

## КВАЛІФІКАЦІЙНА РОБОТА

на тему Метод аналізу тональності відгуків клієнтів з використанням машинного навчання

Рівень вищої освіти другий (магістерський)

Галузь знань 12 – Інформаційні технології

Спеціальність 122 – Комп'ютерні науки  
Шифр і найменування

Освітня програма Комп'ютерні науки  
Код і найменування  
Назва

Виконав: студент 2 курсу, група КНм-24-1  
Курс, група виконавця Богдан ЗАГУРСЬКИЙ  
Підпис Ім'я, ПРІЗВИЩЕ

Керівник: д.т.н., професор кафедри КН  
Науковий ступінь, посада Едуард МАНЗЮК  
Підпис Ім'я, ПРІЗВИЩЕ

Нормоконтроль: к.т.н., доцент кафедри КН  
Науковий ступінь, посада Руслан БАГРІЙ  
Підпис Ім'я, ПРІЗВИЩЕ

До захисту допускаю:

Зав. кафедри КН, д.т.н., професор  
Підпис Олександр БАРМАК  
Ім'я, ПРІЗВИЩЕ

15 грудня 2025 р.

ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

Факультет інформаційних технологій

Кафедра комп'ютерних наук

Освітній ступінь магістр

Галузь знань 12 – Інформаційні технології

Спеціальність 122 – Комп'ютерні науки

ЗАТВЕРДЖУЮ

Завідувач кафедри комп'ютерних наук

(підпис)

д.т.н., професор Олександр БАРМАК


« 28 » 08 2025 року

**ЗАВДАННЯ  
НА КВАЛІФІКАЦІЙНУ РОБОТУ**

1. Тема кваліфікаційної роботи: «Метод аналізу тональності відгуків клієнтів з використанням машинного навчання»
2. Завдання видано студенту Богдану ЗАГУРСЬКОМУ  
(Ім'я, ПРІЗВИЩЕ)
3. Керівник роботи професор кафедри КН Едуард МАНЗЮК  
(Ім'я, ПРІЗВИЩЕ)
4. Затверджені наказом університету від « 25 » 08 2025 р. № 65
5. Дата видачі завдання студенту: « 28 » 08 2025 р.
6. Зміст пояснювальної записки (перелік задач) та вихідні дані:  
Метою роботи є підвищення точності автоматичної класифікації тональності клієнтських відгуків через розробку гібридного методу, що базується на інтеграції різнотипних алгоритмів машинного навчання. Задачі: здійснити огляд наявних методологій та технологічних рішень у сфері автоматичного визначення тональності текстових даних; розробити метод класифікації емоційного забарвлення відгуків, який використовує послідовну інтеграцію; розробити модифікацію базового методу шляхом включення аналізу біграм; створити програмну реалізацію запропонованого методу; провести експериментальне дослідження розробленого методу.

7. Календарний план виконання кваліфікаційної роботи:

№	Назва етапів (розділів) кваліфікаційної роботи магістра	Термін виконання	Примітка
1	Вибір напрямку дослідження та узгодження теми кваліфікаційної роботи з керівником, складання календарного графіка виконання роботи	вересень 2025	Виконано
2	Ознайомлення з предметною областю, аналіз існуючих методів і моделей, формулювання мети та завдань дослідження, визначення об'єкта й предмета дослідження	вересень 2025	Виконано
3	Розробка методу чи моделі для вирішення обраного завдання, опис архітектури рішення	жовтень 2025	Виконано
4	Програмна реалізація методу чи моделі	жовтень 2025	Виконано
5	Дослідження ефективності та експериментальна перевірка результатів, порівняння з відомими підходами	листопад 2025	Виконано
6	Написання пояснювальної записки, оформлення відповідно до вимог, врахування зауважень керівника	листопад 2025	Виконано
7	Підготовка презентаційних матеріалів та попередній захист	листопад 2025	Виконано
8	Перевірка пояснювальної записки на відповідність вимогам оформлення (нормоконтроль) та перевірка на академічну доброчесність. Отримання відгуку керівника та рецензії.	грудень 2025	Виконано
9	Публічний захист кваліфікаційної роботи	грудень 2025	Виконано

Виконавець: студент групи КНм-24-1  Богдан ЗАГУРСЬКИЙ  
Група виконавця Підпис Ім'я, ПРІЗВИЩЕ

Керівник: д.т.н., проф. каф. КН  Едуард МАНЗЮК  
Науковий ступінь, посада Підпис Ім'я, ПРІЗВИЩЕ

## Реферат

Кваліфікаційна робота магістра полягає у створенні методу для автоматичного визначення емоційного забарвлення клієнтських відгуків із застосуванням технологій машинного навчання.

**Актуальність теми.** Практичне значення дослідження обумовлене зростаючою потребою компаній у швидкій та об'єктивній обробці великих масивів текстових відгуків від споживачів. Традиційний підхід до аналізу, який передбачає ручну обробку коментарів спеціалістами, характеризується низькою швидкістю виконання, високою вартістю, схильністю до суб'єктивних оцінок і неможливістю масштабування для великих підприємств.

Сучасні досягнення у сфері обробки природної мови та машинного навчання відкривають нові можливості для створення автоматизованих систем, здатних з високою точністю визначати настрій текстових повідомлень. Впровадження таких систем дозволяє організаціям оперативно виявляти проблемні аспекти своїх продуктів чи послуг, своєчасно реагувати на негативні тенденції, покращувати якість обслуговування клієнтів та підвищувати загальну конкурентоспроможність на ринку. Використання гібридних підходів, які комбінують традиційні методи машинного навчання з елементами глибоких нейронних мереж, дає змогу досягти покращеної якості. Це робить такі рішення придатними для впровадження в реальних бізнес-середовищах з обмеженими ресурсами. Вказані чинники підкреслюють актуальність розробки як для практичного застосування в галузі обслуговування клієнтів, так і для розвитку наукових досліджень у напрямку застосування інтелектуальних технологій до аналізу текстових даних.

**Мета роботи.** Метою роботи є підвищення точності автоматичної класифікації тональності клієнтських відгуків через розробку гібридного методу, що базується на інтеграції різнотипних алгоритмів машинного навчання.

### **Задачі дослідження:**

– здійснити огляд наявних методологій та технологічних рішень у сфері автоматичного визначення тональності текстових даних, включаючи традиційні статистичні підходи та сучасні нейромережеві архітектури;

– розробити метод класифікації емоційного забарвлення відгуків, який використовує послідовну інтеграцію для досягнення підвищеної точності розпізнавання;

– розробити модифікацію базового методу шляхом включення аналізу біграм на етапі формування текстових ознак, що дозволить враховувати локальний контекст слів та покращити розпізнавання складних лінгвістичних конструкцій;

– створити програмну реалізацію запропонованого методу з використанням модульної архітектури, що забезпечить можливість незалежного розвитку окремих компонентів системи;

– провести експериментальне дослідження розробленого методу на репрезентативному наборі реальних клієнтських відгуків з оцінкою ключових метрик якості класифікації.

**Об'єкт дослідження** – процес автоматизованого визначення емоційного забарвлення у текстових відгуках споживачів.

**Предмет дослідження** – архітектури, алгоритми та програмні засоби для класифікації тональності текстів із використанням комбінованих методів машинного навчання.

