

**МАГІСТЕРСЬКА РОБОТА**

**МР. ШМ - 31.00.00.000 ПЗ**

**Група ШМ-24-3**

**Феценець Ангеліна**

**2025**

**Івано-Франківський національний технічний університет нафти і газу**

**Факультет інформаційних технологій**

**Кафедра інженерії програмного забезпечення**

**Феценець Ангеліна Петрівна**

(прізвище, ім'я, по батькові)

УДК 004.9  
(індекс)

## **МАГІСТЕРСЬКА РОБОТА**

**Моделі та методи класифікації відгуків користувачів засобами обробки**

**природної мови та глибокого навчання**

(назва роботи)

**Інженерія програмного забезпечення**

(назва освітньої програми)

**121 - Інженерія програмного забезпечення**

(шифр і назва спеціальності)

**Феценець А.П.**

(підпис, ініціали та прізвище здобувача освітнього ступеня)

**Науковий керівник**

**Вовк Роман Богданович, к.т.н., доцент**

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

**Допущено до захисту**

**Завідувач кафедри**

**доц.**

**Бандура В.В.**

(посада) (підпис) (дата) (ініціали та прізвище)

**Нормоконтроль**

**доц.**

**Вовк Р.Б.**

(посада) (підпис) (дата) (ініціали та прізвище)

Робота містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

**Івано-Франківськ – 2025**

**Івано-Франківський національний технічний університет нафти і газу**

Факультет інформаційних технологій

Кафедра інженерії програмного забезпечення

Освітній рівень магістр

Спеціальність 121 – Інженерія програмного забезпечення

ЗАТВЕРДЖУЮ:

Зав. кафедрою

ІІЗ

доц.

В.В. Бандура

“ 04 ” вересня 2025 р.

# ЗАВДАННЯ

## НА МАГІСТЕРСЬКУ РОБОТУ СТУДЕНТУ

**Феценець Ангеліна Петрівна**

(прізвище, ім'я, по-батькові)

**1. Тема магістерської роботи “ Моделі та методи класифікації відгуків користувачів засобами обробки природної мови та глибокого навчання”**

керівник проекту (роботи) Вовк Роман Богданович, к.т.н., доцент

затверджені наказом закладу вищої освіти від “ 05 ” листопада 2025 р. № 695/7

**2. Строк подання студентом проекту (роботи) 15 грудня 2025 р.**

**3. Вихідні дані до проекту (роботи) Концепції та формальні моделі і методи побудови інформаційних технологій обробки відгуків користувачів засобами глибокого навчання**

**4. Зміст розрахунково - пояснювальної записки(перелік питань, які потрібно розробити)**

1. Аналіз проблематики класифікації відгуків користувачів засобами обробки природної мови

2. Огляд сучасних досліджень та методів в аналізі настроїв

3. Методи та методологія класифікації відгуків користувачів засобами обробки природної мови

4. Імплементация моделей класифікації відгуків користувачів засобами обробки природної мови

**5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень)**

1. Алгоритм NLP (рис. 1.1)

2. Модель LSA<sub>n</sub>ArTe. Лексичний аналіз настроїв текстів із соціальних медіа (рис. 1.2)

3. Комплексна методологію аналізу настроїв на основі даних із Twitter (рис. 1.3)

4. Основні етапи виконання аналізу настроїв та виявлення емоцій (рис. 1.4)

5. Методики для аналізу настроїв та виявлення емоцій (рис. 1.5)

## 6. Консультанти розділів проекту (роботи)

Розділ	Консультант	Підпис, дата
Перевірка на плагіат	доц., к.т.н. Вовк Р.Б.	

7. Дата видачі завдання 04 вересня 2025 р.

Керівник \_\_\_\_\_

(підпис)

Завдання прийняв до виконання \_\_\_\_\_

(підпис)

## КАЛЕНДАРНИЙ ПЛАН

№ п/п	Назви етапів магістерської роботи	Строк виконання етапів роботи	Примітка
1	Підбір і вивчення літератури по темі магістерської роботи	17.09.2025	виконано
2	Аналіз проблематики класифікації відгуків користувачів засобами обробки природної мови	30.09.2025	виконано
3	Огляд сучасних досліджень та методів в аналізі настроїв	16.10.2025	виконано
4	Методи та методологія класифікації відгуків користувачів засобами обробки природної мови	10.11.2025	виконано
5	Імплементация моделей класифікації відгуків користувачів засобами обробки природної мови	22.11.2025	виконано
6	Порівняльний аналіз класифікаційних моделей на конкретних тестових випадках	03.12.2025	виконано
7	Затвердження пояснювальної записки роботи завідувачем кафедри	17.12.2025	виконано

Студент – магістр \_\_\_\_\_

(підпис)

Керівник роботи \_\_\_\_\_

(підпис)

## АНОТАЦІЯ

**Магістерська робота:** 80 с., 16 рис., 6 табл., 40 джерел.

**Тема:** Моделі та методи класифікації відгуків користувачів засобами обробки природної мови та глибокого навчання

**Мета роботи** - розробити та дослідити моделі й методи класифікації користувацьких відгуків із використанням засобів обробки природної мови та глибокого навчання.

**Об'єкт дослідження** - процес автоматизованої класифікації користувацьких відгуків у системах аналізу текстових даних.

**Предмет дослідження** - моделі, методи та інструменти класифікації текстів, що базуються на технологіях обробки природної мови та глибокого навчання.

### **Результати дослідження**

В роботі виконано імплементацію розроблених методів класифікації, проведено навчання та тестування моделей на експериментальному наборі даних, а також здійснено кількісну оцінку результатів.

### **Висновок**

Проведено порівняльний аналіз ефективності різних архітектур нейронних мереж для класифікації текстових даних та визначено оптимальні параметри гіпертюнінгу моделей глибокого навчання для задач аналізу тональності відгуків.

**ОБРОБКА ПРИРОДНОЇ МОВИ, КЛАСИФІКАЦІЯ ТЕКСТІВ, АНАЛІЗ ВІДГУКІВ, ГЛИБОКЕ НАВЧАННЯ, НЕЙРОННІ МЕРЕЖІ, BERT, LSTM, МАШИННЕ НАВЧАННЯ, АНАЛІЗ ТОНАЛЬНОСТІ.**

## ABSTRACT

**Master Thesis:** 80 pp., 16 fig., 6 tab., 40 sources.

**Topic:** Models and methods for classifying user feedback using natural language processing and deep learning

**The aim of the work** is to develop and investigate models and methods for classifying user feedback using natural language processing and deep learning.

**The object of the research** is the process of automated classification of user feedback in text data analysis systems.

**The subject of the research** is models, methods and tools for text classification based on natural language processing and deep learning technologies.

### **Research results**

The work implements the developed classification methods, trains and tests models on an experimental data set, and quantitatively evaluates the results.

### **Conclusion**

A comparative analysis of the effectiveness of different neural network architectures for text data classification is carried out and the optimal hypertuning parameters of deep learning models for tone analysis tasks are determined.

**NATURAL LANGUAGE PROCESSING, TEXT CLASSIFICATION, FEEDBACK ANALYSIS, DEEP LEARNING, NEURAL NETWORKS, BERT, LSTM, MACHINE LEARNING, TONE ANALYSIS.**

## ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ .....	10
ВСТУП.....	11
<b>РОЗДІЛ 1. АНАЛІЗ ПРОБЛЕМАТИКИ КЛАСИФІКАЦІЇ ВІДГУКІВ КОРИСТУВАЧІВ ЗАСОБАМИ ОБРОБКИ ПРИРОДНОЇ МОВИ .....</b>	<b>15</b>
1.1. Порівняння методологій класифікації та аналіз настроїв у галузі обробки природної мови .....	15
1.2. Значення відгуків на Amazon та застосування обробки природної мови (NLP).....	16
1.2.1. Роль клієнтських відгуків в електронній комерції.....	16
1.2.2. Теоретичні основи та алгоритм обробки природної мови (NLP)....	17
1.2.3. Застосування аналізу настроїв у дослідженні відгуків .....	20
1.3. Підходи до аналізу настроїв на основі лексиконів до глибокого навчання .....	20
1.3.1. Традиційні підходи на основі лексиконів .....	20
1.3.2. Методи класифікації на основі класичного машинного навчання..	21
1.3.3. Сучасні підходи глибокого навчання .....	21
1.3.4. Порівняльний аналіз лексичних підходів і глибокого навчання в аналізі настроїв .....	22
1.4. Огляд сучасних досліджень та методів в аналізі настроїв.....	24
1.4.1. Традиційні підходи та методи машинного навчання.....	24
1.4.2. Дослідження мовної адаптації та міжмовного аналізу .....	28
1.4.3. Застосування глибокого навчання та спеціалізованих фреймворків .....	30
Висновки до розділу .....	32
<b>РОЗДІЛ 2. МЕТОДИ ТА МЕТОДОЛОГІЯ КЛАСИФІКАЦІЇ ВІДГУКІВ КОРИСТУВАЧІВ ЗАСОБАМИ ОБРОБКИ ПРИРОДНОЇ МОВИ ТА ГЛИБОКОГО НАВЧАННЯ .....</b>	<b>34</b>

2.1. Розробка класифікаторів настроїв для відгуків Amazon.....	34
2.1.1. Архітектура системи аналізу настроїв.....	34
2.1.3. Технічні характеристики системи.....	37
2.2. Методологія підготовки набору даних для процесу аналізу настроїв ..	38
2.3. Проектний дизайн системи аналізу настроїв.....	41
2.3.1. Збір та попередня обробка даних.....	42
2.3.2. Аналіз настроїв та встановлення міток.....	42
2.3.3. Підготовка навчальних даних .....	42
2.3.4. Архітектура та навчання моделі .....	43
2.3.5. Розгортання, вивід та звітність .....	43
2.4. Збір та попередня обробка даних в процесі аналізу настроїв .....	44
2.4.1. Збір даних (Data Acquisition).....	44
2.4.2. Очищення та нормалізація тексту.....	44
Висновки до розділу .....	46

### РОЗДІЛ 3. ІМПЛЕМЕНТАЦІЯ МОДЕЛЕЙ ТА МЕТОДІВ ДЛЯ КЛАСИФІКАЦІЇ ВІДГУКІВ КОРИСТУВАЧІВ ЗАСОБАМИ ОБРОБКИ ПРИРОДНОЇ МОВИ .....

3.1. Методології класифікації настроїв. Лексичний підхід та використання бібліотеки TextBlob.....	48
3.1.1. Класифікація на основі Opinion Lexicon .....	48
3.1.2. Класифікація за допомогою бібліотеки TextBlob .....	49
3.2. Класифікація настроїв за допомогою моделі глибокого навчання на визначених мітках.....	50
3.2.1. Архітектура та компоненти моделі.....	50
3.2.2. Етап навчання моделі та оптимізація .....	51
3.2.3. Компіляція та оцінка.....	52
3.3. Визначення оптимального порогу класифікації настроїв.....	52
3.3.1. Емпіричне тестування порогових значень Opinion Lexicon .....	53
3.3.2. Емпіричне тестування порогових значень TextBlob.....	54

3.4. Оцінка продуктивності моделей класифікації настроїв.....	55
3.4.1. Методологія оцінки на основі матриці невідповідності.....	55
3.4.2. Метрики продуктивності.....	56
3.4.3. Результати оцінки моделей.....	56
3.5. Порівняльний аналіз методів класифікації настроїв у відгуках користувачів.....	59
3.6. Порівняльний аналіз класифікаційних моделей на конкретних тестових випадках.....	61
3.7. Реалізація дизайну користувацького інтерфейсу та взаємодії системи	64
Висновки до розділу.....	69
ВИСНОВКИ.....	70
ПЕРЕЛІК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ.....	73

## ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

NLP - Natural Language Processing - Обробка природної мови

LSTM - Long Short-Term Memory - Довга короткочасна пам'ять

RNN - Recurrent Neural Network - Рекурентна нейронна мережа

GloVe - Global Vectors for Word Representation - Глобальні вектори для  
представлення слів

NLTK - Natural Language Toolkit - Інструментарій природної мови

PPV - Positive Predictive Value - Позитивна прогностична цінність

TPR - True Positive Rate - Істинно позитивний показник

SGD - Stochastic Gradient Descent - Стохастичний градієнтний спуск

## ВСТУП

### **Актуальність теми.**

Сучасний розвиток цифрової економіки супроводжується стрімким зростанням кількості текстової інформації, яку генерують користувачі в онлайн-середовищі. Відгуки про товари, послуги, мобільні застосунки чи контент формують не лише уявлення про якість продукту, а й впливають на бізнес-рішення, маркетингові стратегії та репутацію компаній. Обробка таких текстових даних у великих масштабах вручну є неможливою, тому зростає потреба у використанні методів автоматизованої класифікації відгуків на основі обробки природної мови (NLP) та глибокого навчання.

Розвиток штучного інтелекту та нейронних мереж відкриває нові можливості для аналізу емоційного тону, визначення категорій відгуків, виявлення ключових тем і трендів. Особливого значення набуває створення ефективних моделей, здатних працювати з неструктурованими текстами, багатомовними корпусами та даними з різних джерел (соціальні мережі, форуми, платформи електронної комерції).

У цьому контексті дослідження спрямоване на розроблення моделей і методів класифікації відгуків користувачів із застосуванням сучасних підходів глибокого навчання, що базуються на архітектурах нейронних мереж, таких як LSTM, GRU, CNN та трансформери (BERT, RoBERTa тощо).

Робота поєднує теоретичний аналіз існуючих підходів до класифікації текстів і практичну реалізацію моделі, яка здатна забезпечити високу точність, узагальненість і стійкість до шуму даних. Результати дослідження мають вагомe значення для галузей аналітики даних, автоматизації клієнтських сервісів, маркетингових досліджень і побудови інтелектуальних систем рекомендацій.

Збільшення обсягів користувацького контенту в інтернеті зумовлює потребу в інтелектуальних інструментах його автоматизованого аналізу. Компанії, які працюють у сфері електронної комерції, IT та послуг, усе

частіше покладаються на аналітику відгуків користувачів для прийняття стратегічних рішень. Проте традиційні методи текстового аналізу не забезпечують необхідної точності через складність природної мови, багатозначність, сарказм, неформальні висловлювання та емоційні нюанси.

Методи глибокого навчання, зокрема рекурентні та трансформерні нейронні мережі, показують значне покращення у вирішенні завдань класифікації текстів, зокрема у виявленні тональності та тематичного розподілу. Поєднання таких моделей із методами обробки природної мови дає змогу створювати системи, що автоматично навчаються на реальних даних і здатні адаптуватися до контексту.

Таким чином, актуальність дослідження зумовлена необхідністю розробки надійних і високоточних моделей класифікації відгуків користувачів, які враховують сучасні тенденції розвитку NLP, глибинних нейронних архітектур і практичні потреби цифрових платформ. Результати дослідження можуть бути застосовані у реальних бізнес-процесах для підвищення ефективності взаємодії з клієнтами та автоматизації аналітики великих текстових масивів.

**Мета роботи** - розробити та дослідити моделі й методи класифікації користувацьких відгуків із використанням засобів обробки природної мови та глибокого навчання.

**Об'єкт дослідження** - процес автоматизованої класифікації користувацьких відгуків у системах аналізу текстових даних.

**Предмет дослідження** - моделі, методи та інструменти класифікації текстів, що базуються на технологіях обробки природної мови та глибокого навчання.

**Завдання дослідження:**

1. Проаналізувати сучасні підходи, методи та інструменти класифікації текстів і відгуків користувачів.
2. Визначити основні проблеми й виклики у використанні традиційних алгоритмів машинного навчання для обробки текстових даних.

3. Обґрунтувати вибір архітектури нейронних мереж для вирішення задачі класифікації відгуків.

4. Розробити модель класифікації текстів із використанням технологій глибокого навчання.

5. Провести навчання та тестування моделі на відкритих корпусах користувацьких відгуків.

6. Оцінити ефективність розробленої моделі та визначити напрями подальшого вдосконалення.

### **Методи дослідження**

Для досягнення мети використано методи аналізу, синтезу, узагальнення та систематизації наукових джерел; статистичні методи оцінки результатів; методи машинного та глибокого навчання (зокрема, LSTM, CNN, BERT, DistilBERT); методи обробки природної мови (токенізація, лематизація, векторизація, Word2Vec, TF-IDF); а також експериментальні методи тестування моделей на реальних даних.

### **Наукова новизна отриманих результатів**

Запропоновано удосконалена модель класифікації користувацьких відгуків, яка поєднує механізми глибокого навчання з попередньою мовною обробкою текстів. Удосконалено підхід до формування векторних представлень текстів на основі комбінованого використання методів TF-IDF і контекстних ембедингів. Визначено оптимальні параметри гіпертунінгу моделей глибокого навчання для задач аналізу тональності відгуків.

### **Практичне застосування результатів**

Розроблена модель класифікації може бути впроваджена у системи моніторингу клієнтських відгуків, CRM-платформи, сервіси аналітики соціальних медіа, а також у системи автоматизованого управління зворотним зв'язком у сфері електронної комерції. Отримані результати сприяють підвищенню якості аналітичних процесів і точності прогнозування поведінки користувачів.

**Структура магістерської роботи.** Представлена робота складається зі вступу, трьох розділів та висновків. Загальний обсяг роботи становить 80 сторінок, і містить 16 рисунків, 6 таблиць, перелік використаних джерел із 40 позицій і додаток.

# РОЗДІЛ 1. АНАЛІЗ ПРОБЛЕМАТИКИ КЛАСИФІКАЦІЇ ВІДГУКІВ КОРИСТУВАЧІВ ЗАСОБАМИ ОБРОБКИ ПРИРОДНОЇ МОВИ

## 1.1. Порівняння методологій класифікації та аналіз настроїв у галузі обробки природної мови

Аналіз настроїв (Sentiment Analysis) є актуальним та інтенсивно досліджуваним напрямком у галузі обробки природної мови (NLP). Метою даного дослідження є порівняльна оцінка продуктивності трьох різних методологій класифікації настроїв.

