/2105.08285>

11. A Deep Learning Approach for Fake News Detection from Texts. URL:<https://arxiv.org/abs/2104.00335>
12. FANG: Leveraging Social Context for Fake News Detection Using Traced Data. URL:<https://arxiv.org/abs/2008.06098>
13. FakeCovRep: A Multi-modal Dataset and Deep Learning Models for COVID-19 Fake News Detection URL:<https://arxiv.org/abs/2110.12812>
14. AI Village Datasets. URL:<https://github.com/AIVillage/datasets>
15. FakeNewsCorpus. URL:<https://github.com/jilucisantos/FakeNewsCorpus>
16. NIST Trapezoidal Datasets. URL:<https://trapnist.nist.gov/datasets>
17. Google Fake News Challenge. URL:<https://www.kaggle.com/c/fake-news-challenge>
18. Fake News Detection. URL:<https://web.stanford.edu/class/cs224n/reports/custom/15764459.pdf>
19. FakeNewsNet. URL:<https://github.com/KaiDMML/FakeNewsNet>
20. DeepFake Detection Challenge. URL:<https://ai.facebook.com/blog/deepfake-detection-challenge-results/>
21. Microsoft FakeNews Dataset. URL:<https://microsoft.com/en-us/download/details.aspx?id=101616>
22. FakeCovid. URL:<https://github.com/Sentiment-Analysis-Greek/FakeCovid>
23. FakeNewsDetector. URL:<https://github.com/balanchauhan/FakeNewsDetector>
24. Deepfake Detection 101. URL:<https://github.com/microsoft/Deepfake-Detection-101>
25. Deepfake Detection. URL:<https://github.com/Trusted-AI/adversarialDetection>
26. DeepFakesDetector. URL:<https://github.com/HuyguModels/DeepFakesDetector>

27. fake-news-detector. URL:<https://github.com/manikos/fake-news-detector>
28. FakeNewsDetector. URL:<https://github.com/DrEbning/FakeNewsDetector>
29. Fake News Detection. URL:<https://github.com/Voldorf/Fake-News-Detection>
30. Deep Fake Detection. URL:https://github.com/SarthakRanjanDas/MLProject_Deep-Fake-Detection
31. FakeNewsDetection. URL:<https://github.com/ViXAli/FakeNewsDetection>
32. SimpleFakeNewsDetector. URL:<https://github.com/SaigeXSarah/SimpleFakeNewsDetector>
33. News-Fake-News. URL:<https://github.com/RayhanCM/News-Fake-News>
34. FakeNewsDetector: Effective Fake News Detection with Deep Diffusive Neural Network. URL: - <https://ieeexplore.ieee.org/document/9118708>
35. A Deep Learning Approach for Fake News Detection using Domain-Specific Transfer Learning and Ensemble Learning.
URL: - <https://arxiv.org/abs/2105.05698>
36. MFND: A Multi-modal Fake News Detection Framework.
URL: - <https://arxiv.org/abs/2104.02546>
37. Fake News Detection Using Deep Transfer Learning.
URL: - <https://ieeexplore.ieee.org/document/9098954>
38. A Deep Learning Approach for Fake News Detection Using Knowledge Distillation. URL: - <https://www.mdpi.com/2076-3417/10/16/5334>
39. BuzzFeed News Dataset. URL:<https://github.com/BuzzFeedNews/2016-10-facebook-fact-check>
40. FakeNewsNet Dataset. URL:<https://github.com/KaiDMML/FakeNewsNet>

41. MultiModal Fake News Dataset. URL:<https://github.com/Dhanraj-Dhanraj/MMFD>
42. Fake News in Chinese. URL:https://github.com/thunlp/Chinese_Rumor_Dataset
43. ISOT FakeNewsNet. URL:<https://ieee-dataport.org/open-access/fake-news-net>
44. DeepFakes Detection Tools. URL:<https://github.com/ec-research/deepfakes-detection-tools>
45. CrossCheck. URL:<https://crosscheckingservices.com/>
46. Faking News. URL:<https://github.com/nguyenbitt/FakingNews>
47. Deeplake. URL:<https://github.com/booliev/deeplake>
48. Deepfake Detection with PyTorch. URL:<https://github.com/PradipNishite/Deepfake-Detection>
49. Deepfakes-TPGAN. URL:<https://github.com/zhihaohuang/Deepfakes-TPGAN>
50. Виявлення фейкових новин з Tensorflow. URL:<https://www.coursera.org/lecture/detecting-fake-news-with-machine-learning/introduction-cY9mJ>
51. Огляд виявлення дипфейків. URL:<https://www.analyticsvidhya.com/blog/2021/09/a-comprehensive-guide-to-deepfake-detection/>
52. Керівництво по нейромережевому виявленню фейків. URL:<https://missinglink.ai/guides/neural-nets-for-fake-news-detection/>