**Методи дослідження.** Використано алгоритми векторизації текстових даних, метод опорних векторів, архітектури повнозв'язних нейронних мереж, технології попередньої обробки природномовних текстів, експериментальне тестування на валідованих датасетах.

**Наукова новизна отриманих результатів.** Удосконалено метод визначення тональності клієнтських відгуків унаслідок застосування послідовної архітектури з повнозв'язною нейронною мережею, що використовує комбіновані вхідні дані, який відрізняється від наявних використанням як векторизованих текстових ознак із біграмами, так і ймовірнісних оцінок від попередньо навченого класифікатора на

основі методу опорних векторів, що дало змогу підвищити точність визначення емоційного забарвлення текстів.

**Апробація результатів кваліфікаційної роботи магістра та публікації.**

Загурський Б.З., Манзюк Е.А., Скрипник Т.К. Метод аналізу тональності відгуків клієнтів з використанням гібридної моделі машинного навчання. Збірник наукових праць за матеріалами XVII Всеукраїнської науково-практичної конференції «Актуальні проблеми комп'ютерних наук АПКН-2025». - Хмельницький, 2025. - С.118 – 123.

**Структура та обсяг роботи.** Робота складається зі вступного розділу, чотирьох основних розділів, заключних висновків, переліку використаних літературних джерел та додаткових матеріалів. Загальний обсяг основного тексту становить близько 83 сторінки, 19 рисунків, 2 таблиць та 45 посилань.

**Ключові терміни:** аналіз тональності, відгуки клієнтів, метод опорних векторів, нейронні мережі, текстова класифікація, обробка природної мови, машинне навчання, гібридні моделі.

## Зміст

Перелік скорочень .....	4
Вступ.....	5
Розділ 1 Огляд методів аналізу тональності відгуків клієнтів .....	8
1.1 Характеристика задачі аналізу тональності відгуків клієнтів.....	8
1.2 Аналіз відомих досліджень та наукових підходів .....	10
1.3 Огляд архітектур, методів машинного навчання для аналізу тональності .....	13
1.4 Мета та постановка задачі.....	15
Розділ 2 Метод аналізу тональності відгуків клієнтів та критерії його оцінювання .	17
2.1 Концепція та схема методу аналізу тональності .....	17
2.2 Архітектура гібридної моделі класифікації .....	19
2.3 Модифікація моделі та покращення точності класифікації .....	23
2.4 Формування та підготовка навчальних даних .....	25
2.5 Критерії та метрики оцінювання роботи методу.....	29
Висновок до розділу 2 .....	35
Розділ 3 Програмна реалізація методу аналізу тональності відгуків .....	37
3.1 Загальна архітектура програмної системи.....	37
3.2 Реалізація модуля попередньої обробки текстів.....	40
3.3 Реалізація модуля класифікації .....	43
3.4 Реалізація модуля векторизації тексту.....	46
3.5 Взаємодія компонентів системи у процесі класифікації .....	50
Висновки до розділу 3 .....	52
Розділ 4 Експериментальні дослідження методу аналізу тональності відгуків .....	54
4.1 Формування набору даних для експериментальних досліджень .....	54
4.2 Налаштування параметрів базової моделі SVM .....	56
4.3 Процес навчання Dense Neural Network .....	59
4.4 Аналіз результатів класифікації Dense Neural Network .....	63
4.5 Оцінка якості гібридної моделі класифікації .....	67
4.6 Порівняльний аналіз результатів експериментів.....	70
Висновки до розділу 4 .....	74

Загальні висновки.....	75
Перелік посилань.....	76
Додатки	

**Перелік скорочень**

<b>Скорочення, термін, позначення</b>	<b>Пояснення</b>
SVM	Support Vector Machine (Метод опорних векторів)
RBF	Radial Basis Function (Радіально-базисна функція)
NN	Neural Network (Нейронна мережа)
TF-IDF	Term Frequency – Inverse Document Frequency
Dense NN	Dense Neural Network (Повнозв'язна нейронна мережа)
NLP	Natural Language Processing (Обробка природної мови)
ReLU	Rectified Linear Unit (Випрямлена лінійна одиниця)
F1-Score	Гармонійне середнє precision та recall

## Вступ

Кваліфікаційна робота магістра полягає у створенні методу для автоматичного визначення емоційного забарвлення клієнтських відгуків із застосуванням технологій машинного навчання.

**Актуальність теми.** Практичне значення дослідження обумовлене зростаючою потребою компаній у швидкій та об'єктивній обробці великих масивів текстових відгуків від споживачів. Традиційний підхід до аналізу, який передбачає ручну обробку коментарів спеціалістами, характеризується низькою швидкістю виконання, високою вартістю, схильністю до суб'єктивних оцінок і неможливістю масштабування для великих підприємств. Сучасні досягнення у сфері обробки природної мови та машинного навчання відкривають нові можливості для створення автоматизованих систем, здатних з високою точністю визначати настрої текстових повідомлень.

Впровадження таких систем дозволяє організаціям оперативно виявляти проблемні аспекти своїх продуктів чи послуг, своєчасно реагувати на негативні тенденції, покращувати якість обслуговування клієнтів та підвищувати загальну конкурентоспроможність на ринку. Використання гібридних підходів, які комбінують традиційні методи машинного навчання з елементами глибоких нейронних мереж, дає змогу досягти покращеної якості.

Це робить такі рішення придатними для впровадження в реальних бізнес-середовищах з обмеженими ресурсами. Вказані чинники підкреслюють актуальність розробки як для практичного застосування в галузі обслуговування клієнтів, так і для розвитку наукових досліджень у напрямку застосування інтелектуальних технологій до аналізу текстових даних.

**Мета роботи.** Метою роботи є підвищення точності автоматичної класифікації тональності клієнтських відгуків через розробку гібридного методу, що базується на інтеграції різнотипних алгоритмів машинного навчання.

**Задачі дослідження:**

– здійснити огляд наявних методологій та технологічних рішень у сфері автоматичного визначення тональності текстових даних, включаючи традиційні статистичні підходи та сучасні нейромережеві архітектури;

– розробити метод класифікації емоційного забарвлення відгуків, який використовує послідовну інтеграцію для досягнення підвищеної точності розпізнавання;

– розробити модифікацію базового методу шляхом включення аналізу біграм на етапі формування текстових ознак, що дозволить враховувати локальний контекст слів та покращити розпізнавання складних лінгвістичних конструкцій;

– створити програмну реалізацію запропонованого методу з використанням модульної архітектури, що забезпечить можливість незалежного розвитку окремих компонентів системи;

– провести експериментальне дослідження розробленого методу на репрезентативному наборі реальних клієнтських відгуків з оцінкою ключових метрик якості класифікації.

**Об'єкт дослідження** – процес автоматизованого визначення емоційного забарвлення у текстових відгуках споживачів.

**Предмет дослідження** – архітектури, алгоритми та програмні засоби для класифікації тональності текстів із використанням комбінованих методів машинного навчання.

**Методи дослідження.** Використано алгоритми векторизації текстових даних, метод опорних векторів, архітектури повнозв'язних нейронних мереж, технології попередньої обробки природномовних текстів, експериментальне тестування на валідованих датасетах.