Для емпіричної оцінки використовувався великий набір даних відгуків Amazon, що містить понад 100000 рецензій від клієнтів. Класифікація настроїв виконувалась із застосуванням трьох підходів:

### 1. Лексичний підхід (Opinion Lexicon).

Оцінка настрою здійснювалася шляхом зіставлення слів у відгуках із попередньо визначеним словником (Opinion Lexicon), який містить слова з позитивною та негативною полярністю. Настрій визначався на основі співвідношення та кількості позитивних і негативних термінів, що з'являються в тексті.

### 2. Бібліотечний підхід (TextBlob).

Використовувалася бібліотека TextBlob для мови програмування Python, яка надає готові функціональні можливості для оцінки полярності настрою тексту.

### 3. Метод глибокого навчання (Нейронна мережа).

Розроблено та навчено модель нейронної мережі. Навчання моделі базувалося на зіркових оцінках, що супроводжували відгуки.

Було проведено порівняння продуктивності класифікаційних моделей, зосереджене на показниках точності:

- Точність лексичного методу (Opinion Lexicon): 64,38%
- Точність бібліотечного методу (TextBlob): 65,71%

- Точність моделі нейронної мережі: 96,46%

Отримані результати демонструють значну перевагу моделі нейронної мережі, яка досягла точності 96,46%, що істотно перевершує показники традиційних лексичних та бібліотечних методів (TextBlob та Opinion Lexicon).

Розроблена модель нейронної мережі пропонує високоефективне рішення для автоматизованої класифікації клієнтських відгуків на позитивні, негативні або нейтральні. Ця модель може бути імплементована брендами для оперативного та точного аналізу майбутніх відгуків, сприяючи вдосконаленню обслуговування клієнтів та стратегічному прийняттю рішень.

## **1.2. Значення відгуків на Amazon та застосування обробки природної мови (NLP)**

### *1.2.1. Роль клієнтських відгуків в електронній комерції*

Клієнтські відгуки на платформі Amazon є критично важливим елементом у процесі прийняття рішень під час онлайн-покупок. Вони слугують цінним джерелом емпіричної інформації та зворотного зв'язку від інших споживачів щодо якості, функціональності, ергономічності та ефективності продуктів.

Екосистема відгуків Amazon являє собою масштабне сховище різноманітних оцінок, починаючи від кількісних рейтингів (зірочок) до якісних, деталізованих текстових коментарів. Цей масив даних є незамінним ресурсом для споживачів, які прагнуть мінімізувати ризики та здійснити обґрунтовану покупку.

Для забезпечення автентичності та надійності система Amazon включає механізм "перевіреної покупки" (verified purchase), що підвищує довіру до наданої інформації. Таким чином, відгуки виступають потужним двостороннім інструментом:

- Для споживачів сприяють підвищенню поінформованості.

- Для виробників стимулюють зростання задоволеності клієнтів, лояльності та формування довіри, що є визначальними факторами в умовах висококонкурентного ринку онлайн-роздрібної торгівлі. Цей зворотний зв'язок дозволяє підприємствам отримувати критичні інсайти для ітеративного вдосконалення продуктів і послуг, формуючи клієнтоорієнтовану екосистему.

### *1.2.2. Теоретичні основи та алгоритм обробки природної мови (NLP)*

Обробка природної мови (NLP) є підгалуззю штучного інтелекту (AI), яка сфокусована на забезпеченні можливості комп'ютерного розуміння, інтерпретації та генерації людської мови. NLP інтегрує принципи лінгвістики, комп'ютерних наук та методів AI для розробки алгоритмів і моделей, призначених для аналізу та обробки даних природної мови.

Навчання цих моделей, як правило, відбувається на великих корпусах текстових даних (наприклад, наукові статті, літературні твори, соціальні медіа). Для оптимізації точності та продуктивності моделей NLP активно застосовуються методи машинного навчання. Враховуючи експоненційне зростання обсягів цифрового тексту, значущість NLP як інструменту для автоматизованого вилучення цінних знань та виконання завдань, які традиційно вимагають людського втручання, постійно зростає.

Представлена графічна схема (рис. 1.1) ілюструє конвеєр (pipeline) обробки тексту за допомогою моделі глибокого навчання, типової для задач аналізу настроїв (Sentiment Analysis). Алгоритм складається з послідовних етапів, які трансформують вхідний текстовий відгук у кінцевий класифікований вихід.

#### 1. Етап введення та векторизації (Input and Vectorization Processing)

Вхідний Текст (Input Text)

- Джерело: Сирий текст, наприклад, відгук Amazon ("This product is I love it").

- Мета: Текст є неструктурованими даними, які не можуть бути безпосередньо оброблені моделлю.

Токенізатор (Tokenizer) - перетворює вхідний текст у послідовність цілочисельних індексів (токенів), згідно з попередньо побудованим словником. Кожне слово (або частина слова) отримує унікальний числовий ідентифікатор.

Вихід: Послідовність чисел, наприклад, [1,5,2,8,3,4].

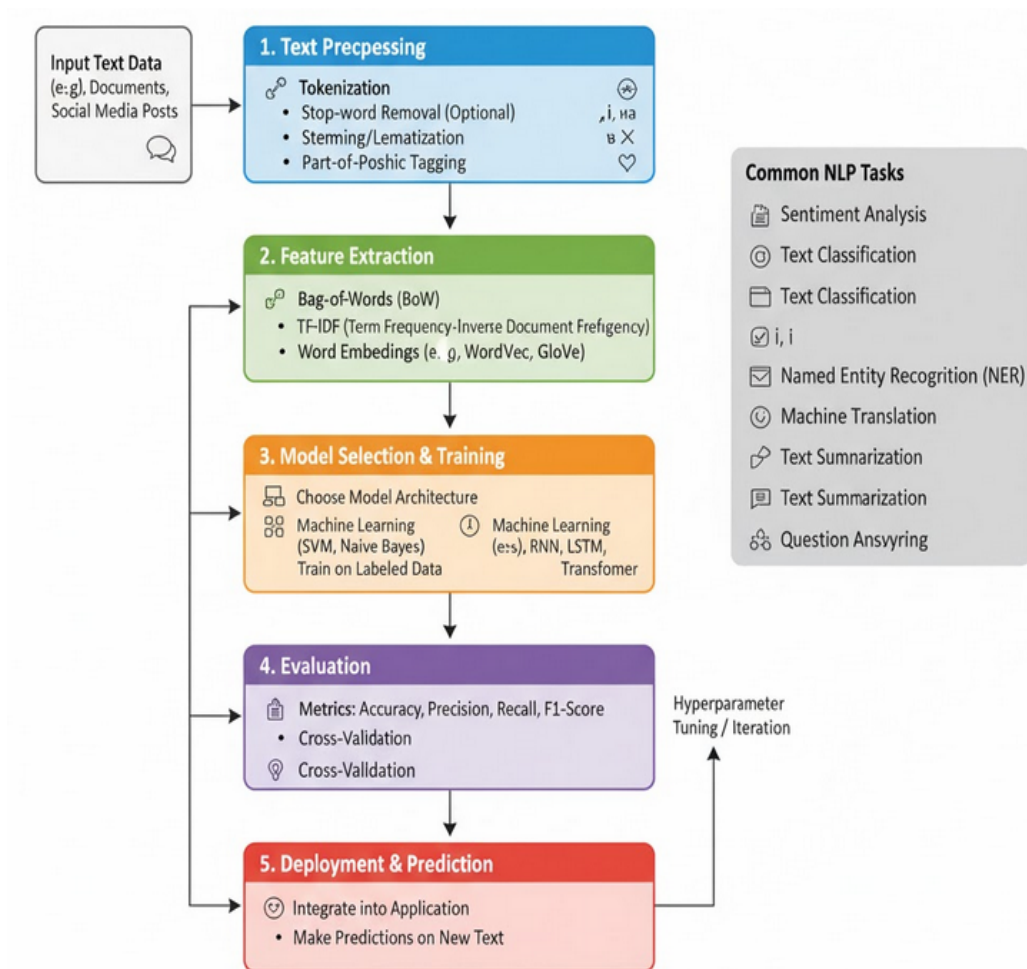


Рис. 1.1. Алгоритм NLP

## 2. Етап представлення ознак (Feature Consolidation)

Шар Векторизації (Embedding Layer) - це перший шар нейронної мережі, який перетворює дискретні індекси токенів на щільні вектори (embeddings). Ці вектори фіксують семантичні властивості слів: слова зі схожим значенням розташовуються близько в багатовимірному просторі.

Мета: перехід від символічного представлення до числового, яке містить лінгвістичну інформацію.

Вихід: Тривимірний тензор (Batch Size, Sequence Length, Embedding Dimension).

3. Етап консолідації та трансформації (Consolidation and Non-Linear Transformation)

Цей етап обробляє всю послідовність векторів, отриманих із Embedding Layer, для створення єдиного, фіксованого представлення всього вхідного відгуку. Типові механізми:

- Global Max Pooling вибирає максимальне значення ознаки по всій послідовності.

- Attention Mechanism присвоює різні ваги важливості різним токенам у тексті, щоб модель могла сфокусуватися на найбільш значущих словах (наприклад, "не" чи "жахливо").

Шар Dropout включається для запобігання перенавчанню (overfitting), випадковим чином "виключаючи" частину нейронів під час навчання (наприклад, Dropout 0.2).

Вихід: Вектор фіксованої довжини, що резюмує весь зміст відгуку.

4. Етап класифікації (Classification)

Вихідний шар (Output Layer) - повнозв'язний шар із кількістю юнітів, що дорівнює кількості класів (у даному випадку 3: позитивний, негативний, нейтральний). Функція активації (Softmax) застосовується для перетворення вихідних значень на ймовірності, сума яких дорівнює одиниці. Це дозволяє моделі визначити, до якого класу відгук належить найбільше.

Модель видає ймовірність належності до кожного класу (наприклад, негативний 1.35%, позитивний 96.46%, нейтральний 2.19%), класифікуючи відгук до класу з найвищою ймовірністю (у цьому прикладі – позитивний).

Таким чином, алгоритм NLP успішно трансформує неструктурований текст у кількісний, класифікований результат, що має пряме практичне застосування для бізнес-аналітики.

### *1.2.3. Застосування аналізу настроїв у дослідженні відгуків*

У межах даного дослідження буде застосовано ключову функцію NLP — аналіз настроїв (Sentiment Analysis).

Аналіз настроїв (або думки) — це процес класифікації існуючих текстових відгуків за категоріями позитивний, негативний або нейтральний. Така класифікація дозволяє компаніям ідентифікувати сильні та слабкі сторони їхніх продуктів.

Застосування аналізу настроїв надає можливість:

- Отримувати якісну інформацію про загальні теми та емоційну тональність відгуків.
- Визначати конкретні аспекти продукту, які викликають задоволення або проблеми у клієнтів.

Цей метод дозволяє брендам та виробникам вийти за межі простого зіркового рейтингу, надаючи глибинний інструмент для аналізу та інтерпретації клієнтського зворотного зв'язку.

## **1.3. Підходи до аналізу настроїв на основі лексиконів до глибокого навчання**

Аналіз настроїв (Sentiment Analysis) є динамічною та високоактивною дослідницькою галуззю, що перебуває на стику обробки природної мови (NLP) та машинного навчання. Вона охоплює низку підходів і методик, призначених для аналізу, вилучення та класифікації емоційної тональності в текстових даних.

### *1.3.1. Традиційні підходи на основі лексиконів*

Лексичний підхід є однією з фундаментальних методологій аналізу настроїв. Він спирається на використання лексиконів настроїв (sentiment lexicons) або словників, які містять слова з попередньо визначеною полярністю (позитивною, негативною) та відповідними ваговими оцінками.

Лексикони можуть бути створені вручну або автоматично за допомогою лінгвістичних евристик.

Загальноприйняті техніки класифікації на основі лексикону включають:

- Обчислення загального настрою тексту через підрахунок частоти входжень позитивних і негативних слів.
- Обчислення інтегральної оцінки настрою шляхом агрегування полярності на рівні слів, враховуючи їхні ваги.

### *1.3.2. Методи класифікації на основі класичного машинного навчання*

Машинне навчання широко застосовується в аналізі настроїв. У цьому контексті алгоритми навчаються на мічених наборах даних для автоматичної класифікації настрою.

#### 1. Класифікатори.

Використовуються такі алгоритми, як наївний Байєс (Naive Bayes), машини опорних векторів (Support Vector Machines, SVM) та дерева рішень (Decision Trees).

#### 2. Інженерія ознак (Feature Engineering).

Цей етап є критично важливим і полягає у вилученні релевантної інформації з тексту, включаючи синтаксичні патерни, n-грами та різноманітні лінгвістичні ознаки.

#### 3. Оптимізація.

Основний фокус навчання моделі та оптимізації гіперпараметрів спрямований на досягнення високої точності та узагальнюючої здатності моделі на невидимих даних.

### *1.3.3. Сучасні підходи глибокого навчання*

Останнім часом моделі глибокого навчання (Deep Learning) набули значної популярності, демонструючи високу ефективність у задачах аналізу настроїв, особливо завдяки їхній здатності автоматично вивчати ознаки та моделювати складні залежності.

Рекурентні нейронні мережі (Recurrent Neural Networks, RNN), як-от:

- Прості RNN.
- Об'єднані рекурентні нейронні мережі (Gated Recurrent Unit, GRU).
- Мережі довгої короткочасної пам'яті (Long Short-Term Memory, LSTM) успішно використовуються для захоплення послідовних залежностей та контекстуальної інформації в текстових послідовностях, що є життєво важливим для точного розуміння емоційного забарвлення.

Ця еволюція від лексичних методів до складних моделей глибокого навчання відображає постійний пошук більш точних, адаптивних та масштабованих рішень для аналізу настроїв.

#### *1.3.4. Порівняльний аналіз лексичних підходів і глибокого навчання в аналізі настроїв*

Порівняння традиційних лексичних методів та сучасних підходів глибокого навчання виявляє значні відмінності у вимогах, гнучкості та кінцевій продуктивності в задачі аналізу настроїв.

Лексичний підхід (Lexicon-Based) спирається на попередньо визначені словники слів із фіксованою полярністю (настроєм).

Таблиця 1.1.

Переваги і недоліки лексичного підходу

<b>Переваги</b>	<b>Недоліки</b>
Простота та швидкість	Низька точність та гнучкість
Не вимагає тривалого навчання чи великих обчислювальних ресурсів.	Не розрізняє контекстуальні нюанси, іронію чи саркастичні висловлювання.
Відсутність необхідності в мічених даних	Проблема неологізмів
Не потребує великих мічених наборів даних для навчання, що спрощує розгортання.	Не може обробляти нові слова, сленг, специфічну термінологію (залежить від наявності в лексиконі).
Легка інтерпретованість	Залежність від якості лексикону

Переваги	Недоліки
Результати класифікації легко пояснити, оскільки вони базуються на конкретних словах лексикону.	Ефективність прямо пропорційна повноті та якості словника; вимагає ручного оновлення.

Отже, лексичні методи є швидкими та прозорими, але крихкими і погано масштабуються для розуміння складної людської мови.

Підхід глибокого навчання (Deep Learning — RNN/LSTM) використовує багат шарові нейронні мережі для автоматичного вивчення ознак і патернів настрою.

Таблиця 1.2.

#### Переваги і недоліки підходу глибокого навчання

Переваги	Недоліки
Висока точність та надійність	Високі вимоги до ресурсів
Здатний досягати найвищої точності (як показано раніше, понад 96%) завдяки моделюванню складних залежностей.	Вимагає значної обчислювальної потужності (GPU) та тривалого часу на навчання.
Розуміння контексту та семантики	Необхідність у великих мічених даних
Моделі, як-от LSTM, ефективно захоплюють контекст, послідовні залежності та можуть розрізняти іронію/заперечення.	Для досягнення високої продуктивності необхідні великі, якісно мічені набори даних.
Автоматичне вивчення ознак	Низька інтерпретованість (Black Box)
Усуває необхідність у ручній інженерії ознак, оскільки модель автоматично вивчає найбільш релевантні ознаки з сирого тексту.	Складно пояснити, чому модель зробила конкретну класифікацію, оскільки процес є нелінійним і багат шаровим.

Отже, глибоке навчання пропонує максимальну продуктивність і здатність до узагальнення, але вимагає значних інвестицій у дані та обчислювальні потужності.

Таким чином, для швидкого прототипування або в умовах обмежених даних можуть бути прийнятними лексичні методи. Однак, для отримання найвищої точності та аналізу складних, реальних клієнтських відгуків з контекстуальними нюансами, підходи глибокого навчання (LSTM/GRU) є науково обґрунтованими та кращими.

#### **1.4. Огляд сучасних досліджень та методів в аналізі настроїв**

Аналіз настроїв (Sentiment Analysis) є активною дослідницькою областю в обробці природної мови (NLP), що характеризується розробкою та апробацією різноманітних методологій для класифікації емоційної тональності текстових даних. Упродовж останніх років було реалізовано значну кількість проектів, що охоплюють як традиційні, так і сучасні парадигми.

##### *1.4.1. Традиційні підходи та методи машинного навчання*

Ряд робіт зосереджено на застосуванні класичних алгоритмів машинного навчання та лексичних методів.

Класифікація відгуків у Twitter була реалізована в роботі [3]. Автори запропонували метод класифікації твітів на позитивні, негативні або нейтральні, використовуючи API Twitter для збору даних та застосовуючи алгоритми Наївного Байєса для аналізу настроїв. В дослідженні представили метод автоматичного збору корпусу, який може бути використаний для навчання класифікатора настроїв. Було застосовано TreeTagger для розмітки частин мови (POS-tagging) де спостерігали різницю у розподілі між позитивними, негативними та нейтральними наборами. На основі спостережень зроблено висновок, що автори використовують синтаксичні структури для опису емоцій або констатації фактів. Деякі мітки частин мови (POS-tags) можуть бути сильними індикаторами емоційно забарвленого тексту.

Зібраний корпус було використано для навчання класифікатора настроїв який здатний визначати позитивні, негативні та нейтральні настрої документів. Класифікатор базується на мультиноміальному класифікаторі Наївного Байєса (multinomial Naïve Bayes), який використовує N-грами та мітки частин мови (POS-tags) як ознаки.