**Наукова новизна отриманих результатів.** Удосконалено метод визначення тональності клієнтських відгуків унаслідок застосування послідовної архітектури з повнозв'язною нейронною мережею, що використовує комбіновані вхідні дані, який відрізняється від наявних використанням як векторизованих текстових ознак із біграмами, так і ймовірнісних оцінок від попередньо навченого класифікатора на

основі методу опорних векторів, що дало змогу підвищити точність визначення емоційного забарвлення текстів.

**Апробація результатів кваліфікаційної роботи магістра та публікації.**

Загурський Б.З., Манзюк Е.А., Скрипник Т.К. Метод аналізу тональності відгуків клієнтів з використанням гібридної моделі машинного навчання. Збірник наукових праць за матеріалами XVII Всеукраїнської науково-практичної конференції «Актуальні проблеми комп'ютерних наук АПКН-2025». - Хмельницький, 2025. - С.118 – 123.

**Структура та обсяг роботи.** Робота складається зі вступного розділу, чотирьох основних розділів, заключних висновків, переліку використаних літературних джерел та додаткових матеріалів. Загальний обсяг основного тексту становить близько 83 сторінки, 19 рисунків, 2 таблиць та 45 посилань.

## **Розділ 1 Огляд методів аналізу тональності відгуків клієнтів**

### **1.1 Характеристика задачі аналізу тональності відгуків клієнтів**

Аналіз тональності відгуків клієнтів становить одне з найважливіших завдань у сфері обробки мови, що дає змогу автоматично визначати емоційну оцінку текстових повідомлень. Ця задача набуває критичної важливості в контексті електронної комерції, де відгуки користувачів на різноманітних платформах безпосередньо впливають на купівельні рішення та формують бізнес-стратегії компаній. Основна складність полягає в тому, що відгуки часто містять суб'єктивні судження, іронію, сарказм, культурно-специфічні вирази та різноманітні шумові елементи, що значно ускладнює процес автоматизованої обробки [1, 2].

Головна мета аналізу тональності полягає в класифікації текстового контенту для виявлення переваг продуктів, їх недоліків та загального рівня задоволеності клієнтів. Це забезпечує можливість систематичного покращення товарів і послуг на основі зворотного зв'язку від споживачів. Задача включає кілька ключових етапів: збір даних з різних платформ електронної комерції, попередню обробку тексту, що передбачає видалення стоп-слів, нормалізацію та очищення від пунктуації, вилучення значущих ознак за допомогою спеціалізованих векторизаторів, та безпосередню класифікацію [3, 4].

У сучасних відгуках клієнтів часто виділяються конкретні аспекти продуктів або послуг, такі як ціна, якість, швидкість доставки, обслуговування клієнтів. Це вимагає застосування аспектно-орієнтованого аналізу для більш глибокого розуміння споживчих уподобань. Метод CESC ефективно використовує сегментацію слів для витягування ознак, комбінуючи архітектуру довготривалої короткотермінової пам'яті з моделлю тем, що дозволяє точніше ідентифікувати уподобання клієнтів на основі онлайн-відгуків [2]. Подібний підхід демонструє модель MWVH, яка комбінує вектори слів з паралельними каналами обробки для локальних та глобальних ознак, досягаючи середньої точності більше 90% на компактних наборах даних [5].

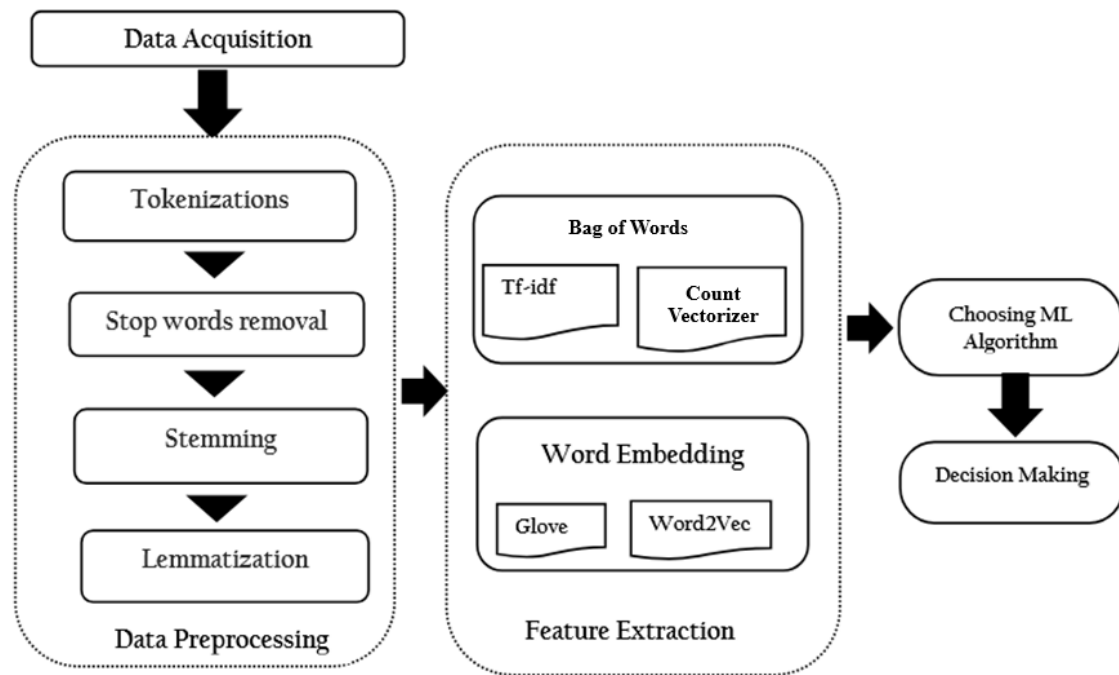


Рисунок 1.1 – Діаграма процесу [3]

Важливими параметрами задачі є врахування багатомовності контенту, наявність шумових даних різної природи та специфічний контекст конкретних категорій продуктів. У відгуках про електронні пристрої модель DC-BCGM успішно застосовує подвійний канал для обробки локальних семантичних зв'язків та глобальних залежностей, демонструючи високу точність на наборах даних туристичних відгуків, хоча це супроводжується збільшенням часу навчання на 40% [6]. Для оцінки ефективності використовуються стандартні метрики – точність, повнота, F1-міра, які дозволяють комплексно вимірювати якість класифікації. Метод LPQA ефективно виявляє шумові компоненти та покращуючи загальну якість класифікації думок [7].

Серед основних обмежень задачі виділяються нестача якісно розмічених даних для навчання, наявність упередженості у відгуках користувачів, складність обробки іронії та сарказму, що вимагає розробки гібридних підходів для досягнення кращої узагальненості моделей [1, 3]. Інтеграція контексту продукту відіграє значну роль у підвищенні якості аналізу.

Складність задачі також проявляється в необхідності обробки багатокласової класифікації, де оцінки варіюються від сильно негативних до сильно позитивних. Система з архітектурою актор-критик вирішує цю проблему через механізм винагород, що базується на точності передбачення при одночасній мінімізації кількості використаних слів [8]. Практичні реалізації на платформі Kaggle, що включають аналіз відгуків з використанням VADER та Roberta, виявляють суттєві відмінності в оцінках між методами та підходами глибокого навчання, підкреслюючи обмеження традиційних методів [9–12]. Окремі наукові дослідження використовують штучний інтелект для встановлення емоцій та характеристик продуктів, досягаючи точності 85% при групуванні відгуків за 11 ключовими факторами [13].

Загальна характеристика задачі вимагає знаходження оптимального балансу між точністю моделі та її інтерпретованістю. Сучасні виклики включають обробку неструктурованих даних великого обсягу, адаптацію до змін у мовних патернах користувачів, забезпечення стійкості моделей до атак та маніпуляцій.

## **1.2 Аналіз відомих досліджень та наукових підходів**

Сучасні публікації в галузі аналізу тональності демонструють чітку тенденцію до розробки гібридних методів, що поєднують переваги різних підходів для досягнення вищої точності. Еволюція методів від традиційних систем до складних архітектур глибокого навчання відображає зростаючу складність задач та вимог до точності. Систематичний огляд застосовності машинного навчання для аналізу настроїв клієнтів у сфері гостинності та туризму виявляє критичні прогалини в існуючих підходах, зокрема недостатню ефективність вилучення ознак при обробці великих обсягів даних [14–16]. Комплексний огляд підходів до аналізу настроїв охоплює різноманітні домени застосування, включаючи текст, голосові повідомлення та зображення, при цьому виділяючи ключові виклики, пов'язані з неоднозначністю контексту [17, 18].

Трансформерні архітектури займають провідне місце в сучасних дослідженнях. Модель RoBERTa демонструє високу ефективність в оцінці настроїв

на відгуках Amazon, точно відображаючи емоційні тони, хоча й має обмеження в узагальненні на нові домени [19, 20]. Гібридний фреймворк TWSSenti представляє інноваційний підхід через комбінацію BERT, GPT-2, RoBERTa, XLNet та DistilBERT з класичними методами TF-IDF та BoW. Ця система досягає точності 94-95% на наборах даних Sentiment140 та IMDB, використовуючи такі параметри як динамічне маскування в RoBERTa для покращення контекстуального розуміння [21]. Основним обмеженням залишається недостатня ефективність при обробці сильно зашумлених даних.

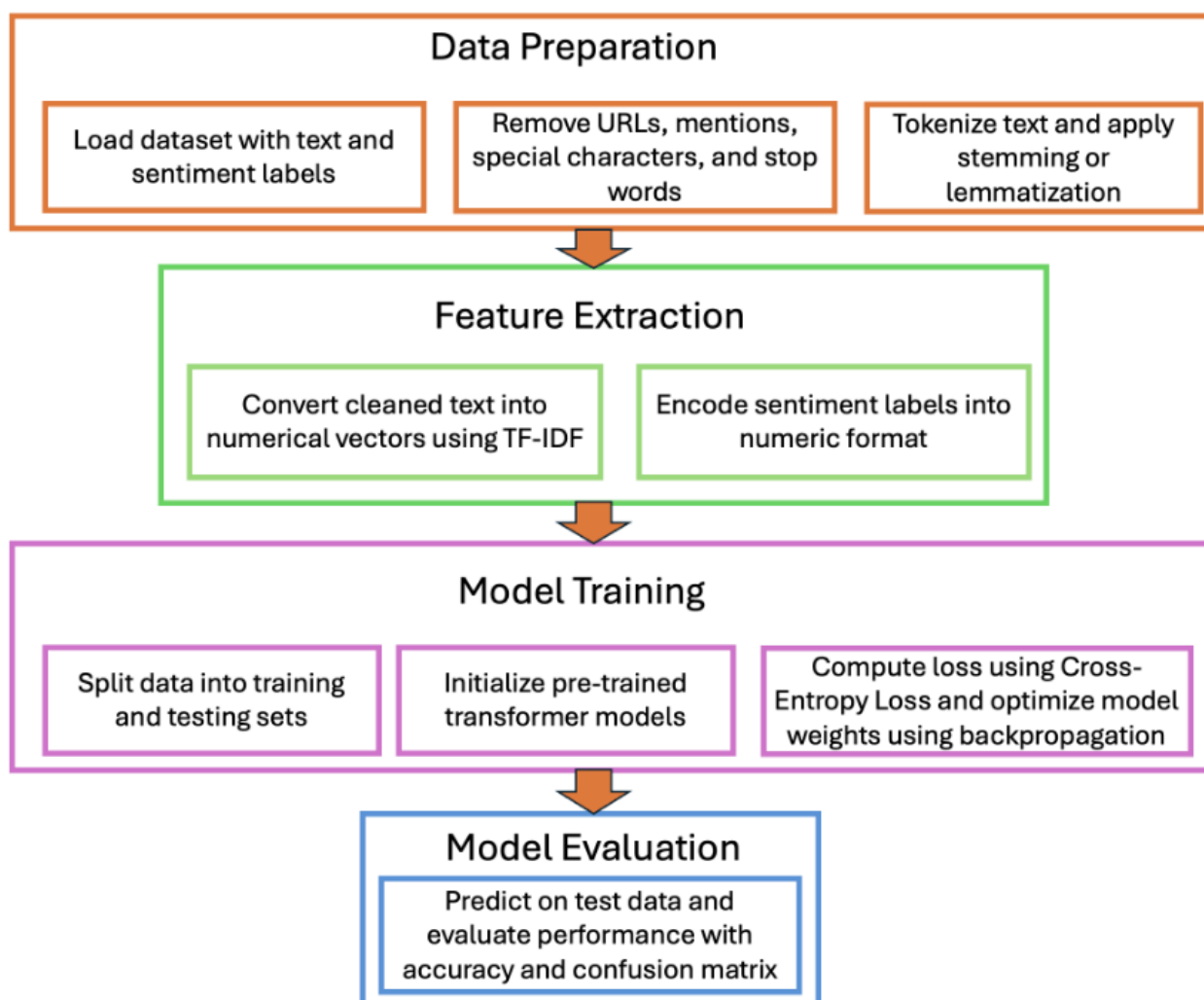


Рисунок 1.2 – Гібридний фреймворк TWSSenti [21]

Аналіз репозиторіїв GitHub виявляє різноманітність практичних реалізацій. Проект Sentiment-Analysis-on-Online-Product-Reviews застосовує SVM та SGDClassifier з TF-IDF векторизацією, де TF розраховується як відношення кількості

появи терміну до загальної кількості термінів, а IDF визначається через логарифм відношення загальної кількості документів до кількості документів з терміном [22]. Альтернативний репозиторій використовує TF-IDF у поєднанні з SVM, Naïve Bayes та деревом рішень, оцінюючи точність через графічну візуалізацію результатів [23].

Підходи з глибоким навчанням, представлені в Customer-Reviews-Analysis, застосовують ансамблеві методи, комбінуючи Random Forest, SVM, LSTM та CNN з STV-голосуванням. Ефективність оцінюється через стандартні метрики – точність, повнота, F1-міра, хоча конкретні параметри моделей не завжди розкриваються [24–26].

Комплексний аналіз відгуків з використанням NLP включає метрики CSAT та NPS, застосовуючи інструменти як Brandwatch для категоризації контенту, хоча безплатні версії мають суттєві функціональні обмеження [27–29]. Особливу увагу привертає задача виявлення фейкових відгуків, де графові нейронні мережі використовуються для аналізу зв'язків між користувачами та продуктами.