Тему розширення лексиконів досліджували в роботі [4] і навели підхід на основі словника, складеного вручну, але виявили його недостатність через обмежене покриття. Вони наголосили на необхідності розширення лексикону та створення нового тезаурусного лексикону для підвищення точності класифікації.

Було запропоновано модель LSA<sub>n</sub>ArTe, що використовує лексичний підхід, залучаючи самостійно створений та вручну анотований лексикон, який містить приблизно 10769 речень. Цей розширений лексикон слугує основою для аналізу настроїв (SA), дозволяючи ідентифікувати та категоризувати настрої, висловлені в онлайн-відгуках про кавові продукти.

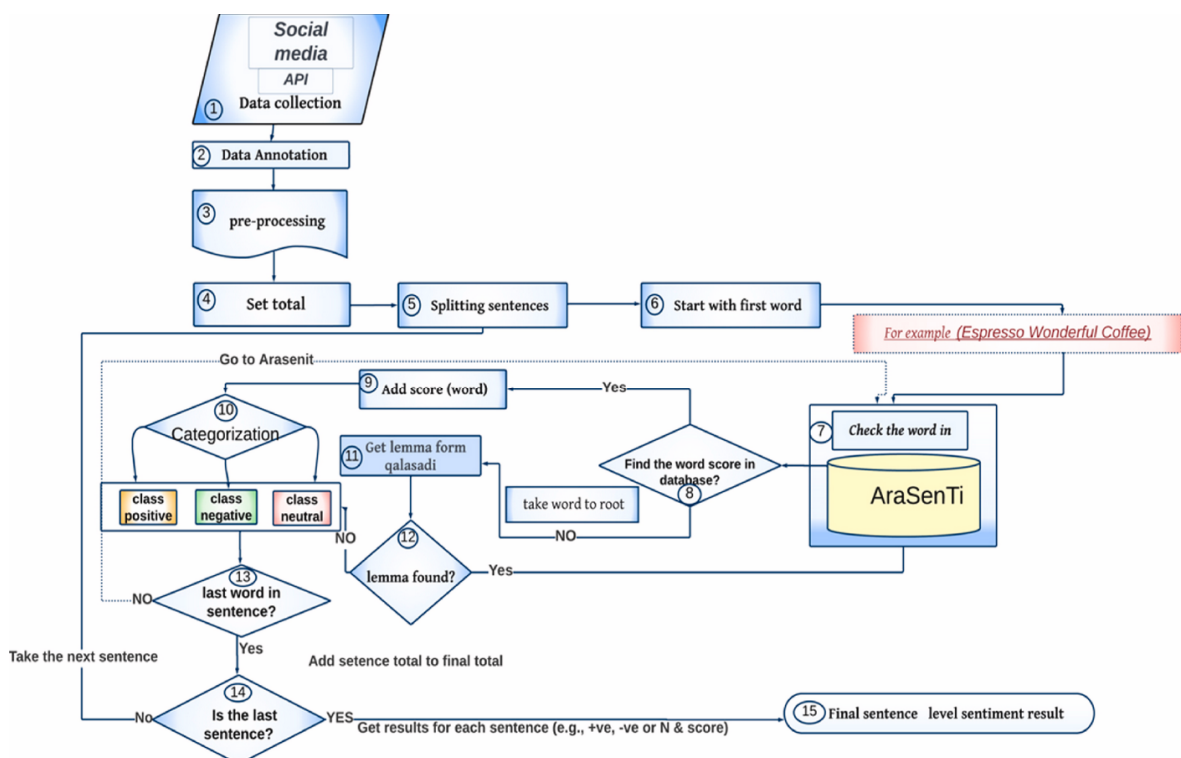


Рис. 1.2. Модель LSA<sub>n</sub>ArTe. Лексичний аналіз настроїв текстів із соціальних медіа

Робочий процес моделі LSA<sub>n</sub>ArTe детально описаний на рис. 1.2, який ілюструє поетапний процес від збору даних та аналізу настроїв до фінальної класифікації настроїв. Ця візуалізація сприяє розумінню комплексного підходу, застосованого для аналізу настроїв в тексті, пов'язаному з кавою, підкреслюючи залежність моделі від спеціалізованого лексикону та її аналітичні можливості.

Порівняльний аналіз ML-моделей виконано в [10]. Автори провели аналіз настроїв COVID-19-пов'язаних твітів, порівнюючи кілька алгоритмів машинного навчання, зокрема випадковий ліс (RF), XGBoost, машини опорних векторів (SVC) та Класифікатор Extra Trees (ETC). Вони використовували такі методи вилучення ознак, як TF-IDF та Bag-of-Words (BoW), і дійшли висновку, що ETC показав найвищу точність (93%).

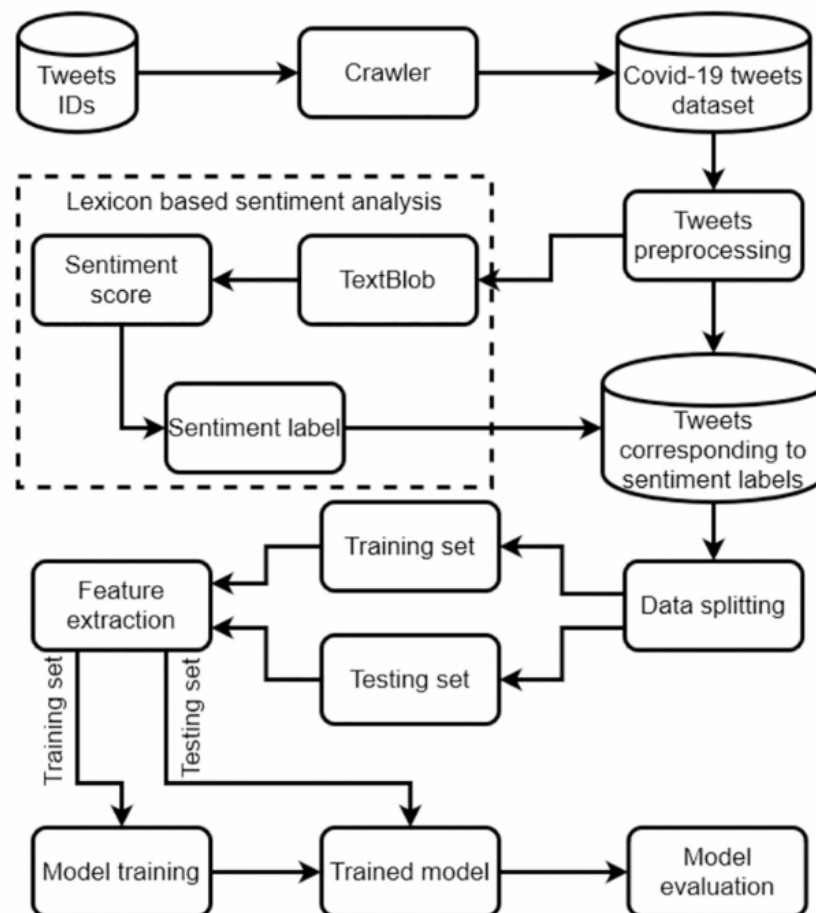


Рис. 1.3. Комплексна методологію аналізу настроїв на основі даних із Twitter (твітів, пов'язаних із COVID-19)

Тематична схема запропонованої методології представлена на рис. 1.3. Після вилучення твітів за допомогою розробленого внутрішнього краулера (in-house built crawler), дані піддаються попередній обробці (preprocessing). Попередня обробка є критично важливим етапом, який впливає на точність навчальних моделей. Тут видаляються з твітів стоп-слова, імена користувачів, посилання, розділові знаки та числові значення, а також застосовуються техніки стеммінгу.

Схема (рис. 1.3) демонструє гібридний підхід, де лексичний інструмент (TextBlob) використовується для автоматичного мічення великого корпусу, який потім використовується для навчання більш універсального класифікатора машинного навчання.

В дослідженні [11] описуються гібридні підходи, зазначивши, що хоча лексичний підхід є ефективним для аналізу настроїв та емоцій, словниковий підхід є більш адаптивним та легшим у застосуванні. Процес аналізу настроїв та виявлення емоцій проходить через кілька етапів, як показано на рис. 1.4, включаючи збір набору даних, попередню обробку, вилучення ознак, розробку моделі та оцінку.

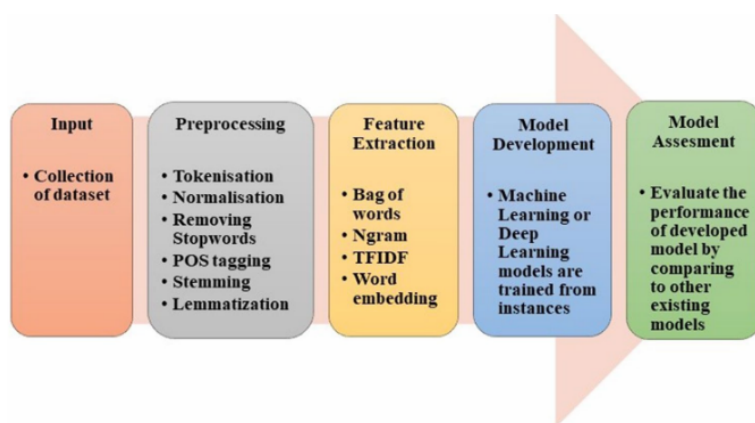


Рис. 1.4. Основні етапи виконання аналізу настроїв та виявлення емоцій

На рис. 1.5 представлено різні методиками для аналізу настроїв та виявлення емоцій, які широко класифікуються на: лексичний підхід, підхід на основі машинного навчання та підхід на основі глибокого навчання.

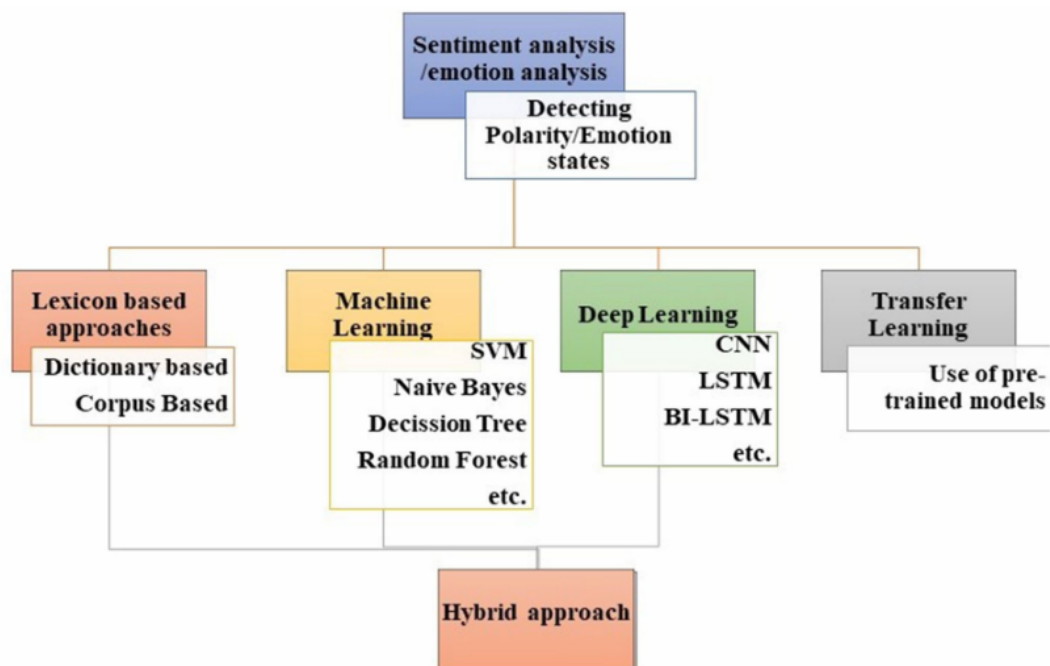


Рис. 1.5. Методики для аналізу настроїв та виявлення емоцій

Гібридний підхід є комбінацією статистичних методів та методів машинного навчання, що застосовується для подолання недоліків обох цих підходів. Трансферне навчання (Transfer learning) також є підмножиною машинного навчання, яке дозволяє використовувати попередньо навчену модель в інших подібних предметних областях.

#### 1.4.2. Дослідження мовної адаптації та міжмовного аналізу

Окремий напрямок зосереджений на розширенні аналізу настроїв на інші мови та міжкультурні контексти.

В роботі [3] досліджувались арабські лексикони. Автори працювали над покращенням існуючих арабських лексиконів настроїв з низьким покриттям. Вони створювали арабські переклади англійських лексиконів настроїв, використовуючи зібрані арабські твіти та арабський переклад лексикону емоцій NRC (National Research Council).

Було створено нові арабські лексикони настроїв із застосуванням методів віддаленого нагляду (distant supervision), та була продемонстрована їхня корисність в аналізі настроїв публікацій у соціальних медіа. Крім того,

було здійснено переклад існуючих англійських лексиконів настроїв (чотирьох, створених вручну, та двох, згенерованих автоматично) на арабську мову за допомогою Google Translate. Було показано, що ці лексикони покращують продуктивність порівняно з конкурентоспроможним базовим класифікатором, який використовує різні ознаки, засновані на поверхневій формі. Арабський Діалектний Хештег Лексикон виявився особливо корисним, але додавання ознак із перекладених лексиконів додатково підвищило точність класифікації. Нарешті, було проаналізовано підмножину автоматично перекладених записів лексикону настроїв, щоб оцінити ступінь збереження настрою після перекладу. Також було ідентифіковано різні причини, які можуть призводити до помилкових записів у перекладеному лексиконі. Усі розроблені лексикони надаються у вільний доступ.

Щодо міжнародного аналізу настроїв, то його досліджували в [7] як концепцію міжнародного аналізу настроїв за допомогою системи Lydia, яка розпізнає іменовані сутності та аналізує їхній розподіл у дев'яти мовах. Автори підкреслили необхідність розробки системи, яка дозволяє виконувати аналіз без втрати значення через простий переклад на англійську, і виявили послідовність розрахованих оцінок настроїв у різних мовах. Вони також запропонували розрахунок полярності настроїв для міжмовного порівняння.

В роботі [8] запропонували перший великомасштабний набір даних настроїв Twitter із ручною анотацією для чотирьох поширених мов у Нігерії, що включав 30000 анотованих твітів на мову. Вони також створили лексикони для трьох нігерійських мов на основі набору даних NaijaSenti і використовували для аналізу різні локалізовані варіанти BERT (mBERT, RemBERT, AfriBERTa, mDEBERTaV3 та XLM-R).

Пропонуються методи, що уможливлюють збір, фільтрацію та анотування даних для таких мов з низьким рівнем ресурсів. Крім того, запроваджується вручну анотований лексикон настроїв для трьох мов (хауса, іґбо та йоруба). Надано порівняльні експерименти (benchmark experiments) на

наборі даних настроїв Twitter з використанням сучасних попередньо навчених мовних моделей та трансферного навчання. Результати свідчать, що моделі, специфічні для конкретних мов, та мовно-адаптивне тонке налаштування демонструють найкращу продуктивність у середньому. NaijaSenti має потенціал стимулювати інтерес до аналізу настроїв та інших низхідних задач NLP у зазначених мовах.

### 1.4.3. Застосування глибокого навчання та спеціалізованих фреймворків

Сучасні дослідження активно використовують трансформерні архітектури та генеративні моделі.

В роботі [13] представили уніфіковану генеративну структуру для аналізу настроїв на основі аспектів (ABSA). Вони прагнули вирішити кілька завдань ABSA (ідентифікація аспектних термінів, полярності та термінів думки) у межах єдиної моделі, використовуючи попередньо навчену модель послідовність-послідовність BART.

Аналіз настроїв на основі аспектів (Aspect-based Sentiment Analysis, ABSA) є задачею деталізованого (fine-grained) аналізу настроїв (SA), метою якої є ідентифікація терміна аспекту (a), його відповідної полярності настрою (s) та терміна думки (opinion term) (o).

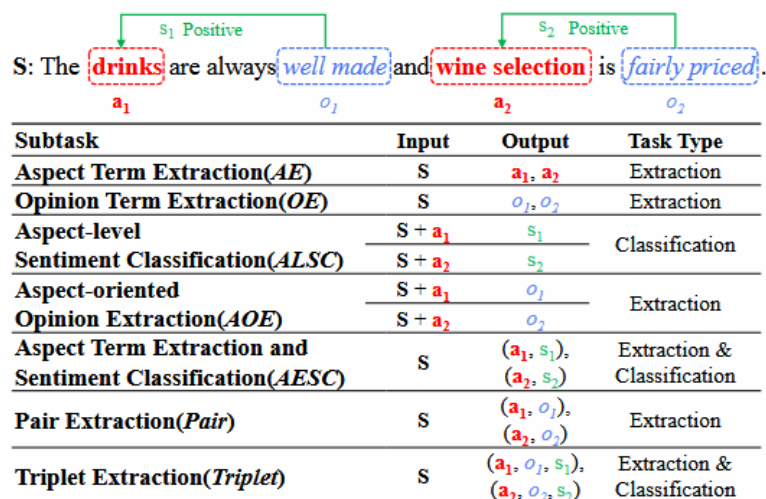


Рис. 1.6. Ілюстрація семи підзадач у межах ABSA

Наприклад, у реченні: "The drinks are always well made and wine selection is fairly priced" (Напої завжди добре приготовані, а вибір вин має помірну ціну), термінами аспектів є "drinks" (напої) та "wine selection" (вибір вин), їхня полярність настрою є "positive" (позитивна) для обох, а термінами думки є "well made" (добре приготовані) та "fairly priced" (помірна ціна).

На основі комбінації (a,s,o) існує сім підзадач у межах ABSA. Ці підзадачі узагальнено на рис. 1.6.