Дослідження описують метод CESC для витягування ознак, що поєднує LSTM з LDA для ефективного виявлення уподобань користувачів, хоча конкретні параметри моделі часто не розкриваються [30, 31]. Модель MWVN інтегрує CNN та RNN з механізмом уваги, досягаючи точності від 94.62% до 95.14%, при цьому основним обмеженням залишається розмір доступних навчальних наборів [32].

Важливим напрямком є розробка методів для роботи з мультимодальними даними, де текстові відгуки доповнюються зображеннями та відео. Це створює нові можливості для більш точного аналізу, але водночас ускладнює архітектуру моделей та вимагає значних обчислювальних ресурсів.

Сучасні тенденції включають використання перенесеного навчання для адаптування моделей до нових доменів, застосування методів федеративного навчання для збереження приватності користувачів, розробку методів пояснюваного штучного інтелекту для підвищення довіри до результатів аналізу. Критичним залишається питання масштабованості рішень та їх адаптації до специфіки різних мов та культурних контекстів.

### 1.3 Огляд архітектур, методів машинного навчання для аналізу тональності

Сучасні архітектури машинного навчання для аналізу тональності демонструють значну еволюцію від простих статистичних моделей до складних нейромережових систем. Трансформерні архітектури займають центральне місце в сучасних рішеннях. RoBERTa застосовує механізм динамічного маскування в процесі навчання, що дозволяє моделі краще узагальнювати контекстуальну інформацію. Емпіричні дослідження показують високу ефективність цієї архітектури для оцінки настроїв, хоча виявляються обмеження при застосуванні в специфічних доменах, таких як поведінкова економіка [20, 33, 34].

Архітектура DC-BCGM представляє комплексний підхід до аналізу тональності через інтеграцію кількох компонентів. BERT використовується для генерації динамічних векторних представлень, BiGRU обробляє глобальні залежності з параметрами оновлення, CNN витягує локальні семантичні патерни, а механізм уваги визначає відносну важливість різних частин тексту [35, 36]. Система досягає точності 95.85% та F1-міри 97.65% на стандартних наборах даних, хоча високі обчислювальні вимоги обмежують її застосування в реальному часі [37].

Методи з підкріпленням навчанням відкривають нові можливості для оптимізації аналізу тональності. Архітектура актор-критик використовує актора для вибору оптимальних дій при обробці тексту та критика для оцінки якості цих дій. Система винагород враховує як семантичну точність класифікації, так і ефективність використання контексту. FastXCatStack демонструє ефективність ансамблевого підходу, комбінуючи FastText для швидкої векторизації тексту з градієнтним бустингом XGBoost та CatBoost [38–40].

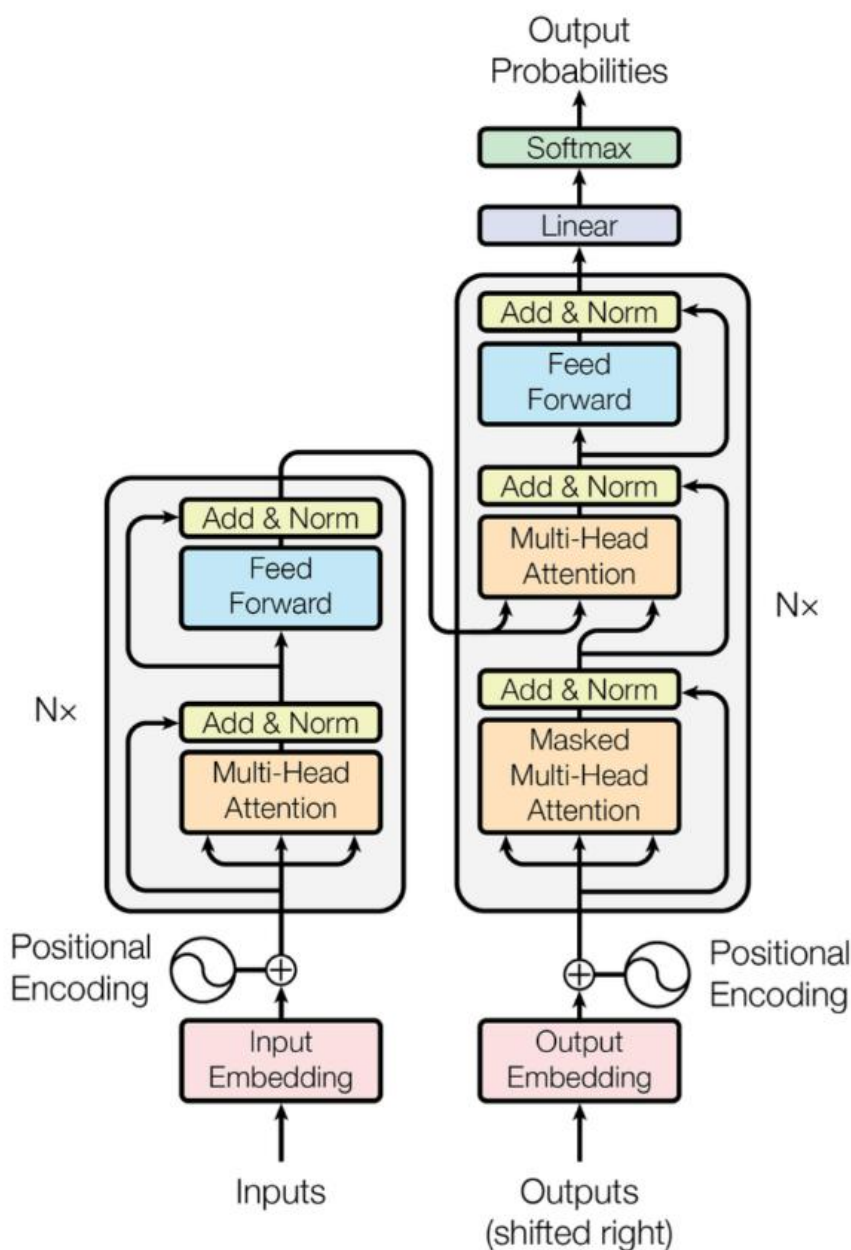


Рисунок 1.2 – Архітектура моделі трансформатора [20]

Для інтерпретації результатів використовується LIME, що дозволяє зрозуміти вплив окремих ознак на фінальне рішення [41, 42]. Реалізації на платформі Hugging Face, такі як Reviews-Sentiment-Analysis, використовують попередньо навчені моделі з механізмами оцінки впевненості в передбаченнях, що особливо важливо для практичних застосувань [43].

На платформі Kaggle представлені різноманітні реалізації LSTM для аналізу відгуків, що демонструють точність в діапазоні 80-90% при використанні TF-IDF векторизації. Ці моделі показують нижчі помилки порівняно з традиційними

методами, особливо при обробці довгих текстів з складними залежностями. Архітектура CNN для аналізу тональності використовує згорткові шари різних розмірів для захоплення n-грамних патернів. Типова конфігурація включає фільтри розміром 3, 4 та 5 слів, що дозволяє виявляти локальні семантичні зв'язки. Механізм max-pooling витягує найбільш значущі ознаки з кожної карти ознак, формуючи компактне представлення для фінальної класифікації. RNN та їх варіанти LSTM та GRU ефективно обробляють послідовну природу тексту. LSTM використовує механізми воріт для контролю потоку інформації [44, 45].