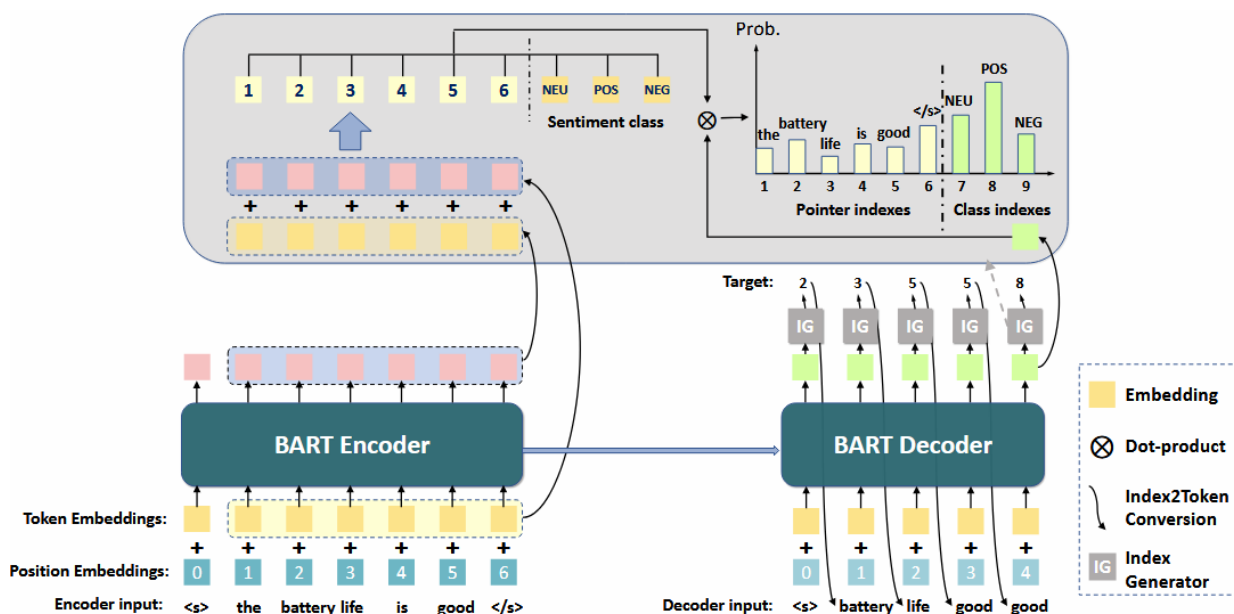


Рис. 1.7. Архітектура генеративної моделі

На рисунку 1.7 показано приклад процесу генерації для підзадачі трійка (Triplet subtask), де вхідна послідовність (source) – "<s> the battery life is good </s>", а цільова послідовність (target) – "2 3 5 5 8 6" (показано лише часткову послідовність декодера, де 6 (</s>) має бути наступним індексом генерації).

Модуль "Index2Token Conversion" конвертує індекси на відповідні токени. Зокрема, pointer index буде перетворено на його відповідний токен у source text, а class index буде перетворено на відповідні токени класу.

Вектори вбудовування у всіх блоках отримуються з тієї ж матриці вбудовування (embedding matrix). Для підвищення продуктивності генерації

використовуються різні position embeddings у вхідній та цільовій послідовностях.

Ця архітектура є типовою для генеративних моделей у NLP, які переформулюють завдання класифікації (наприклад, ABSA) у завдання послідовність-послідовність.

В роботі [15] застосовували модель BERT на наборах даних Twitter для вивчення громадської думки щодо впливу COVID-19, а в дослідженні [16] вивчали настрої, пов'язані з вакцинами проти COVID-19, використовуючи пакет R *syuzhet* та лексикон емоцій NRC. Вони зафіксували пік емоції довіри після оголошення Pfizer про ефективність вакцини, що свідчить про вищу прийнятність вакцин порівняно з попередніми аналогічними кампаніями.

Узагальнюючи, сфера аналізу настроїв демонструє постійну методологічну еволюцію від лексиконів до складних трансформерних моделей, із зростаючим акцентом на міжмовну адаптацію та аналіз спеціалізованих, великомасштабних наборів даних у реальному часі.

### **Висновки до розділу**

У першому розділі роботи проведено теоретичний аналіз проблематики класифікації відгуків користувачів засобами обробки природної мови (NLP). Визначено, що актуальність цієї проблеми зумовлена зростанням обсягів текстових даних у цифровому середовищі та необхідністю їх автоматизованої інтерпретації. Досліджено ключові методології аналізу настроїв, зокрема лексикон-орієнтовані підходи, класичні алгоритми машинного навчання та сучасні методи глибокого навчання. Порівняльний аналіз показав, що традиційні лексиконні моделі відзначаються простотою реалізації, проте обмежені у відтворенні контексту та семантики. Методи машинного навчання продемонстрували більшу адаптивність, однак потребують ретельного налаштування ознак. Найвищу ефективність виявлено у глибоких моделях, які здатні самостійно формувати семантичні представлення текстів.

Розглянуто роль клієнтських відгуків у формуванні бізнес-аналітики, зокрема на прикладі платформи Amazon, де вони є ключовим джерелом інформації про якість продуктів. Проведений огляд сучасних досліджень виявив тенденцію до використання багатомовних трансформерних моделей, здатних до міжмовної адаптації. Узагальнено, що проблема класифікації настроїв є міждисциплінарною та потребує поєднання лінгвістичних і математичних підходів. Результати аналізу створили теоретичну основу для формування методології побудови моделей класифікації, представленої в наступному розділі.

## РОЗДІЛ 2. МЕТОДИ ТА МЕТОДОЛОГІЯ КЛАСИФІКАЦІЇ ВІДГУКІВ КОРИСТУВАЧІВ ЗАСОБАМИ ОБРОБКИ ПРИРОДНОЇ МОВИ ТА ГЛИБОКОГО НАВЧАННЯ

### 2.1. Розробка класифікаторів настроїв для відгуків Amazon

Метою даного проекту є розробка та впровадження моделі, здатної точно класифікувати відгуки Amazon за трьома категоріями настроїв: позитивний, негативний або нейтральний. Досягнення високої точності класифікації забезпечується інтеграцією кількох методів машинного навчання та глибокого навчання. Такий аналіз настроїв має високу прикладну значущість як для споживачів, які прагнуть отримати об'єктивну інформацію про продукт чи послугу, так і для фахівців із маркетингу, яким необхідно оцінити громадську думку та репутацію свого бренду.

#### 2.1.1. Архітектура системи аналізу настроїв

Запропонована система для досягнення точної класифікації настроїв інтегрує етапи попередньої обробки тексту, лексичні ресурси та моделі глибокого навчання (LSTM).

##### 1) Попередня обробка тексту

Початковим етапом є попередня обробка тексту для забезпечення узгодженості даних і підвищення ефективності моделей. Цей процес включає:

- Нормалізація регістру - переведення тексту в нижній регістр.
- Очищення - видалення пунктуації та зайвих пробілів.
- Токенізація - розбиття тексту на окремі лексичні одиниці.
- Видалення стоп-слів (Stop-word removal) - усунення часто вживаних, але малоінформативних слів.
- Обробка скорочень - розгортання або нормалізація скорочень слів.

##### 2) Лексичні ресурси та оцінка настроїв

Для присвоєння оцінок настроїв використовуються існуючі лексичні ресурси:

- Opinion Lexicon - використовується для ідентифікації слів із попередньо визначеною позитивною чи негативною полярністю та агрегації цих полярностей для обчислення оцінки настрою для всього тексту.

- TextBlob - застосовується як альтернативний метод, що використовує власні внутрішні можливості аналізу настроїв для визначення емоційного тону тексту. Внутрішній аналіз TextBlob може забезпечити підвищену точність та гнучкість класифікації порівняно з простими лексичними підрахунками Opinion Lexicon.

### 3) Модель глибокого навчання

Для більш точної класифікації використовуються розширені моделі аналізу настроїв, зокрема мережі довгої короткочасної пам'яті (LSTM). LSTM-мережі здатні захоплювати контекстуальну інформацію та послідовні залежності у тексті, що є критичним для розуміння складних мовних конструкцій.

#### 2.1.2. *Методологія реалізації та оцінки*

Процес реалізації та оцінки моделі детально ілюструється блок-схемою моделі (рис. 2.1).

##### 1. Збір даних.

Дані відгуків для окремих продуктів збираються з Amazon за допомогою спеціалізованого розширення браузера Google Chrome. Сирі відгуки компілюються у структурований набір даних.

##### 2. Класифікація та розмітка даних.

Текст відгуку класифікується трьома незалежними методами, які слугують для мічення та порівняння:

- Класифікація на основі зіркових рейтингів (основна мітка для навчання нейронної мережі).

- Класифікація на основі Opinion Lexicon (агрегована полярність).

- Класифікація на основі оцінки полярності TextBlob.



Рис. 2.1. Алгоритм роботи запропонованої моделі

### 3. Навчання Моделі.

Нейронна мережа навчається за допомогою послідовної моделі TensorFlow Keras. Модель навчається на мітках, визначених методом зіркових рейтингів, для розпізнавання позитивних, негативних та нейтральних текстів.

### 4. Оцінка.

Після навчання та тестування продуктивність кожної класифікації (LSTM, Opinion Lexicon, TextBlob) оцінюється за допомогою метрик матриці невідповідності (confusion matrix metrics).

### 2.1.3. Технічні характеристики системи

Апаратні та програмні характеристики наведені в таблицях 2.1 та 2.2 відповідно.

Таблиця 2.1.

#### Апаратні характеристики

Компонент	Тестовий сценарій (Opinion Lexicon Threshold)	Тестовий сценарій (TextBlob Threshold)
<b>Процесор</b>	Intel i5 7500 або AMD Ryzen 3 3300X	Intel i7 10700K або AMD Ryzen 9 3900X
<b>Графічний процесор (GPU)</b>	NVIDIA GTX 1050	NVIDIA GTX 2080 Ti або новіший
<b>Оперативна пам'ять (RAM)</b>	8 ГБ	16 ГБ
<b>Дисковий простір</b>	5 ГБ	5 ГБ
Включає бібліотеки. Розмір набору даних може варіюватися.		

Таблиця 2.2.

#### Програмні характеристики

Категорія	Специфікація
<b>Мови Програмування</b>	Python 3.10
<b>Середовище розробки (IDE)</b>	PyCharm, Jupyter Notebook
<b>Фреймворк</b>	Flask
<b>Ключові бібліотеки</b>	TensorFlow, Keras (для процесів глибокого навчання), Pandas (для обробки даних), TextBlob, NLTK (для NLP-задач), Scikit-Learn (для класичного ML та метрик), Matplotlib (для візуалізації).

## 2.2. Методологія підготовки набору даних для процесу аналізу настроїв

Підготовка великого та неструктурованого текстового набору даних, як-от відгуки Amazon, є критично важливим етапом для забезпечення ефективності та точності моделей обробки природної мови (NLP). Процес включає очищення, нормалізацію та форматування даних.

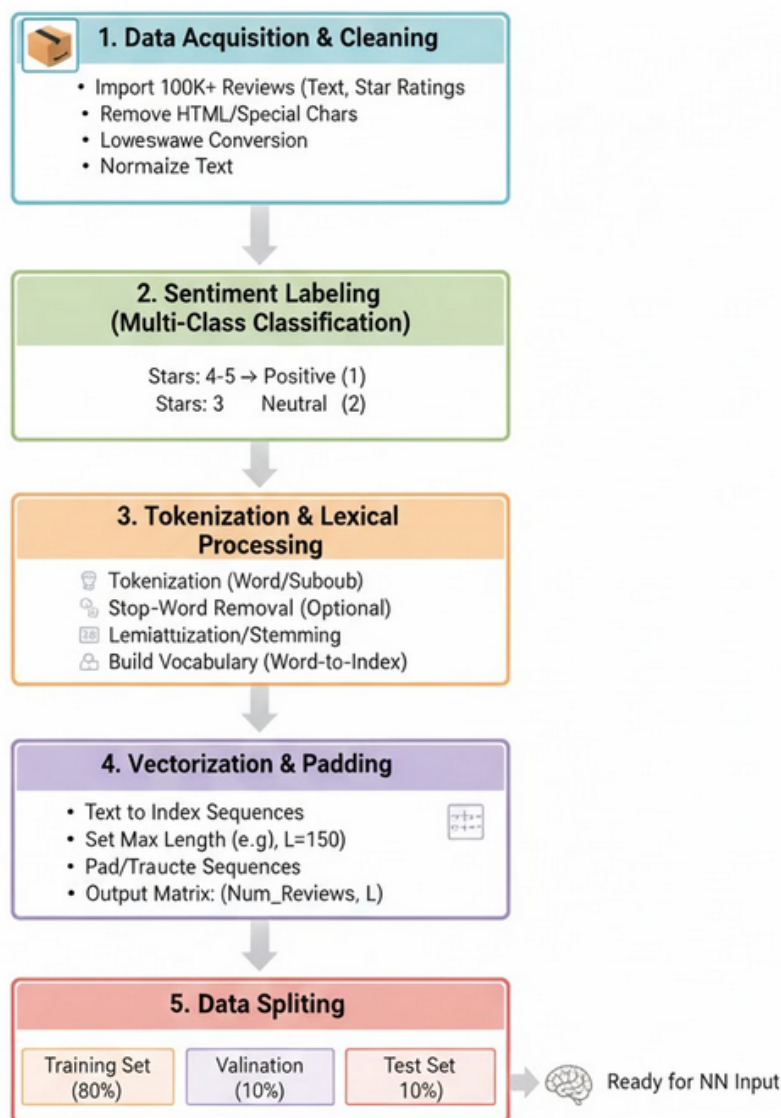


Рис. 2.2. Етапи методології підготовки набору даних Amazon для аналізу настроїв

### 1. Збір та попередня обробка даних (Data Acquisition and Cleaning)

Таблиця 2.3.

## Етапи попередньої обробки даних

Етап	Опис	Технічні операції
<b>Збір</b>	Отримання набору даних, що містить понад 100000 відгуків. Кожна одиниця даних повинна включати <b>текст відгуку</b> та відповідний <b>зірковий рейтинг</b> (1-5 зірок).	Імпорт даних (наприклад, з файлів CSV або JSON).
<b>Очищення</b>	Усунення артефактів, які можуть негативно вплинути на модель.	Видалення HTML-тегів, спеціальних символів та небуквених знаків. Перетворення тексту до нижнього регістру (lowercase).
<b>Нормалізація</b>	Обробка скорочень, сленгу та емодзі, якщо це необхідно.	Заміна загальноприйнятих скорочень на повні форми.

## 2. Маркування настрою (Sentiment Labeling)

Оскільки моделі машинного навчання потребують міток (labels), зіркові рейтинги використовуються для деривації емоційної полярності (настрою).

Таблиця 2.4.

## Маркування даних

Рейтинг (зірки)	Присвоєна мітка настрою	Категорія
4 та 5	Позитивний	1
3	Нейтральний	0
1 та 2	Негативний	-1 або 2 (залежно від схеми кодування)

Застосування трьох категорій (позитивний, негативний, нейтральний) перетворює задачу на багатокласову класифікацію.

## 3. Токенізація та лексична обробка (Tokenization and Lexical Processing)

Цей етап перетворює текст у формат, придатний для обробки нейронною мережею.

Токенізація - розбиття тексту відгуку на індивідуальні одиниці (токени), які можуть бути словами або підсловами.

Видалення стоп-слів (Stop-Word Removal) - видалення поширених, але малоінформативних слів (наприклад, "і", "в", "на"). Примітка: для аналізу настроїв цей крок іноді пропускають, оскільки стоп-слова можуть містити заперечення ("не добре").

Лемматизація / Стеммінг - зведення слів до їхньої базової форми (наприклад, "біг", "біжить", "бігли" → "бігти") для зменшення розміру словника та підвищення узагальнюючої здатності моделі.

Побудова словника всіх унікальних токенів та присвоєння кожному токену унікального цілочисельного індексу.

4. Векторизація та вирівнювання послідовностей (Vectorization and Padding)

Нейронні мережі працюють із числовими векторами фіксованої довжини.

Перетворення в індекси - кожен відгук перетворюється з послідовності слів на послідовність цілочисельних індексів згідно зі словником.

Вирівнювання (Padding) - оскільки відгуки мають різну довжину, всі послідовності приводяться до єдиної максимальної довжини (наприклад,  $L=150$  токенів). Коротші послідовності доповнюються нулями (padding), а довші — обрізаються (truncating).

$$\text{Input Matrix Size} = (\text{Кількість Відгуків}) \times L$$

5. Розподіл набору даних (Data Splitting)

Набір даних розділяється на три підмножини для коректної оцінки моделі:

- Тренувальний набір (Training Set) - використовується для навчання моделі (наприклад, 80% даних).

- Валідаційний набір (Validation Set) - використовується для налаштування гіперпараметрів та запобігання перенавчанню (наприклад, 10% даних).

- Тестовий набір (Test Set) - використовується для фінальної, незалежної оцінки продуктивності моделі (наприклад, 10% даних).

Цей структурований процес гарантує, що дані готові для подальшої подачі в шар Embedding Layer нейронної мережі.

### **2.3. Проєктний дизайн системи аналізу настроїв**

Представлений дизайн системи описує модульну архітектуру для розробки та розгортання моделі точної класифікації настроїв на основі текстових даних (наприклад, класифікація відгуків клієнтів платформи Amazon за трьома категоріями настроїв: позитивний, негативний або нейтральний).

#### *Модуль 1. Користувацький інтерфейс (User Module)*

Цей модуль забезпечує взаємодію користувача з аналітичною системою, надаючи такі функціональні можливості:

- Перегляд даних: надання доступу до візуалізації або перегляду структурованого набору даних.

- Текстове введення: можливість інтерактивного введення довільного текстового фрагменту.

- Функція прогнозування настрою (Predict Sentiment): запуск механізму класифікації для визначення настрою введеного тексту.

- Функція перекладу настрою (Translate Sentiment): запуск механізму для інтерпретації або перекладу визначеного настрою.

#### *Модуль 2. Системна архітектура та обробка даних (System Module)*

Цей модуль охоплює повний конвеєр обробки даних, навчання та розгортання моделі.

### *2.3.1. Збір та попередня обробка даних*

Збір даних (Data Acquisition). Система здійснює збір відгуків клієнтів з визначеного джерела (наприклад, онлайн-ринку). Зібрані дані мають переважно текстовий формат, доповнений кількісними індикаторами настрою (наприклад, рейтингами зірочками).