Гібридні архітектури поєднують переваги різних підходів. CNN-LSTM використовує CNN для витягування локальних ознак, які потім обробляються LSTM для врахування послідовних залежностей. Така комбінація особливо ефективна для довгих текстів з складною структурою. Механізми уваги революціонізували аналіз тональності, дозволяючи моделям системи фокусуватися на релевантних загальних частинах тексту. Самоувага в трансформерах обчислює відносну важливість кожного токена відносно всіх інших, формуючи багате контекстуальне представлення.

Методи трансферного навчання дозволяють застосовувати знання, отримані на великих корпусах тексту, для специфічних задач аналізу тональності. BERT та його варіації попередньо навчаються на мільярдах токенів тексту, після чого дотренуються на конкретних датасетах відгуків. Це значною мірою знижує вимоги до розміру навчальної вибірки та покращує узагальнення. Основні обмеження сучасних архітектур включають високі обчислювальні вимоги для трансформерних моделей, складність інтерпретації результатів глибоких нейронних мереж, чутливість до розподілу даних у навчальній вибірці, необхідність великих обсягів розмічених даних для досягнення необхідною точності [46].

#### **1.4 Мета та постановка задачі**

Відповідно до виконаного дослідження існуючих підходів сформульовано цілі та конкретні завдання дослідження.

Метою роботи є підвищення точності автоматичної класифікації тональності клієнтських відгуків через розробку гібридного методу, що базується на інтеграції різнотипних алгоритмів машинного навчання.

Задачі дослідження:

– здійснити огляд наявних методологій та технологічних рішень у сфері автоматичного визначення тональності текстових даних, включаючи традиційні статистичні підходи та сучасні нейромережеві архітектури;

– розробити метод класифікації емоційного забарвлення відгуків, який використовує послідовну інтеграцію для досягнення підвищеної точності розпізнавання;

– розробити модифікацію базового методу шляхом включення аналізу біграм на етапі формування текстових ознак, що дозволить враховувати локальний контекст слів та покращити розпізнавання складних лінгвістичних конструкцій;

– створити програмну реалізацію запропонованого методу з використанням модульної архітектури, що забезпечить можливість незалежного розвитку окремих компонентів системи;

– провести експериментальне дослідження розробленого методу на репрезентативному наборі реальних клієнтських відгуків з оцінкою ключових метрик якості класифікації.

## **Розділ 2 Метод аналізу тональності відгуків клієнтів та критерії його оцінювання**

### **2.1 Концепція та схема методу аналізу тональності**

Аналіз тональності відгуків клієнтів являє собою задачу класифікації тексту, де необхідно визначити емоційне забарвлення повідомлення. У контексті цієї роботи розглядається тризначна класифікація, де відгуки можуть бути різної тональності. Така класифікація дозволяє отримати загальне уявлення про ставлення клієнтів до продукту або послуги.

Концепція запропонованого методу базується на використанні методів навчання для автоматизованого визначення тональності. Ідея методу полягає в тому, щоб модель розпізнавала закономірності в текстових даних, які вказують на позитивну, негативну або нейтральну тональність. Для цього використовується набір попередньо розмічених відгуків, на основі яких модель навчається виявляти характерні ознаки кожного класу тональності.

Метод складається з декількох етапів послідовної обробки інформації. На першому етапі здійснюється збирання та попереднє оброблення текстових даних. Цей необхідний етап включає очищення тексту від зайвих символів, видалення стоп-слів, зведення до нижнього регістру, та інші операції, що покращують якість подальшої обробки. На другому етапі здійснюється перетворення даних тексту у числа, яке може бути оброблене моделлю машинного навчання. Третій етап передбачає навчання класифікаційної моделі на підготовлених даних. Останній етап включає застосування моделі для класифікування нових відгуків та оцінку якості її роботи.

Загальна схема методу аналізу тональності наведена на рисунку 2.1. Схема демонструє послідовність кроків від отримання вхідних даних до видачі фінального результату класифікації. Вхідні дані надходять у вигляді необроблених текстових відгуків клієнтів. Ці відгуки проходять через модуль попередньої обробки, де виконується токенізація, нормалізація та очищення тексту. Після попередньої обробки дані потрапляють у модуль векторизації, де текст перетворюється на числові вектори за методом TF-IDF.

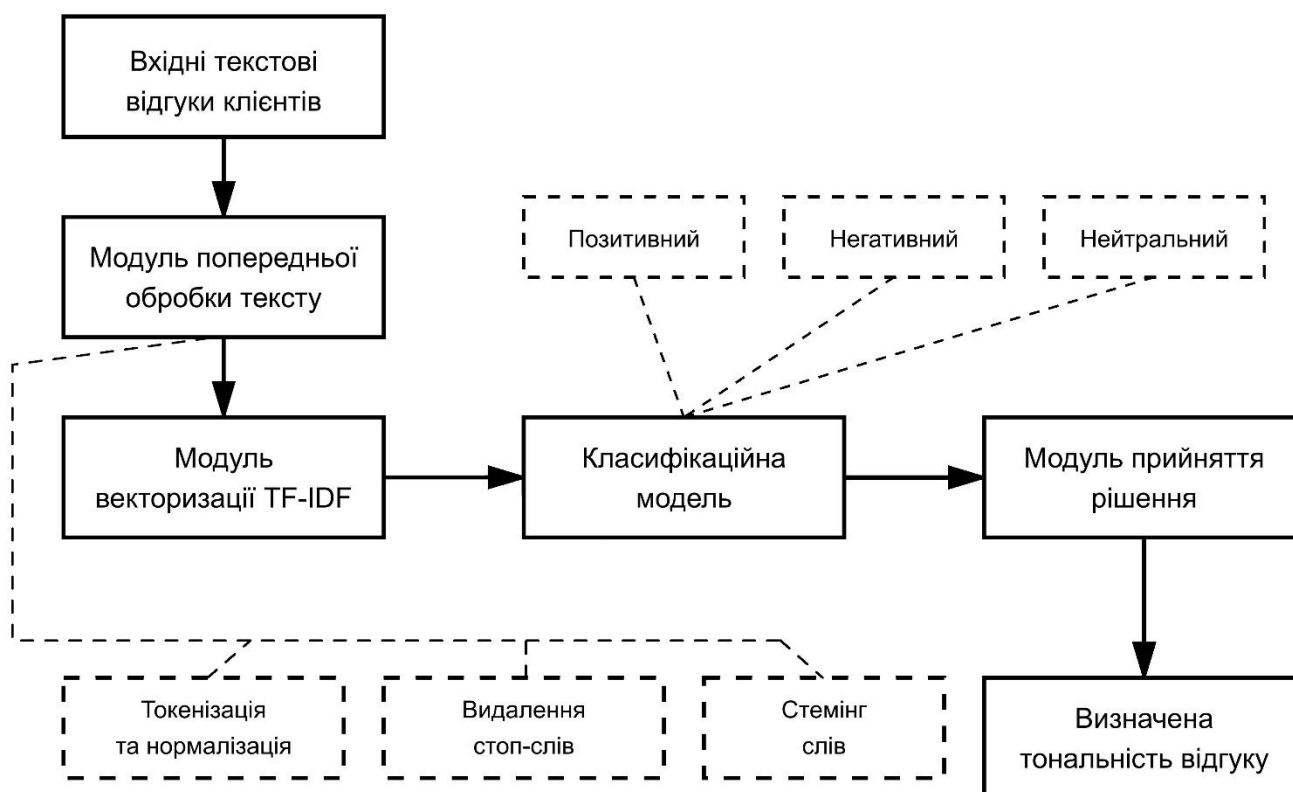


Рисунок 2.1 – Загальна схема методу аналізу тональності

Підготовлені векторні представлення текстів подаються на вхід класифікаційної моделі. Модель аналізує ознаки тексту та визначає ймовірність належності до класів кожної тональності. На основі цих ймовірностей приймається рішення про фінальний клас. Результатом роботи методу є визначена тональність відгуку з відповідною впевненістю моделі в своєму рішенні.