Попередня обробка (Data Preprocessing) - зібраний масив даних піддається очищенню та нормалізації для підвищення якості аналізу:

- Очищення тексту - застосовуються методи для видалення шуму, зокрема пунктуації та стоп-слів. Весь текст конвертується у нижній регістр.
- Нормалізація - скорочення слів розгортаються до їхніх повних форм. Очищений текст зберігається в окремому полі для подальшої обробки.

### *2.3.2. Аналіз настроїв та встановлення міток*

Система виконує первинний аналіз настроїв та готує мітки.

Настрій класифікується за трьома категоріями (позитивний, негативний, нейтральний) на основі зіркових рейтингів, пов'язаних із відгуками.

Лексична оцінка (Opinion Lexicon) розраховується оцінка настрою (sentiment score) для кожного відгуку за допомогою Opinion Lexicon. Ця оцінка забезпечує кількісну метрику для настрою.

Оцінка полярності (TextBlob) розраховується оцінка полярності настрою (sentiment polarity score) для кожного відгуку за допомогою TextBlob. Ця оцінка надає кількісну характеристику емоційного тону.

### *2.3.3. Підготовка навчальних даних*

Система готує дані для навчання моделі глибокого навчання.

Відгуки та відповідні мітки настроїв розділяються на навчальні (training) та тестові (testing) набори.

Кодування цільової змінної - цільові мітки настрою кодуються за допомогою one-hot encoding для сумісності з моделями класифікації.

Токенізація та вбудовування - текстові дані токенізуються за допомогою токенізатора. Вбудовування слів (word embeddings) перетворює токени на числові вектори. Попередньо навчена модель вбудовування GloVe використовується для створення матриці вбудовування, що відображає слова у відповідні векторні представлення із захопленням семантичної інформації.

#### *2.3.4. Архітектура та навчання моделі*

Розробляється модель глибокого навчання, яка включає:

- Шар вбудовування (Embedding Layer) - використовує попередньо навчену матрицю вбудовування.

- Рекурентні шари (Recurrent Layers) - наприклад, LSTM, які призначені для захоплення послідовних залежностей та контекстуальної інформації в тексті.

- Щільні шари (Dense Layers) виконують фінальну класифікацію та генерують прогнози.

Модель навчається на підготовленому навчальному наборі з визначеною кількістю епох та розміром пакету. Продуктивність моделі контролюється за допомогою валідаційного набору для моніторингу метрик втрат та точності. Застосовуються механізми ранньої зупинки (early stopping) та збереження моделі для запобігання перенавчанню та фіксації оптимальної моделі.

#### *2.3.5. Розгортання, вивід та звітність*

Навчена та оцінена модель зберігається та може бути завантажена для класифікації настрою нових, необроблених відгуків клієнтів. Розгорнута модель приймає попередньо оброблений текст, виконує токенізацію/доповнення та генерує прогнози.

Система видає передбачені мітки настроїв. Прогнози використовуються для аналізу тенденцій настроїв. Генеруються звіти та

візуалізації (наприклад, розподіл настроїв, динаміка настроїв у часі, порівняння настроїв між продуктами), що надають стратегічні інсайти.

## **2.4. Збір та попередня обробка даних в процесі аналізі настроїв**

Етап підготовки даних є фундаментально важливим для забезпечення надійності та високої точності моделей машинного навчання та глибокого навчання.

### *2.4.1. Збір даних (Data Acquisition)*

Збір даних здійснювався за допомогою розширення для браузера Google Chrome, спеціально адаптованого для вилучення відгуків про окремі продукти з платформи Amazon.com.

Розширення мало технічне обмеження, дозволяючи завантажувати не більше 1000 відгуків за один раз для кожного продукту.

Для формування достатнього обсягу вибірки було виконано ручний відбір понад 80–90 продуктів, з подальшим індивідуальним завантаженням відгуків для кожного.

Оскільки одночасне завантаження відгуків із кількох продуктів було неможливим, зібрані файли CSV були вручну об'єднані в єдиний агрегований файл. Із початкового набору стовпців (включаючи "Резюме відгуку", "Ідентифікатор продукту", "Голоси" тощо) для подальшого аналізу були відібрані лише дві змінні: "Загалом" (кількісний рейтинг зірок) та "Текст відгуку" (необроблений текстовий контент). Копії цих двох стовпців були використані для створення нового DataFrame Pandas.

### *2.4.2. Очищення та нормалізація тексту*

Очищення тексту має на меті підготовку текстових даних до аналізу та моделювання, забезпечуючи узгодженість та чистоту даних. Для попередньої обробки були застосовані такі методи:

- Переведення в нижній регістр (lowercasing) - увесь текстовий корпус був приведений до єдиного нижнього регістру. Ця процедура є критичною для підтримки узгодженості та запобігання дублюванню токенів (наприклад, "Good" і "good" розглядаються як ідентичні).

- Розгортання скорочень (contraction expansion) - для забезпечення коректного аналізу та ідентичності подібних фраз використовувалася бібліотека скорочень. Вона розгортає поширені скорочення (наприклад, "can't" → "cannot", "I'm" → "I am").

- Видалення пунктуації (punctuation removal) - застосовувалася функція `remove_punctuation()`, яка видаляє всі знаки пунктуації, визначені у стандартному наборі `string.punctuation` (включно з символами: `!`, `"`, `#`, `\$`, `%`, тощо).

- Видалення стоп-слів (stop-word removal) - для підвищення інформативності тексту використовувалася бібліотека NLTK Stopwords (набір англійських стоп-слів). Функція `remove_stopwords()` фільтрує слова, які мають низьку семантичну цінність для класифікації настроїв (наприклад, `the`, `a`, `is`, `was`, `through`, `by`). Крім того, ми також видаляємо будь-які слова, що складаються лише з цифр.

- Додаткове очищення: крім того, з тексту видаляються будь-які слова, що складаються виключно з числових символів.

- Видалення крайніх пробілів (Whitespace Trimming) - будь-які початкові або кінцеві пробіли були видалені для забезпечення узгодженості формату тексту. Основна мета цієї операції – забезпечення узгодженості даних та уникнення помилок при подальшій токенизації та аналізі. Тексти, які семантично ідентичні, але містять різні кінцеві пробіли (наприклад, "відгук " та "відгук"), розглядаються моделлю як різні. Whitespace trimming усуває цю різницю.

Зайві пробіли можуть призвести до створення некоректних або порожніх токенів на етапі токенизації, що негативно впливає на якість словника та продуктивність моделі.

Приклад трансформації подано в таблиці 2.5.

Таблиця 2.5.

Приклад тексту до і після обробки

<b>До попередньої обробки (приклад сирого відгуку)</b>	<b>Після попередньої обробки (очищений текст)</b>
"I had two 27in monitors that pump out heat big time. I needed another for work, so I bought one of these. Nice, light, bright, and no heat. After a week I bought another for my gaming PC. They're beautiful. I didn't get the ones with the adjustable stand. I didn't need that and they're perfect.Pro tip: They're shipped in the box that shows what's inside, so the 2nd time I checked the box in checkout for Amazon to pack it in one of their boxes."	"two 27 in monitors pump heat big time needed another work bought one nice light bright heat week bought another gaming pc beautiful get ones adjustable stand need perfectpro tip shipped box shows inside 2nd time checked box checkout amazon pack one boxes"

### Висновки до розділу

У другому розділі розроблено методологічну основу побудови системи класифікації відгуків користувачів із використанням засобів NLP та глибокого навчання. Визначено загальну архітектуру системи аналізу настроїв, що включає етапи збору, попередньої обробки, нормалізації, токенизації та класифікації текстових даних. Обґрунтовано вибір технічних засобів, бібліотек і фреймворків для реалізації процесу обробки, зокрема Python, TensorFlow, Keras, NLTK та TextBlob. Запропонована методика підготовки корпусу даних базується на поєднанні ручного анотування та автоматизованої лінгвістичної нормалізації. Проведено формалізацію процесу створення навчальної вибірки, що забезпечує репрезентативність і баланс позитивних та негативних відгуків. Розроблено структуру бази даних, яка підтримує зберігання текстів, метаданих та результатів аналізу. Сформовано алгоритм навчання моделі, що охоплює етапи вибору архітектури нейронної мережі, визначення функції втрат і методів

оптимізації. Описано проєктний дизайн системи, включно з компонентами збору, аналізу та звітності. Показано, що інтеграція NLP-методів із глибоким навчанням забезпечує підвищення точності класифікації. Розроблена методологія є універсальною основою для практичної реалізації моделі, імплементованої в третьому розділі.

## РОЗДІЛ 3. ІМПЛЕМЕНТАЦІЯ МОДЕЛЕЙ ТА МЕТОДІВ ДЛЯ КЛАСИФІКАЦІЇ ВІДГУКІВ КОРИСТУВАЧІВ ЗАСОБАМИ ОБРОБКИ ПРИРОДНОЇ МОВИ

### 3.1. Методології класифікації настроїв. Лексичний підхід та використання бібліотеки TextBlob

Для класифікації настроїв у текстових даних були застосовані два основні підходи: метод, заснований на попередньо визначеному лексиконі (Opinion Lexicon) та використання бібліотеки TextBlob.

#### 3.1.1. Класифікація на основі Opinion Lexicon

Лексикон у контексті NLP відноситься до спеціалізованого словникового запасу, що використовується для аналізу. Лексичний аналіз настроїв використовує попередньо визначений словник, у якому слова марковані як позитивні, негативні або нейтральні. Для визначення настрою тексту виконується його токенізація, після чого кожен токен зіставляється з наявними словами у лексиконі.

Для цього дослідження використовується Opinion Lexicon, створений Мінкінгом Ху та Бінгом Лю, що містить приблизно 6800 слів із позитивними та негативними конотаціями.

З Opinion Lexicon було виділено два списки:

- pos\_words: Містить приблизно 2006 позитивних слів (наприклад, "love", "happy", "excellent", "amazing", "beautiful").
- neg\_words: Містить приблизно 4783 негативних слова (наприклад, "hate", "sad", "terrible", "disgusting", "ugly").

Обчислення оцінки настрою `get_sentiment_score()` була розроблена для кількісної оцінки настрою тексту.

Кожен текст спершу токенізується на окремі речення за допомогою TreebankWord токенизатора. Для кожного речення підраховується кількість

позитивних і негативних слів. Оцінка настрою для кожного речення інкрементується на +1 за кожне позитивне слово та декрементується на -1 за кожне негативне слово.

Кінцева оцінка для тексту обчислюється як середнє арифметичне оцінок настрою всіх речень у тексті. Функція повертає значення з плаваючою комою у діапазоні [-1;1]. Позитивні значення вказують на позитивний настрій, негативні — на негативний, а значення, близькі до 0, позначають нейтральний настрій.

### 3.1.2. Класифікація за допомогою бібліотеки *TextBlob*

*TextBlob* є універсальною бібліотекою Python для NLP, що полегшує виконання широкого спектру завдань (токенізація, розмітка частин мови, аналіз настроїв). Бібліотека побудована на основі NLTK (Natural Language Toolkit). На відміну від простого лексичного підходу, *TextBlob* використовує внутрішні методи машинного навчання для аналізу настроїв.

Аналіз настроїв за допомогою *TextBlob* генерує два ключові показники:

1. Полярність (Polarity) - визначає орієнтацію настрою в діапазоні від -1 (повністю негативний) до 1 (повністю позитивний).
2. Суб'єктивність (Subjectivity) - визначає, наскільки текст виражає думку (суб'єктивний) або констатує факт (об'єктивний).

У межах даного проекту використовувалася лише метрика полярності для класифікації відгуків на позитивні, негативні або нейтральні.

Була створена функція обчислення полярності настрою `calculate_sentiment_polarity()`, яка використовує внутрішні механізми *TextBlob*:

- *TextBlob* токенізує текст на речення та слова.
- Обчислюється полярність настрою для кожного речення.

Оцінки полярності на рівні речень агрегуються для отримання загальної оцінки полярності для всього тексту. Результат є значенням з плаваючою комою в діапазоні [-1;1], де інтерпретація ідентична лексичному

підходу: негативні значення вказують на негативний настрій, позитивні — на позитивний, а 0 — на нейтральний.

## **3.2. Класифікація настроїв за допомогою моделі глибокого навчання на визначених мітках**

Для підвищення точності та узгодженості класифікації настроїв (позитивний, негативний, нейтральний) застосовується модель глибокого навчання, що навчається на попередньо визначених мітках. Архітектура цієї нейронної мережі ретельно структурована для ефективної обробки послідовних текстових даних.

### *3.2.1. Архітектура та компоненти моделі*

Архітектура моделі являє собою послідовну структуру, яка складається із вхідних шарів, прихованих шарів та вихідного шару, кожен з яких містить взаємопов'язані нейрони.

Шар вбудовування (Embedding Layer) є першим шаром, що перетворює цілочисельно кодовані токени слів на щільні числові вектори (embeddings). Цей процес дозволяє захоплювати семантичні властивості та взаємозв'язки між словами в тексті. Для ініціалізації цього шару використовується попередньо навчена матриця вбудовування (наприклад, на основі GloVe), що забезпечує використання семантичних знань, отриманих із великих корпусів.

Як рекурентний шар (LSTM Layer) використовується шар довгої короткочасної пам'яті (Long Short-Term Memory, LSTM) — різновид рекурентної нейронної мережі (RNN), добре адаптований для послідовних даних.

LSTM шар обробляє послідовності вхідних векторів, ефективно захоплюючи тимчасові залежності та довгострокові патерни в тексті, що є вирішальним для розуміння контексту. У цій моделі використовується один LSTM шар зі 128 одиницями.

Щільні та вихідні шари (Dense and Output Layers). Після LSTM додається щільний шар з 64 одиницями та функцією активації сигмоїда. Цей шар з'єднує всі нейрони попереднього шару та виконує нелінійне перетворення даних, адаптуючись до складних представлень тексту. Фінальний шар містить 3 одиниці (за кількістю класів: позитивний, негативний, нейтральний) та використовує функцію активації Softmax. Softmax генерує розподіл ймовірностей над класами, забезпечуючи, що сума прогнозованих ймовірностей дорівнює 1.

### *3.2.2. Етап навчання моделі та оптимізація*

Навчання моделі передбачає коригування ваг на основі міток класифікації (зіркових рейтингів).

Використовується Categorical Cross-entropy (категорійна крос-ентропія). Вона є стандартною для багатокласової класифікації, де цільові мітки кодуються за допомогою one-hot encoding. Функція обчислює втрати між істинними мітками та розподілом прогнозів.

Як оптимізатор використовується алгоритм Adam. Adam — це адаптивний алгоритм оптимізації швидкості навчання, який динамічно коригує швидкість навчання для кожної ваги, забезпечуючи швидшу та ефективнішу мінімізацію функції втрат.

Для токенизації налаштовується об'єкт `Tokenizer()` (з `Keras preprocessing.text`), який підганяється до попередньо обробленого тексту. Цей процес створює словник (`word_index`), що відображає слова на унікальні цілочисельні індекси. Об'єкт токенизатора зберігається (наприклад, за допомогою бібліотеки `Pickle`) для забезпечення консистентного відображення нових текстових даних на цілочисельні послідовності при розгортанні.

Щодо розподілу даних, то використовується функція `train_test_split()` (`scikit-learn`) для розділення даних на навчальний та тестовий набори, з розміром тестової вибірки 30% (0.3).

Для відповідності вимогам архітектур глибокого навчання щодо вхідних даних фіксованої довжини застосовується доповнення. Функція `sequence_padding()` перетворює токеновані послідовності на числові вектори (за допомогою `texts_to_sequences`) та виконує доповнення (`padding`) або обрізання (`truncating`) для уніфікації їхньої довжини.

Під час кодування міток, цільові мітки настроїв перетворюються на вектори One-Hot Encoding. Функція `read_glove_vector()` читає попередньо навчений файл GloVe, створюючи словник `word_to_vec_map`, де слова відображаються на їхні відповідні векторні представлення (масиви NumPy). Ініціалізується матриця вбудовування з розміром, що відповідає розміру словника та довжині вектора вбудовування. Кожне слово у вивченому словнику, яке має відповідний попередньо навчений вектор GloVe, використовується для заповнення відповідного рядка в матриці. Ця матриця використовується для ініціалізації шару вбудовування моделі.

### 3.2.3. Компіляція та оцінка

Модель створюється за допомогою послідовного API Keras і компілюється з трьома основними метриками:

- функція втрат: `categorical_crossentropy`.
- оптимізатор: Adam.
- метрика: Accuracy (точність), яка оцінює відсоток правильно класифікованих прикладів.

Навчання моделі виконується за допомогою методу `fit()`, використовуючи навчальний набір та мітки, закодовані за допомогою one-hot encoding.

## 3.3. Визначення оптимального порогу класифікації настроїв

У контексті лексичного аналізу настроїв, де оцінки Sentiment Score (Opinion Lexicon) та Polarity Score (TextBlob) знаходяться в діапазоні  $[-1;1]$ ,

визначення порогового значення для нейтрального настрою є критично важливим етапом. Цей поріг повинен максимально узгоджуватися з істинною міткою (Ground Truth) відгуку. Для емпіричного визначення оптимального порогу було проведено серію тестових випадків.

### 3.3.1. Емпіричне тестування порогових значень Opinion Lexicon

Було протестовано шість різних порогових значень для визначення нейтральної зони в класифікації, заснованій на Opinion Lexicon. Істинний настрій вважається нейтральним, якщо оцінка  $S$  задовольняє умові  $S_{min} \leq S \leq S_{max}$ .

Таблиця 3.1.

Результати тестування порогових значень Opinion Lexicon

Умова для нейтрального настрою (Neutral Range)	Точність (Accuracy)	Кількість позитивних	Кількість негативних	Кількість нейтральних
0 < Нейтральні > 0	64.38%	69055	18899	12008
0 < Нейтральні > 0.01	64.23%	68577	18899	12486
0 < Нейтральні > 0.02	63.21%	65539	18899	15524
0 < Нейтральні > 0.05	58.51%	52736	18899	28327
0 < Нейтральні > 0.1	46.82%	48270	18899	32793
-0.1 < Нейтральні > 0.1	39.14%	32793	3523	63646

Аналіз результатів у таблиці 3.1 демонструє, що точність класифікації значно залежить від розміру нейтрального порогового значення. Найвища точність (64.38%) була досягнута при встановленні найсуворішого порогу, де лише оцінки, точно рівні нулю ( $S=0$ ), класифікувалися як нейтральні. Кожне послідовне збільшення порогового значення призводить до суттєвого зниження точності. Це свідчить про те, що у випадках, коли Opinion Lexicon

присвоює ненульову оцінку, вона має більшу кореляцію з істинною полярністю, ніж з нейтральним станом.

### 3.3.2. Емпіричне тестування порогових значень TextBlob

Подібний набір тестових випадків було застосовано до класифікації на основі оцінки полярності TextBlob.

Таблиця 3.2.

Результати тестування порогових значень TextBlob

Умова для нейтрального настрою (Neutral Range)	Точність (Accuracy)	Кількість позитивних	Кількість негативних	Кількість нейтральних
0 < Нейтральні > 0	65.71%	79146	16123	4693
0 < Нейтральні > 0.01	65.62%	78068	16123	5771
0 < Нейтральні > 0.02	65.52%	76776	16123	7063
0 < Нейтральні > 0.05	64.46%	71806	16123	12033
0 < Нейтральні > 0.1	61.51%	61811	16123	22028
-0.1 < Нейтральні > 0.1	57.10%	61811	6936	31215

Результати, представлені в таблиці 3.2, також чітко вказують на те, що найвища точність (65.71%) для класифікації TextBlob досягається, коли нейтральний настрій обмежений випадком, де оцінка полярності точно дорівнює нулю ( $S=0$ ). Подібно до Opinion Lexicon, будь-яке розширення нейтральної зони призводить до зниження загальної точності класифікації моделі.

На основі емпіричного тестування для обох лексичних методів (Opinion Lexicon та TextBlob) оптимальний поріг для нейтрального настрою встановлюється, коли оцінка настрою або полярності точно дорівнює нулю. Це максимізує узгодженість класифікації з істинними мітками в даному наборі даних.

### 3.4. Оцінка продуктивності моделей класифікації настроїв

Оцінка продуктивності розроблених класифікаційних моделей (Opinion Lexicon, TextBlob та нейронна мережа) здійснювалася за допомогою матриці невідповідності (Confusion Matrix) та ключових метрик.

#### 3.4.1. Методологія оцінки на основі матриці невідповідності

Матриця невідповідності є табличним представленням, що візуалізує продуктивність алгоритму класифікації шляхом порівняння істинних (фактичних) та передбачених міток.

Для обчислення матриці використовувалася функція `confusion_matrix` з бібліотеки Scikit-learn.

Функція приймала істинні мітки настроїв (отримані із зіркових рейтингів Amazon) та передбачені мітки настроїв (отримані відповідно від Opinion Lexicon, TextBlob та навченої нейронної мережі).

У контексті бінарної класифікації, матриця складається з чотирьох основних компонентів, які, при розширенні до мультикласової задачі, оцінюються для кожного класу окремо:

- Істинно Позитивні (True Positives, TP): Випадки, коли клас був правильно передбачений як позитивний (для даного класу).
- Істинно Негативні (True Negatives, TN): Випадки, коли клас був правильно передбачений як негативний (для даного класу).
- Хибно Позитивні (False Positives, FP): Випадки, коли клас був неправильно передбачений як позитивний (Помилка I роду).
- Хибно Негативні (False Negatives, FN): Випадки, коли клас був неправильно передбачений як негативний (Помилка II роду).

У даному трикласовому завданні ("позитивний", "негативний", "нейтральний") діагональні елементи матриці відображають кількість правильно класифікованих випадків, тоді як позадіагональні елементи вказують на помилки класифікації.

### 3.4.2. Метрики продуктивності

Для кількісної оцінки якості кожної моделі були обчислені наступні ключові метрики.

Точність (Accuracy) - частка правильно класифікованих прикладів від загальної кількості.

Точність (Precision) - частка істинно позитивних результатів серед усіх позитивно класифікованих (позитивна прогностична цінність, PPV).

Повнота (Recall) - частка істинно позитивних результатів серед усіх фактичних позитивних (чутливість або Hit Rate, TPR).

F1-оцінка (F1-Score) - гармонійне середнє точності та повноти, що забезпечує збалансовану оцінку.

### 3.4.3. Результати оцінки моделей

Класифікація Opinion Lexicon (рис. 3.1).

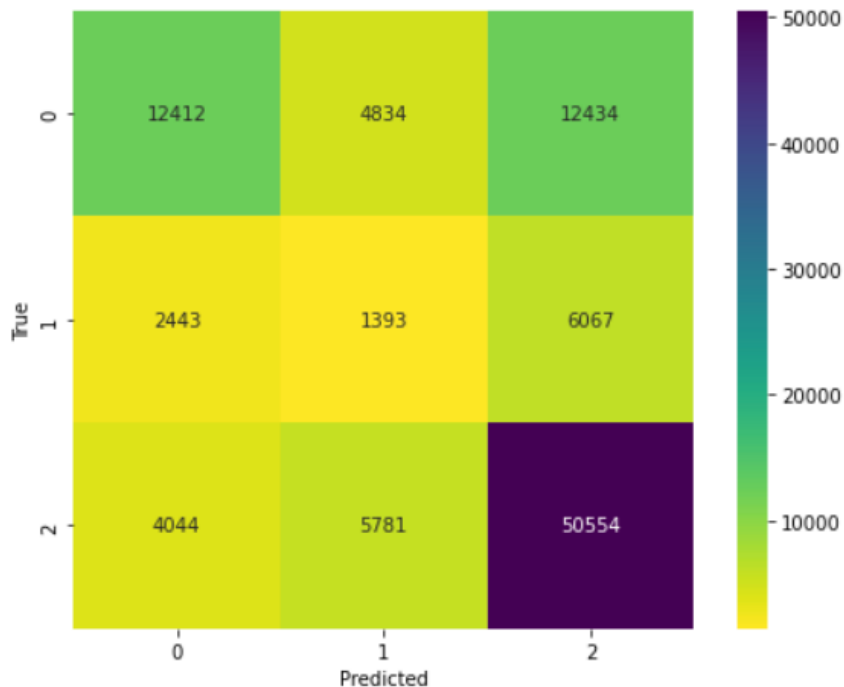


Рис. 3.1. Матриця невідповідності для класифікації Opinion Lexicon

Загальна точність класифікації: 64.38% (Загальна кількість відгуків N=99962).

## Результати оцінки для Opinion Lexicon

Категорія	TP	FN	FP	TN	Точність	Повнота	F1-оцінка
Негативні	12412	17268	6487	63795	65.67%	41.81%	51.10%
Нейтральні	1393	8510	10615	34671	11.60%	14.06%	12.71%
Позитивні	50554	9825	18501	21082	73.20%	83.72%	78.11%

Лексичний підхід демонструє відносно високу продуктивність для класів позитивний (F1: 78.11%) та негативний (F1: 51.10%), але суттєво низьку продуктивність для класу нейтральний (F1: 12.71%). Це вказує на те, що лексикон має обмежену здатність розпізнавати відсутність сильної полярності.

Розглянемо результати класифікації методом TextBlob (рис. 3.2)

Загальна точність класифікації: 65.71% (Загальна кількість відгуків N=99962).

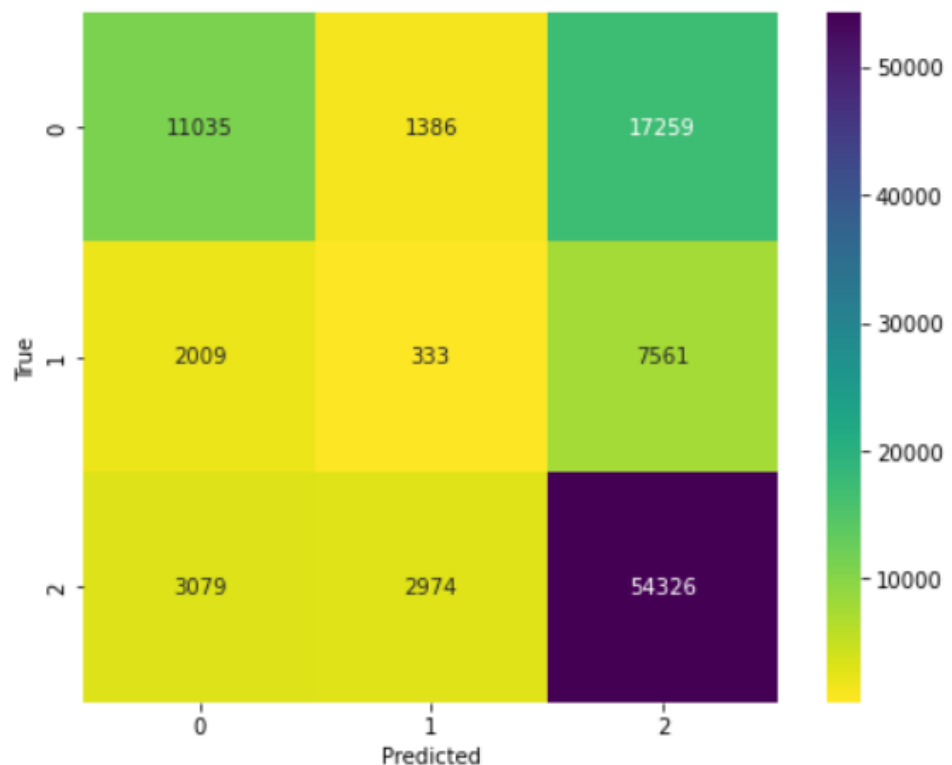


Рис. 3.2. Матриця невідповідності для класифікації TextBlob

Таблиця 3.4.

## Результати оцінки для TextBlob

Категорія	TP	FN	FP	TN	Точність	Повнота	F1-оцінка
Негативні	11035	18645	5088	65194	68.44%	37.17%	48.18%
Нейтральні	333	9570	4360	85699	7.09%	3.36%	4.56%
Позитивні	54326	6053	24820	14763	68.64%	89.97%	77.87%

TextBlob показав дещо вищу загальну точність (65.71%) порівняно з Opinion Lexicon. Однак, клас нейтральний знову демонструє найгірші результати (F1: 4.56%), що є загальною проблемою лексичних та простих статистичних підходів.

Розглянемо класифікацію з використанням нейронної мережі (рис. 3.3).

Загальна точність класифікації: 96.46% (Загальна кількість відгуків N=69973).

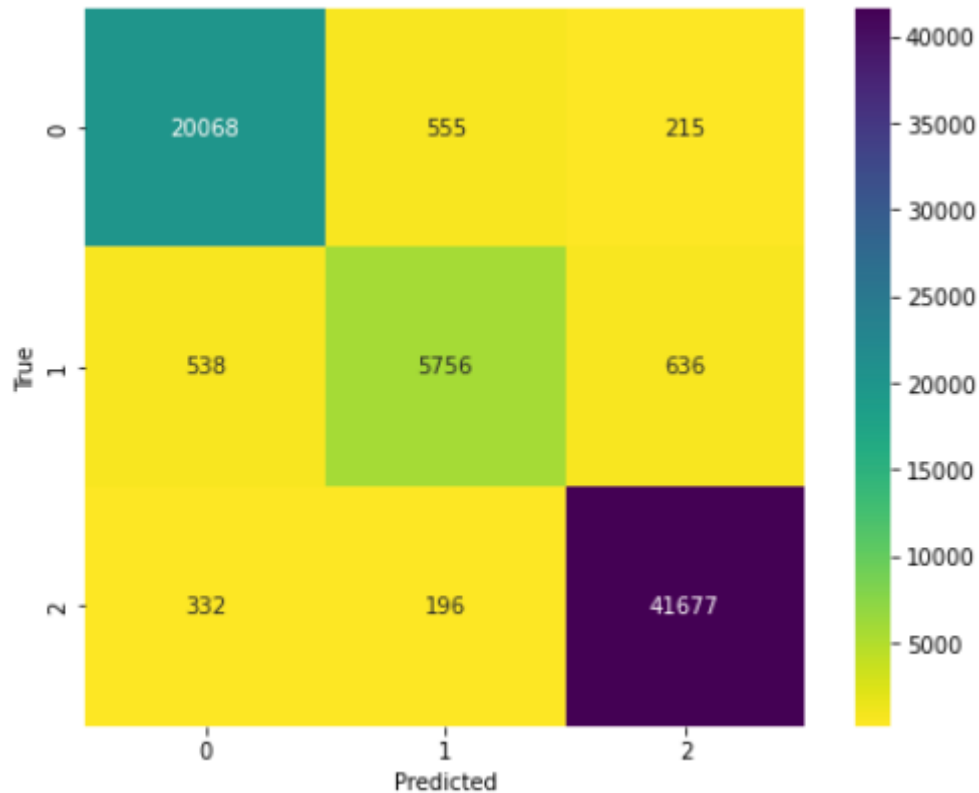


Рис. 3.3. Матриця невідповідності для класифікації з використанням нейронної мережі

Модель нейронної мережі (LSTM) продемонструвала значно вищу продуктивність із загальною точністю 96.46%. Всі класи, включаючи нейтральний (F1: 85.67%), показали високі значення F1-оцінки, що підтверджує ефективність глибокого навчання у захопленні складних контекстуальних залежностей для багатокласової класифікації настроїв.

### **3.5. Порівняльний аналіз методів класифікації настроїв у відгуках користувачів**

На основі емпіричних даних, отриманих під час оцінки трьох моделей класифікації настроїв:

1. Opinion Lexicon,
2. TextBlob.
3. Нейронна мережа / LSTM)

Відповідно було проведено детальний аналіз їхньої ефективності. Ключовим критерієм була здатність кожної моделі точно розрізнити позитивний, негативний та, особливо, нейтральний настрій.

Лексичні методи, незважаючи на свою простоту та швидкість, продемонстрували посередню загальну продуктивність у порівнянні з істинними мітками.

Opinion Lexicon досяг загальної точності 64.38%, тоді як TextBlob показав лише трохи кращий результат — 65.71%.

Найбільш критичним недоліком обох лексичних моделей стала їхня вкрай низька здатність класифікувати нейтральний настрій.

Opinion Lexicon досяг лише 12.71% за метрикою F1-оцінки для нейтрального класу.

TextBlob виявився ще менш ефективним, отримавши лише 4.56% F1-оцінки для нейтрального класу.

Така низька продуктивність зумовлена тим, що лексичні методи переважно ґрунтуються на підрахунку позитивних і негативних слів. Вони не

мають механізмів для розуміння контексту, іронії, заперечень або нюансів, які часто роблять відгук нейтральним (наприклад, "продукт працює, але нічого особливого"). Як наслідок, ці моделі або помилково відносять нейтральні тексти до полярних класів, або не можуть належним чином визначити їхню відсутність полярності.

Обидва методи показали прийнятну продуктивність у класифікації позитивних відгуків (F1-оцінки понад 77%), оскільки прямі позитивні формулювання легко розпізнаються за допомогою списків слів.

Модель, побудована на архітектурі довгої короткочасної пам'яті (LSTM), продемонструвала виняткову ефективність, значно перевершуючи обидва лексичні підходи.

Нейронна мережа досягла загальної точності 96.46%.

Продуктивність для позитивного та негативного класів була майже ідеальною, з F1-оцінками 98.37% та 96.12% відповідно. Це підтверджує, що модель глибокого навчання ефективно вивчила складні контекстуальні ознаки, пов'язані з сильною полярністю.

Найважливішим показником є успіх у класифікації нейтрального настрою, де модель отримала високу F1-оцінку 85.67%. Це досягнення демонструє, що LSTM-мережа здатна захоплювати послідовні залежності та контекстуальну інформацію, яка дозволяє їй точно ідентифікувати тексти, що не містять сильних емоційних висловлювань, на відміну від лексичних моделей.

Хоча модель LSTM має вищу обчислювальну вартість, її висока узгодженість та здатність долати складність мультикласової класифікації робить її найбільш надійним рішенням для цієї задачі.

На основі отриманих результатів, для досягнення найвищої точності та надійного розрізнення всіх трьох класів настроїв, рекомендовано застосовувати архітектуру глибокого навчання (LSTM). Висока продуктивність, зокрема в класифікації нейтральних відгуків, обґрунтовує використання LSTM як основної класифікаційної системи.

### 3.6. Порівняльний аналіз класифікаційних моделей на конкретних тестових випадках

Для подальшої валідації та розуміння поведінки розроблених класифікаційних систем було проведено порівняльний аналіз їхньої роботи на репрезентативних тестових прикладах. Розглядалися три основні моделі: Opinion Lexicon (точність 64.38%), TextBlob (точність 65.71%) та нейронна мережа (точність 96.46%).

Таблиця 3.5.

Тестові випадки позитивного настрою

№	Приклад змісту	Істинна Мітка	Opinion Lexicon	TextBlob	Нейронна Мережа
1	Опис мишки: "Я б купив це знову без вагань!"	Позитивний	Позитивний	Позитивний	Позитивний
2	Опис дисків: "Я виявив, що вони швидкі та надійні, і дуже компактні..."	Позитивний	Позитивний	Нейтральний	Позитивний
3	Змішана думка: "...це працює, добре це працює, але так, це повільно..."	Позитивний	Нейтральний	Негативний	Позитивний

Аналіз результатів позитивного класу.

Консистентна класифікація (рядок 1 в таблиці 3.5), у випадках однозначно позитивного тексту всі три моделі демонструють узгодженість, правильно класифікуючи настрій як "Позитивний".

У рядку 2 TextBlob помилково відносить відгук до "Нейтрального" класу. Це пояснюється тим, що його розрахунок полярності не перевищив встановлений поріг для позитивного результату, незважаючи на загальнопозитивний тон.

У рядку 3, яка містить змішану думку, Opinion Lexicon класифікує настрій як "Нейтральний" (через баланс позитивних і негативних слів, що

знижують оцінку до нуля), а TextBlob — як "Негативний" (імовірно, через внутрішнє врахування заперечень, наприклад, "don't buy").

В обох складних випадках, нейронна мережа правильно класифікувала істинний настрій як "Позитивний", демонструючи свою здатність інтерпретувати кінцевий позитивний намір, незважаючи на внутрішні протиріччя в тексті.

Таблиця 3.6.