Вибір саме такого підходу обґрунтовується декількома факторами. Використання методів навчання дає змогу автоматизувати процес аналізу великої кількості відгуків, що було б неможливо виконати вручну в розумні терміни. Навчена модель здатна виявляти складні закономірності в даних, які не завжди очевидні при ручному аналізі. Такий підхід забезпечує однорідність у класифікації, оскільки модель застосовує однакові критерії до всіх відгуків.

Важливою особливістю запропонованого методу є його модульність. Кожен етап обробки може бути модифікований або замінений без суттєвих змін в інших частинах системи. Така архітектура забезпечує гнучкість та можливість подальшого

вдосконалення методу. Наприклад, можна змінити спосіб векторизації тексту або використати іншу модель класифікації без необхідності переробляти всю систему.

Метод розроблявся з урахуванням практичних вимог до систем аналізу тональності. Зокрема, приділялася увага балансу між точністю класифікації та швидкістю обробки. Занадто складні моделі можуть досягати високої точності, але при цьому вимагають значних обчислювальних ресурсів та часу для обробки кожного відгуку. Запропонований підхід намагається знайти оптимальне співвідношення між цими параметрами.

Концепція методу також передбачає можливість роботи з відгуками різної довжини та структури. Клієнтські відгуки можуть варіюватися від коротких оцінок з декількох слів до розгорнутих текстів з детальним описом досвіду використання продукту. Метод повинен однаково добре справлятися з класифікацією як коротких, так і довгих відгуків, виявляючи ключові ознаки тональності незалежно від обсягу тексту.

## **2.2 Архітектура гібридної моделі класифікації**

Для вирішення задачі аналізу тональності відгуків було обрано гібридний підхід, що поєднує традиційні методи навчання з елементами глибокого навчання. Архітектура побудована за принципом послідовної обробки, де кожен компонент використовує результати попереднього етапу для покращення якості класифікації. Основу архітектури становить послідовна комбінація методу опорних векторів та Dense Neural Network. Такий підхід дозволяє використати переваги обох методів у послідовному порядку. Метод опорних векторів добре працює з високорозмірними даними та здатний знаходити оптимальну розділяючу гіперплощину між класами. Dense Neural Network, у свою чергу, дає змогу виявляти складні нелінійні залежності та навчитися коригувати передбачення базової моделі, використовуючи інформацію про її впевненість у класифікації. Архітектура моделі наведена на рисунку 2.2. На вхід першого етапу моделі подається векторне подання тексту, одержане за допомогою TF\_IDF векторизації з необхідним додаванням біграм. Розмірність обробленого

вектора складає близько 3500 ознак, що забезпечує достатнє представлення тексту з врахуванням контексту слів через біграми.

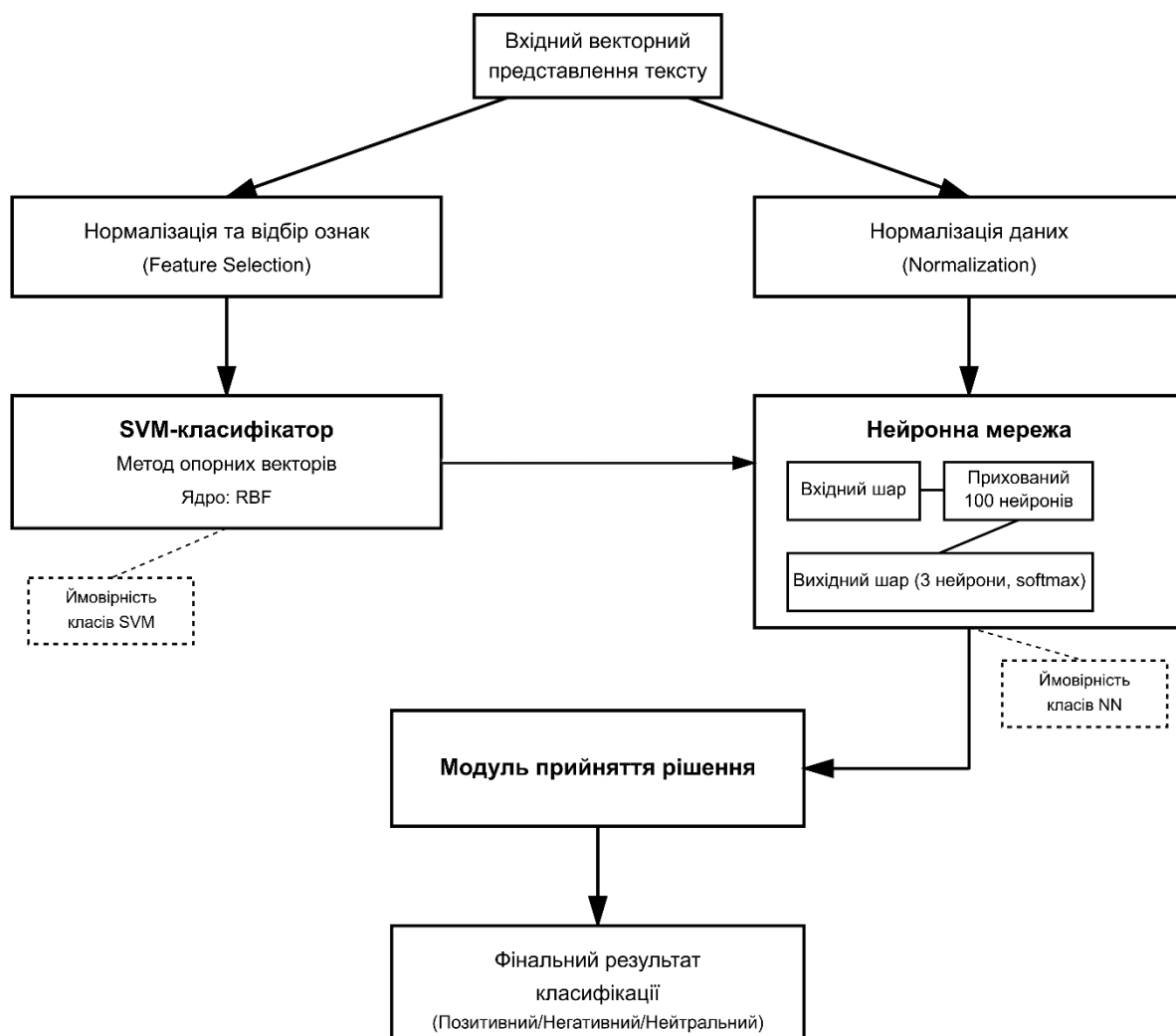


Рисунок 2.2 – Архітектура гібридної моделі класифікації

Перший компонент архітектури реалізований у вигляді методу опорних векторів з радіально-базисною функцією ядра. Вибір RBF-ядра обумовлений його здатністю моделювати складні нелінійні залежності між ознаками. Параметри ядра, такі як коефіцієнт  $\gamma$  та параметр регуляризації, підбираються експериментально для досягнення найкращої якості класифікації. SVM-класифікатор обробляє вхідні ознаки та формує передбачення у вигляді трьох ймовірностей для кожного класу тональності. Другий компонент архітектури представляє собою Dense Neural Network з простою архітектурою. Ключовою особливістю цього компоненту є те, що він