Тестові випадки негативного настрою

№	Приклад змісту	Істинна Мітка	Opinion Lexicon	TextBlob	Нейронна Мережа
1	Опис невдалого продукту: "Я більше ніколи не буду використовувати цього продавця."	Негативний	Негативний	Негативний	Негативний
2	Компромісна думка: "Клавіатура... повністю функціональна... Однак клавіші ESC та F1-F7 мають погане фарбування..."	Негативний	Позитивний	Нейтральний	Негативний
3	Несправність: "Спочатку це працювало чудово... Але сьогодні він перестав працювати."	Негативний	Позитивний	Позитивний	Негативний

Проведемо аналіз результатів негативного класу.

Усі моделі успішно ідентифікували однозначно негативний настрій.

Лексичні моделі продемонстрували вразливість до наявності позитивних слів навіть у відгуках з негативним кінцевим результатом (наприклад, несправність). В рядку 2 таблиці 3.6 Opinion Lexicon класифікував відгук як "Позитивний", ймовірно, через наявність позитивних

слів, які переважили опис недоліків. TextBlob класифікував його як "Нейтральний".

В рядку 3, де несправність була описана після позитивного вступу ("Спочатку це працювало чудово"), обидва лексичні методи помилково видали "Позитивний" результат. Це чітко ілюструє, що ці моделі не враховують часову або причинно-наслідкову логіку тексту, а лише агрегують полярність слів.

Нейронна мережа в обох випадках правильно класифікувала "Негативний" настрій, оскільки вона здатна розпізнати, що кінцева думка, незважаючи на початкові переваги, є негативною через функціональні недоліки.

Таблиця 3.7.

Тестові випадки нейтрального настрою

№	Приклад змісту	Істинна Мітка	Opinion Lexicon	TextBlob	Нейронна Мережа
1	Опис зносу: "Купив його 5 років тому. Він завжди скрипить... Подушка все ще пристойна через 5 років."	Нейтральний	Нейтральний	Нейтральний	Нейтральний
2	Технічний опис проблеми: "ВИКОРИСТАНИЙ Seagate... швидкість знову падає... Мінус 2 зірки за випадкові незручності."	Нейтральний	Негативний	Негативний	Нейтральний
3	Змішані відчуття: "...вона дуже темного рожевого кольору... Миша працює ідеально... зовсім не зручна... моя рука затікає..."	Нейтральний	Позитивний	Позитивний	Нейтральний

Проведемо аналіз результатів нейтрального класу. Усі три моделі успішно класифікували як "Нейтральний" відгук, який описує довгострокове зношування без сильних емоційних висловлювань (рядок 1).

Лексичні моделі продемонстрували високу схильність до помилкової класифікації нейтральних відгуків. У рядку 2, де текст є технічним і містить критику, обидва лексичні методи класифікували його як "Негативний", оскільки негативні слова перевищили поріг. У рядку 3, де є чітко змішаною думкою з балансом позитивних і негативних аспектів, обидва лексичні методи помилково віднесли відгук до "Позитивного" класу, оскільки сума їхніх оцінок перевищила нульовий поріг.

Нейронна мережа в обох випадках зберегла високу стійкість, правильно класифікуючи ці складні, збалансовані тексти як "Нейтральні", підтверджуючи свою здатність до контекстуального аналізу.

Порівняльний аналіз підтверджує, що нейронна мережа є найбільш надійним і стійким методом класифікації. Лексичні моделі (Opinion Lexicon та TextBlob) є крихкими і схильні до помилок у випадках, що містять:

1. Змішану полярність: Наявність як позитивних, так і негативних аспектів.
2. Часову Динаміку: Зміна настрою в межах одного тексту (наприклад, "Спочатку добре, потім зламався").

Нейронна мережа, завдяки архітектурі LSTM, успішно долає ці лінгвістичні виклики, забезпечуючи високу точність у всіх трьох класах.

### **3.7. Реалізація дизайну користувацького інтерфейсу та взаємодії системи**

Інтерфейс користувача (UI) слугує критично важливою ланкою для взаємодії користувача з розробленою системою аналізу настроїв. UI реалізовано за допомогою фронтенд-фреймворку Flask. Головна сторінка програми (рис. 3.4) є центральним пунктом доступу до функціоналу системи.

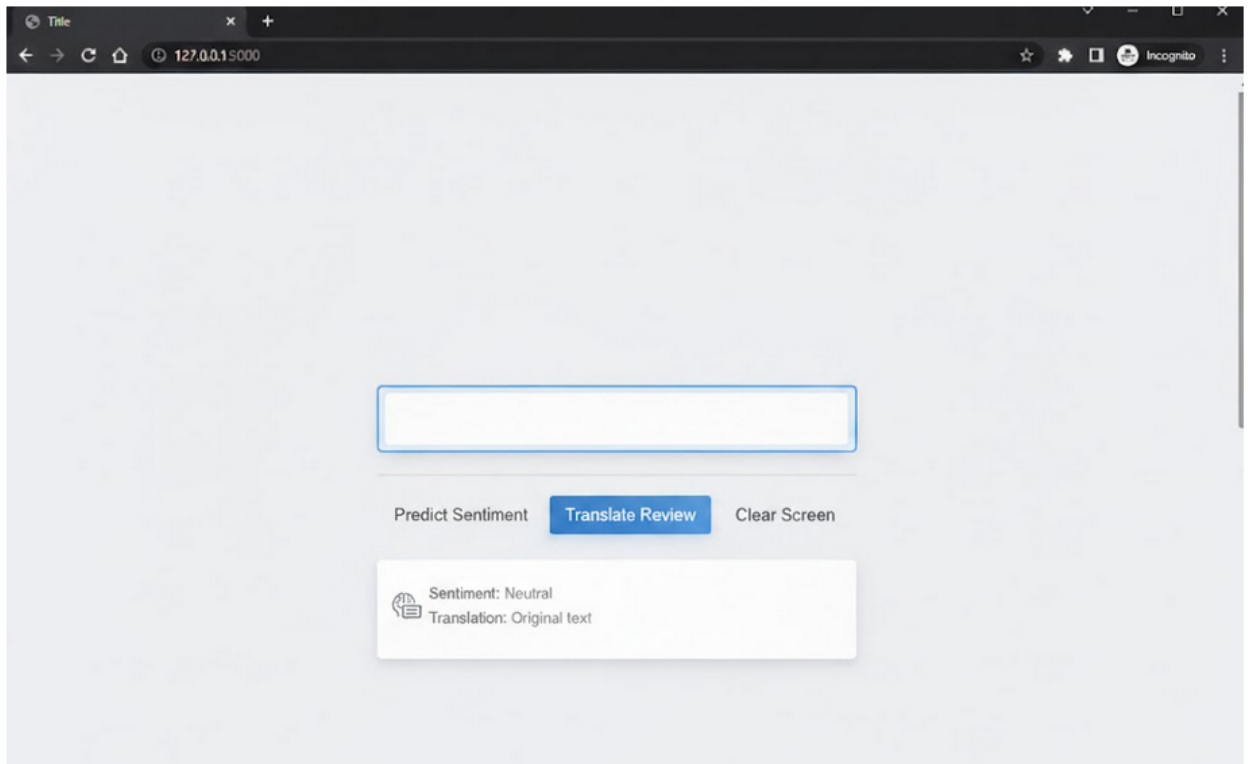


Рис. 3.4. Головна сторінка системи

Вона забезпечує наступні ключові елементи взаємодії:

- Поле введення тексту (Text Input Field) - основний елемент, призначений для введення користувачем довільного текстового фрагменту (наприклад, відгуку) для подальшого аналізу.

- Кнопка "Прогнозувати настрої" (Predict Sentiment) - тригер для запуску процесу класифікації настроїв у введеному тексті за допомогою моделі глибокого навчання.

- Кнопка "Перекласти відгук" (Translate Review) - тригер для ініціалізації комплексного процесу, що включає аналіз настрою та, за необхідності, переклад тексту.

- Кнопка "Очистити Екран" (Clear Screen) - функціонал для скидання відображених даних та підготовки інтерфейсу до нового циклу взаємодії.

Процес виводу прогнозу настрою (Sentiment Prediction Output) зображений на рис. 3.5 і демонструє безпосереднє використання класифікаційної моделі.

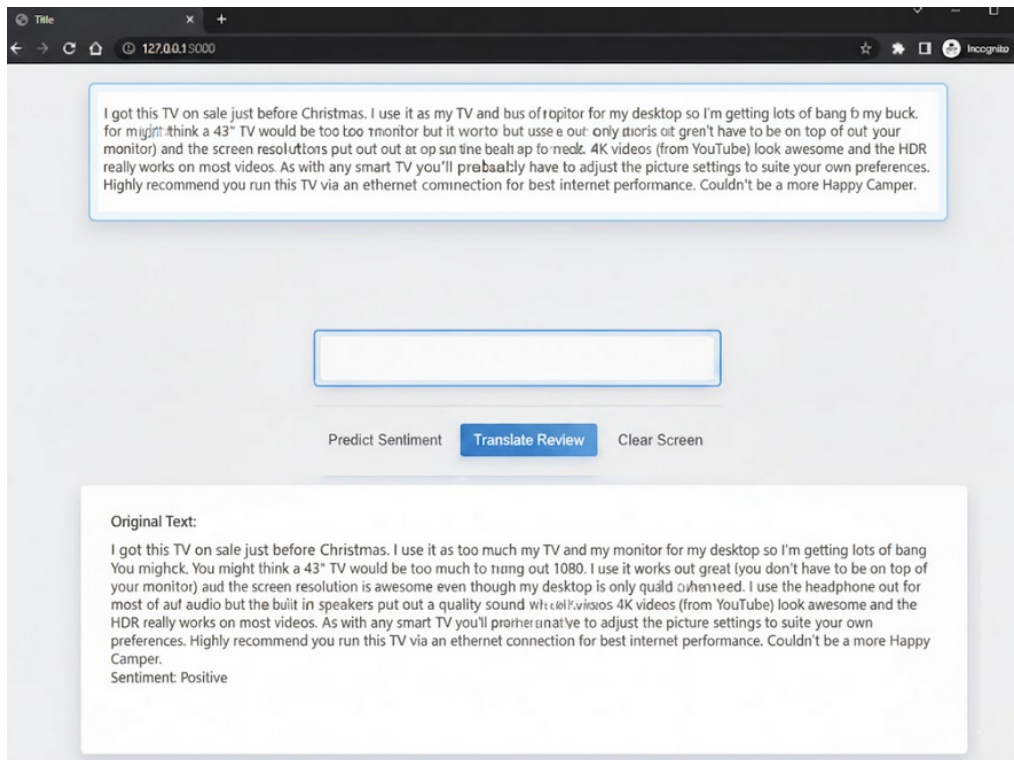


Рис. 3.5. Вивід перекладеного прогнозу настрою з введених користувачем даних

Після активації кнопки "Прогнозувати настрої", введений користувачем текст передається до бек-енд модуля для попередньої обробки та подачі до навченої нейронної мережі.

Система генерує та відображає передбачену мітку настрою (позитивний, негативний або нейтральний) на основі результатів роботи моделі. Це є основною функцією системи.

Вивід перекладеного прогнозу настрою (Translated Prediction Output) зображений на рисунку 3.6, описує розширений функціонал системи, який інтегрує мовний переклад.

Після активації кнопки "Перекласти відгук", система виконує два послідовні кроки:

1. Аналіз настрою: визначається настрої введеного тексту (як описано вище).
2. Переклад: Якщо вихідна мова тексту не відповідає цільовій мові виводу (або згідно з логікою системи), текст перекладається.

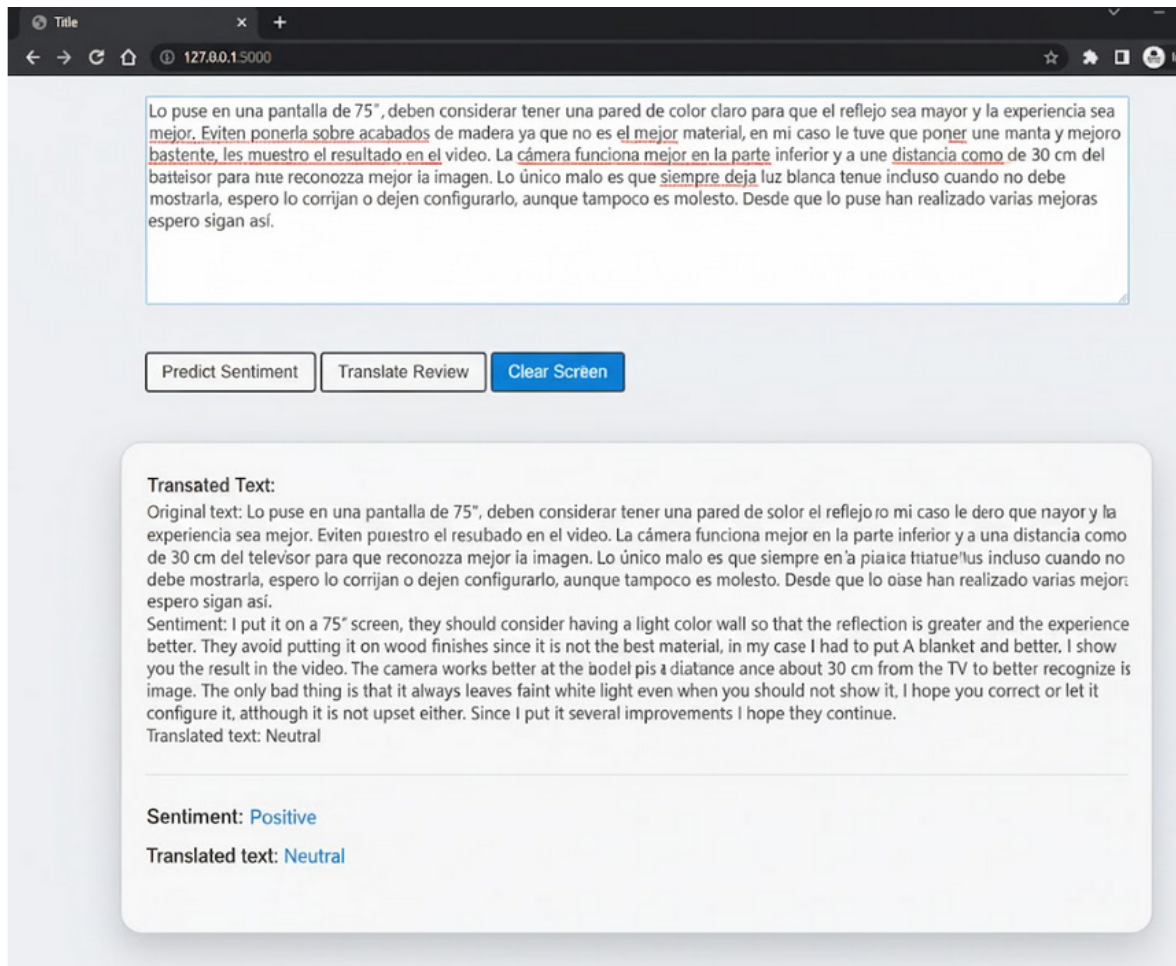


Рис. 3.6. Результат прогнозу настрою перекладеного тексту, наданого користувачем

Користувачеві надається подвійний результат: як сам перекладений текст, так і прогноз настрою, отриманий у результаті аналізу. Це забезпечує можливість аналізу відгуків, написаних різними мовами, підвищуючи універсальність системи.

У межах даного проєкту було здійснено дослідження та реалізацію комплексної системи класифікації настроїв (Sentiment Classification System) на основі текстових даних (відгуків клієнтів) з використанням методів обробки природної мови (NLP) та глибокого навчання (Deep Learning).

Було проведено порівняльне оцінювання трьох різних класифікаційних підходів: двох лексичних методів та однієї моделі глибокого навчання. Емпіричні результати чітко засвідчили, що модель нейронної мережі досягла

найвищої точності (96.46%), що значно перевершує продуктивність лексичних методів (Opinion Lexicon — 64.38%, TextBlob — 65.71%).

Висока продуктивність моделі глибокого навчання пояснюється її архітектурними перевагами:

1. Контекстуальний аналіз. Нейронна мережа, зокрема з рекурентними шарами (LSTM), має здатність до аналізу послідовних залежностей у тексті. Це дозволяє їй виходити за межі простого підрахунку окремих слів (як це роблять лексичні методи) та інтерпретувати значення токенів у контексті речення.

2. Точність класифікації. Ця здатність забезпечує точнішу класифікацію, особливо у випадках із запереченнями, іронією або збалансованою (нейтральною) полярністю, де лексичні підходи є вразливими.

Розроблена система має високу прикладну цінність для суб'єктів бізнесу, зокрема брендів та виробників:

- Система дозволяє автоматизувати процес аналізу великих обсягів (Big Data) клієнтських відгуків.

- Шляхом точного визначення настроїв, система допомагає ідентифікувати системні проблеми з продуктами, виявляти чинники задоволеності клієнтів та приймати обґрунтовані рішення щодо вдосконалення послуг. Це сприяє підвищенню якості продукції та лояльності споживачів.

Подальший розвиток системи може бути зосереджений на наступних аспектах:

- розширення мовної підтримки: додавання підтримки інших мов, окрім англійської, для аналізу глобального фідбеку.

- підвищення точності нейронної мережі шляхом інтеграції додаткових, більш репрезентативних даних та оптимізації архітектури моделі (наприклад, використання моделей-трансформерів).

- дослідження можливості інтеграції системи з платформами соціальних мереж для забезпечення аналізу настроїв у реальному часі, що дозволить брендам оперативно реагувати на зміни громадської думки.

### **Висновки до розділу**

У третьому розділі виконано практичну імплементацію моделей і методів класифікації відгуків користувачів засобами обробки природної мови та глибокого навчання. Реалізовано два основні підходи — лексичний аналіз із використанням бібліотек Opinion Lexicon і TextBlob та модель глибокого навчання, побудовану на нейронній архітектурі. Проведено навчання моделей на попередньо підготовленому корпусі відгуків Amazon із використанням методів очищення, лематизації та токенізації тексту. Здійснено тестування різних конфігурацій моделей для визначення оптимальних гіперпараметрів і порогових значень класифікації. Експериментальні результати продемонстрували, що модель глибокого навчання забезпечує вищі показники точності, повноти та F1-міри порівняно з лексичними методами. Проведено аналіз матриці невідповідностей, що дозволив оцінити характер помилок і визначити напрями вдосконалення моделі. Розроблено механізм візуалізації результатів класифікації у вигляді графічного інтерфейсу користувача, який спрощує інтерпретацію результатів. Здійснено порівняльне дослідження моделей за різними метриками продуктивності, що підтвердило перевагу глибоких нейронних мереж. Отримані результати свідчать про ефективність запропонованої архітектури для практичного застосування у системах аналітики відгуків.

## ВИСНОВКИ

У магістерській роботі проведено дослідження теоретичних, методичних і прикладних аспектів класифікації відгуків користувачів із використанням сучасних засобів обробки природної мови (Natural Language Processing, NLP) та моделей глибокого навчання. Робота спрямована на розв'язання актуальної науково-практичної проблеми — підвищення ефективності автоматизованого аналізу настроїв у текстових даних, що є ключовим елементом інтелектуальних систем підтримки прийняття рішень у сфері електронної комерції, маркетингу та клієнтської аналітики.

У першій частині дослідження здійснено глибокий аналіз теоретичних основ та сучасних тенденцій у сфері обробки природної мови, класифікації текстів і аналізу настроїв. Визначено, що класичні підходи — лексикон-орієнтовані методи та алгоритми машинного навчання — забезпечують базовий рівень точності, але не враховують складних контекстуальних залежностей, властивих природній мові. Натомість моделі глибокого навчання, зокрема архітектури на основі рекурентних нейронних мереж (LSTM), згорткових мереж (CNN) та трансформерів (BERT, RoBERTa, DistilBERT), продемонстрували суттєво вищу здатність до розуміння семантики текстів і тонких емоційних відтінків, що робить їх ефективним інструментом для аналізу відгуків.

У другій частині роботи розроблено методологію побудови класифікаторів настроїв для відгуків користувачів Amazon, що включає етапи збору, очищення, нормалізації та анотації текстових даних. Розроблена методика дозволила забезпечити структуровану підготовку корпусу текстів для подальшого машинного аналізу. Створено архітектуру системи аналізу настроїв, яка поєднує лексичний підхід (з використанням бібліотек Opinion Lexicon та TextBlob) і модель глибокого навчання, що дало змогу порівняти точність і продуктивність різних підходів у реальних умовах.

У третій частині виконано імплементацію розроблених методів класифікації, проведено навчання та тестування моделей на експериментальному наборі даних, а також здійснено кількісну оцінку результатів. Отримані експериментальні дані засвідчили, що модель глибокого навчання демонструє найвищі показники точності (accuracy), повноти (recall) та F1-міри порівняно з класичними підходами. Зокрема, встановлено, що використання трансформерних моделей дозволяє досягати більшої узагальнювальної здатності, стабільності та стійкості до шумових даних, що є особливо важливим у випадку великих і неструктурованих текстових масивів.

Окрім того, проведено порівняльний аналіз результатів класифікації між лексичними методами та глибокими моделями, який показав, що останні не лише забезпечують вищу якість класифікації, але й краще виявляють контекстно-залежні емоції. Розроблена система була доповнена модулем користувацького інтерфейсу, який забезпечує зручну візуалізацію результатів аналізу настроїв та інтерактивну взаємодію з користувачем.

Результати дослідження підтвердили доцільність інтеграції методів NLP із технологіями глибокого навчання у процесі автоматичної класифікації користувацьких відгуків. Наукова новизна роботи полягає у комплексному підході до розробки архітектури системи аналізу настроїв, яка поєднує різні рівні обробки природної мови — від попередньої лінгвістичної нормалізації до високорівневої нейронної семантичної інтерпретації. Практичне значення полягає у створенні прототипу системи, який може бути використаний для моніторингу споживчих настроїв, аналізу репутації продуктів та автоматизованого виявлення емоційних тенденцій у великих текстових масивах.

Підсумовуючи результати дослідження, можна зробити висновок, що поставлена у роботі мета — розроблення моделей і методів класифікації відгуків користувачів із використанням засобів обробки природної мови та глибокого навчання — досягнута повною мірою. Запропоновані підходи

довели свою ефективність у вирішенні завдань аналізу настроїв і можуть бути застосовані для побудови інтелектуальних аналітичних систем у сфері електронної комерції, соціальних медіа та сервісів зворотного зв'язку.

Подальші дослідження доцільно спрямувати на вдосконалення багатомовних моделей, розширення можливостей контекстного аналізу, інтеграцію мультимодальних джерел даних (текст, зображення, аудіо) та впровадження методів підкріпленого навчання для динамічної адаптації систем класифікації до змін у мовних тенденціях і поведінці користувачів.

## ПЕРЕЛІК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. An analysis of customer perception using lexicon-based sentiment analysis of Arabic Texts framework: Heliyon - [https://www.cell.com/heliyon/fulltext/S2405-8440\(24\)06351-5](https://www.cell.com/heliyon/fulltext/S2405-8440(24)06351-5)
2. A performance comparison of supervised machine learning models for Covid-19 tweets sentiment analysis | PLOS One - <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0245909>
3. A review on sentiment analysis and emotion detection from text | Social Network Analysis and Mining - <https://link.springer.com/article/10.1007/s13278-021-00776-6>
4. A Unified Generative Framework for Aspect-based Sentiment Analysis - ACL Anthology - <https://aclanthology.org/2021.acl-long.188/>
5. Sentiment Lexicons for Arabic Social Media / Saif M. Mohammad, Mohammad Salameh, Svetlana Kiritchenko. - <https://aclanthology.org/L16-1006.pdf>
6. International Sentiment Analysis for News and Blogs | Proceedings of the International AAAI Conference on Web and Social Media - <https://ojs.aaai.org/index.php/ICWSM/article/view/18606>
7. View of International Sentiment Analysis for News and Blogs. - <https://ojs.aaai.org/index.php/ICWSM/article/view/18606/18391>
8. [2201.08277] NaijaSenti: A Nigerian Twitter Sentiment Corpus for Multilingual Sentiment Analysis. - <https://arxiv.org/abs/2201.08277>
9. Zhu, L., Xu, M., Bao, Y., Xu, Y., & Kong, X. (2023). Deep learning for aspect-based sentiment analysis: a review. <https://doi.org/10.1016/j.knosys.2022.109356>
10. Sharma, N. A., Ali, A. B. M. S., & Kabir, M. A. (2024). A review of sentiment analysis: tasks, applications, and deep learning techniques. International Journal of Data Science and Analytics, <https://doi.org/10.1007/s41060-024-00594-x>

11. Dang, J., et al. (2023). Challenges and future in deep learning for sentiment analysis. *Artificial Intelligence Review*, <https://doi.org/10.1007/s10462-023-10651-9>
12. Mamani-Coaquira, Y., & Villanueva, E. (2024). A review on text sentiment analysis with machine learning and deep learning techniques. *IEEE Access*, <https://doi.org/10.1109/ACCESS.2024.3513321>
13. Hammouda, S., et al. (2022). Survey on sentiment analysis: evolution of research methods and topics. *Artificial Intelligence Review*, <https://doi.org/10.1007/s10462-022-10386-z>
14. Mamani-Coaquira, Y., & Villanueva, E. (2024). A Review on Text Sentiment Analysis With Machine Learning and Deep Learning Techniques. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2024.3513321>
15. Liu, B. (2012). *Sentiment analysis and opinion mining*. Synthesis Lectures on Human Language Technologies. Morgan & Claypool.
16. Alaparthi, S., & Mishra, M. (2020). Bidirectional Encoder Representations from Transformers (BERT): A sentiment analysis odyssey. arXiv preprint. arXiv:2007.01127
17. Gouthamia, G., & Hegde, M. (2021). A comprehensive survey on sentiment analysis techniques. *International Journal of Technology*, 300 pages.
18. Tsirmpas, D., Gkionis, I., Papadopoulos, G. Th., & Mademlis, I. (2023). Neural natural language processing for long texts: a survey on classification and summarization. arXiv preprint. arXiv:2305.16259
19. Vazan, M., Masoumi, F. S., & Saeedi Majd, S. (2022). A deep convolutional neural networks based multi-task ensemble model for aspect and polarity classification in Persian reviews.
20. Suryawanshi, N. S. (2024). Sentiment analysis with machine learning and deep learning: A survey of techniques and applications. *International Journal of Science and Research Archive*, 12(2), 005–015. <https://doi.org/10.30574/ijrsra.2024.12.2.1205>

21. Singh, N., & Kaur, H. (2023). A comparative study of sentiment analysis on customer reviews using deep learning and traditional machine learning techniques. *Computers*, 13(12), 340. <https://doi.org/10.3390/computers131>
22. Kumar, S., & Jaiswal, A. K. (2020). Soft computing techniques for sentiment analysis of Twitter data: A review. *International Journal of Fuzzy Systems*, 22(3).
23. Mohey El-Din Mohamed Hussein, D. (2016). A survey on sentiment analysis challenges. *Journal of King Saud University – Engineering Sciences*. <https://doi.org/10.1016/j.jksues.2016.04.002>
24. Mamani-Coaquira, Y., & Villanueva, E. (2024). A review of sentiment analysis: tasks, applications and deep learning techniques. *International Journal of Data Science and Analytics*. <https://doi.org/10.1007/s41060-024-00594-x>
25. “What is Sentiment Analysis?” (2023). IBM. Retrieved from <https://www.ibm.com/think/topics/sentiment-analysis>
26. “What is Sentiment Analysis?” (2023). AWS. Retrieved from <https://aws.amazon.com/what-is/sentiment-analysis/>
27. Yadav, K., & Vishwakarma, S. (2020). A review on sentiment analysis and emotion detection from text. *PLOS ONE*. <https://doi.org/10.1371/journal.pone.0271343>
28. Choi, C., & Lee, J. (2022). A deep learning-based model using hybrid feature extraction for sentiment analysis of consumer reviews. *Journal of Big Data*, 9(1), 80. <https://doi.org/10.1186/s40537-022-00680-6>
29. Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*, 2(1–2), 1–135.
30. Schuster, M., & Paliwal, K. (1997). Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing*, 45(11), 2673–2681.
31. Goldberg, Y. (2016). *A Primer on Neural Network Models for Natural Language Processing*. Morgan & Claypool.

32. Jurafsky, D., & Martin, J. H. (2023). *Speech and Language Processing* (4th ed.). Pearson.
33. McCulloch, W., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4), 115–133.
34. Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data Mining: Practical Machine Learning Tools and Techniques* (3rd ed.). Morgan Kaufmann.
35. Vaswani, A., et al. (2017). Attention is All You Need. *Advances in Neural Information Processing Systems (NeurIPS)*, 30, 5998–6008.
36. Liu, Y., et al. (2019). RoBERTa: A robustly optimized BERT pre-training approach. *arXiv preprint. arXiv:1907.11692*
37. Goldberg, R., & Shah, G. (2021). A survey on transformer architectures for NLP. *Journal of Artificial Intelligence Research*, 320 pp.
38. Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4), e1253. <https://doi.org/10.1002/widm.1253>
39. Do, T., & Nguyen, H. (2020). Toward effective aspect-based sentiment analysis: A survey. *IEEE Transactions on Knowledge and Data Engineering*, 32(8), 1597–1618.
40. Liu, H., & Yu, Z. (2021). Cross-lingual sentiment analysis: A survey. *Information Processing & Management*, 58(3), 102590.

## **ДОДАТКИ**

## Додаток А

### Фрагменти програмних кодів

```
import nltk
from nltk.tokenize import TreebankWordTokenizer
from nltk.tokenize import sent_tokenize
from nltk.corpus import stopwords, opinion_lexicon

from ipywidgets import interact, interact_manual, fixed
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.metrics import confusion_matrix, classification_report, accuracy_score, precision_score, recall_score, f1_score
from textblob import TextBlob
nltk.download('stopwords')
nltk.download("opinion_lexicon")

import string
from string import punctuation
import contractions

from tqdm import tqdm
tqdm.pandas()

pd.set_option('display.max_columns', None)
pd.set_option('display.max_colwidth', None)
```

```
from tensorflow.keras.preprocessing import sequence
from tensorflow.keras.layers import Dense,Dropout,Embedding,LSTM, Bidirectional, BatchNormalization,Masking,Activation,SimpleRNN
from tensorflow.keras.callbacks import EarlyStopping,ModelCheckpoint
from tensorflow.keras.losses import categorical_crossentropy
from tensorflow.keras.optimizers import Adam, SGD
from tensorflow.keras.models import Sequential
from tensorflow.keras.preprocessing.text import Tokenizer
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder
from keras.utils import np_utils
import tensorflow
import warnings
warnings.filterwarnings('ignore')
```

```
#read the data

reviews = pd.read_csv("bigdata.csv",low_memory=False)

#keep only targeted columns
col = ['overall','reviewText']
df = reviews[col].copy()

df.dropna(inplace=True)
```

```
def create_model(vocab_size,embed_vector_len,max_sentence_length,emb_matrix):
    model = Sequential()
    model.add(Embedding(input_dim=vocab_size, output_dim=embed_vector_len,weights = [emb_matrix],
                       input_length=max_sentence_length, trainable=False))
    model.add(LSTM(128, return_sequences=False))
    model.add(Dense(32, input_shape=Y_oh_train.shape, activation='sigmoid'))
    model.add(Dense(32, activation='sigmoid'))
    model.add(Dense(3, activation='sigmoid'))
    model.compile(loss='categorical_crossentropy',
                  optimizer='adam',
                  metrics=['accuracy'])
    return model
```



```

#process the Label data
train_labels = np.array(Y_train)
test_labels = np.array(y_test)

#Convert the Labels to One hot encoding vector for softmax for neural network

num_labels = len(np.unique(test_labels))
Y_oh_train = np_utils.to_categorical(train_labels, num_labels)
y_oh_test = np_utils.to_categorical(test_labels, num_labels)

print(f"Shape for Train Labels: {Y_oh_train.shape}\n")
print(f"Shape for test Labels: {y_oh_test.shape}")

# 0: ['1', '0', '0']
# 1: [0, 1, 0]
# 2: [0, 0, 1]

```

Shape for Train Labels: (69973, 3)

Shape for test Labels: (29989, 3)

```

# Create callbacks
callbacks = [EarlyStopping(monitor='val_loss', patience=5),
             ModelCheckpoint('model.h5', save_best_only=True,
                             save_weights_only=False)]

```

```

history = model.fit(train_set, Y_oh_train, epochs=30, batch_size=128, verbose=1, validation_split=0.2, callbacks=callbacks)

```

```

Epoch 1/30
438/438 [=====] - 12s 23ms/step - loss: 0.9051 - accuracy: 0.5951 - val_loss: 0.8949 - val_accuracy: 0.6066
Epoch 2/30
438/438 [=====] - 10s 22ms/step - loss: 0.8951 - accuracy: 0.6023 - val_loss: 0.8956 - val_accuracy: 0.6066
Epoch 3/30
438/438 [=====] - 10s 22ms/step - loss: 0.8952 - accuracy: 0.6022 - val_loss: 0.8963 - val_accuracy: 0.6066
Epoch 4/30
438/438 [=====] - 10s 22ms/step - loss: 0.8950 - accuracy: 0.6023 - val_loss: 0.8955 - val_accuracy: 0.6066
Epoch 5/30
438/438 [=====] - 10s 23ms/step - loss: 0.8943 - accuracy: 0.6029 - val_loss: 0.8952 - val_accuracy: 0.6069
Epoch 6/30
438/438 [=====] - 10s 23ms/step - loss: 0.8935 - accuracy: 0.6040 - val_loss: 0.8933 - val_accuracy: 0.6081
Epoch 7/30
438/438 [=====] - 10s 24ms/step - loss: 0.8931 - accuracy: 0.6042 - val_loss: 0.8910 - val_accuracy: 0.6066
Epoch 8/30
438/438 [=====] - 11s 24ms/step - loss: 0.8905 - accuracy: 0.6082 - val_loss: 0.8936 - val_accuracy: 0.6089
Epoch 9/30
438/438 [=====] - 11s 24ms/step - loss: 0.8944 - accuracy: 0.6033 - val_loss: 0.8936 - val_accuracy: 0.6067
Epoch 10/30
438/438 [=====] - 11s 24ms/step - loss: 0.8929 - accuracy: 0.6035 - val_loss: 0.8924 - val_accuracy: 0.6085
Epoch 11/30
438/438 [=====] - 11s 25ms/step - loss: 0.8898 - accuracy: 0.6044 - val_loss: 0.9190 - val_accuracy: 0.6088
Epoch 12/30
438/438 [=====] - 11s 24ms/step - loss: 0.6989 - accuracy: 0.7294 - val_loss: 0.5954 - val_accuracy: 0.7800
Epoch 13/30
438/438 [=====] - 11s 25ms/step - loss: 0.5373 - accuracy: 0.8090 - val_loss: 0.5424 - val_accuracy: 0.7936
Epoch 14/30
438/438 [=====] - 11s 25ms/step - loss: 0.4458 - accuracy: 0.8418 - val_loss: 0.4280 - val_accuracy: 0.8462
Epoch 15/30
438/438 [=====] - 11s 25ms/step - loss: 0.3656 - accuracy: 0.8713 - val_loss: 0.3745 - val_accuracy: 0.8712
Epoch 16/30
438/438 [=====] - 11s 25ms/step - loss: 0.2987 - accuracy: 0.9023 - val_loss: 0.3237 - val_accuracy: 0.8975
Epoch 17/30
438/438 [=====] - 11s 26ms/step - loss: 0.2458 - accuracy: 0.9278 - val_loss: 0.3099 - val_accuracy: 0.9052
Epoch 18/30
438/438 [=====] - 11s 25ms/step - loss: 0.2113 - accuracy: 0.9420 - val_loss: 0.2903 - val_accuracy: 0.9170
Epoch 19/30
438/438 [=====] - 11s 25ms/step - loss: 0.1855 - accuracy: 0.9532 - val_loss: 0.2781 - val_accuracy: 0.9221
Epoch 20/30

```