приймає на вхід не лише оригінальні векторні представлення тексту, а й ймовірності, отримані від SVM-класифікатора.

Таким чином, вхідний вектор для нейронної мережі складається з 3500 оригінальних ознак та 3 додаткових ознак – ймовірностей від SVM. Загальна величина вхідного вектора для мережі становить 3503 ознаки. Архітектура мережі включає вхідний шар розмірністю 3503 нейрони, один прихований шар з 100 нейронами та вихідний шар з 3 нейронами. Прихований шар застосовує функцію активаційну ReLU, яка добре зарекомендувала себе для задач класифікації тексту. Вихідний шар застосовує функцію активаційну Softmax для отримання ймовірнісного розподілу по класах тональності. Навчання моделі відбувається послідовно у два етапи.

Спочатку навчається SVM-класифікатор на повному наборі навчальних даних з використанням стандартної процедури оптимізації. Після завершення навчання SVM, модель застосовується до всього навчального набору для отримання передбачень у вигляді ймовірностей. На другому етапі навчається Dense Neural Network. Навчальні дані для нейронної мережі формуються шляхом конкатенації оригінальних TF-IDF векторів з біграмами та ймовірностей, отриманих від SVM. Таким чином, нейронна мережа навчається не лише на основі текстових ознак, а й враховує, наскільки впевнений SVM-класифікатор у своїх передбаченнях. Це дозволяє мережі виявляти випадки, коли SVM помиляється або сумнівається, і навчитися коригувати ці помилки.

Така послідовна архітектура реалізує концепцію мета навчання або стекінгу, яка широко використовується в ансамблевих методах машинного навчання. Основна ідея полягає в тому, що друга модель навчається виправляти помилки першої моделі, використовуючи додаткову інформацію про впевненість базової моделі у своїх передбаченнях. Важливою особливістю архітектури є те, що нейронна мережа має доступ як до низькорівневих ознак TF-IDF вектори, так і до високорівневих ознак передбачення SVM. Низькорівневі ознаки дозволяють мережі самостійно виявляти закономірності в тексті, тоді як високорівневі ознаки надають інформацію про те, як базова модель інтерпретує ці закономірності. Комбінація цих двох типів інформації

дозволяє досягти кращої якості класифікації порівняно з використанням лише однієї моделі.

Параметри моделі підбиралися експериментально з використанням крос-валідації. Для SVM-класифікатора параметр  $C$  варіювався в діапазоні від 0.1 до 10, а параметр  $\gamma$  від 0.001 до 1. Для нейронної мережі підбиралася кількість нейронів у прихованому шарі від 50 до 200, а також швидкість навчання від 0.001 до 0.1. Оптимальні значення параметрів визначалися на основі якості класифікації на валідаційній вибірці. Було встановлено, що 100 нейронів у прихованому шарі забезпечують досить добрий баланс між складністю моделі та її здатністю до узагальнення. Навчання нейронної мережі здійснюється з використанням оптимізатора Adam та функції втрат. Оптимізатор Adam обрано через його здатність адаптивно налаштовувати швидкість навчання для кожного параметра, що прискорює збіжність та покращує якість навчання.

Функція втрат є стандартним вибором для задач багатокласової класифікації, оскільки вона ефективно штрафує моделі за невірні передбачення пропорційно до їх віддаленості від правильного класу. Архітектура передбачає можливість подальшого розширення без суттєвої зміни існуючої структури. Наприклад, можна додати ще один шар до нейронної мережі або використати більш складні архітектури, такі як LSTM або трансформери.

Також можна експериментувати з різними способами комбінування передбачень SVM з оригінальними ознаками. Така модульність забезпечує гнучкість системи та можливість її подальшого розвитку відповідно до нових вимог або доступності додаткових обчислювальних ресурсів. Перевагою послідовної архітектури є те, що вона дозволяє інтерпретувати роль кожного компоненту в процесі класифікації. SVM виступає в ролі базового класифікатора, який формує первинне розуміння тональності тексту. Dense Neural Network виконує роль коректора, який аналізує, чи можна довіряти передбаченню SVM, і за необхідності вносить корективи. Таке розділення обов'язків робить систему більш прозорою та зрозумілою порівняно з повністю підходами глибокого навчання.

## 2.3 Модифікація моделі та покращення точності класифікації

Базова архітектура гібридної моделі, описана в попередньому підрозділі, показала прийнятні результати на тестових даних. Однак аналіз помилок класифікації виявив деякі слабкі місця, що можна покращити за допомогою невеликих модифікацій. Основна проблема базової моделі полягає в недостатньо точній класифікації нейтральних відгуків, які часто помилково відносяться до позитивного або негативного класу.

Запропонована модифікація спрямована на покращення обробки ознак перед їх подачею до класифікаторів. Основна ідея полягає у додаванні додаткового етапу аналізу контексту слів. У базовій моделі використовується стандартна TF-IDF векторизація, яка враховує лише частоту появи слів у документі та їх рідкість у всьому корпусі. Проте такий підхід не враховує контекст, в якому з'являються слова, що може призводити до втрати важливої інформації про тональність.

Модифікація передбачає введення додаткового блоку аналізу біграм перед основною векторизацією. Біграми являють собою послідовності з двох сусідніх слів у тексті. Врахування біграм дозволяє частково захопити контекст слів та їх взаємозв'язки. Наприклад, словосполучення "не добре" та "дуже добре" містять слово "добре", але мають протилежну тональність. Використання біграм допомагає моделі розрізняти такі випадки.

Модифікована архітектура наведена на рисунку 2.3. Порівняно з базовою архітектурою, додано блок формування біграм після токенізації тексту. Цей блок створює пари сусідніх слів та додає їх до списку ознак разом з окремими словами. Далі комбінований список слів та біграм проходить через TF-IDF векторизацію, створюючи розширене представлення тексту.

Теоретичне обґрунтування цієї модифікації базується на лінгвістичних особливостях вираження тональності в тексті. Емоційне забарвлення часто передається не окремими словами, а їх комбінаціями. Заперечні конструкції, посилювальні прикметники та модальні слова формують контекст, який впливає на

тональність висловлювання. Врахування біграм дозволяє моделі краще розпізнавати такі конструкції.

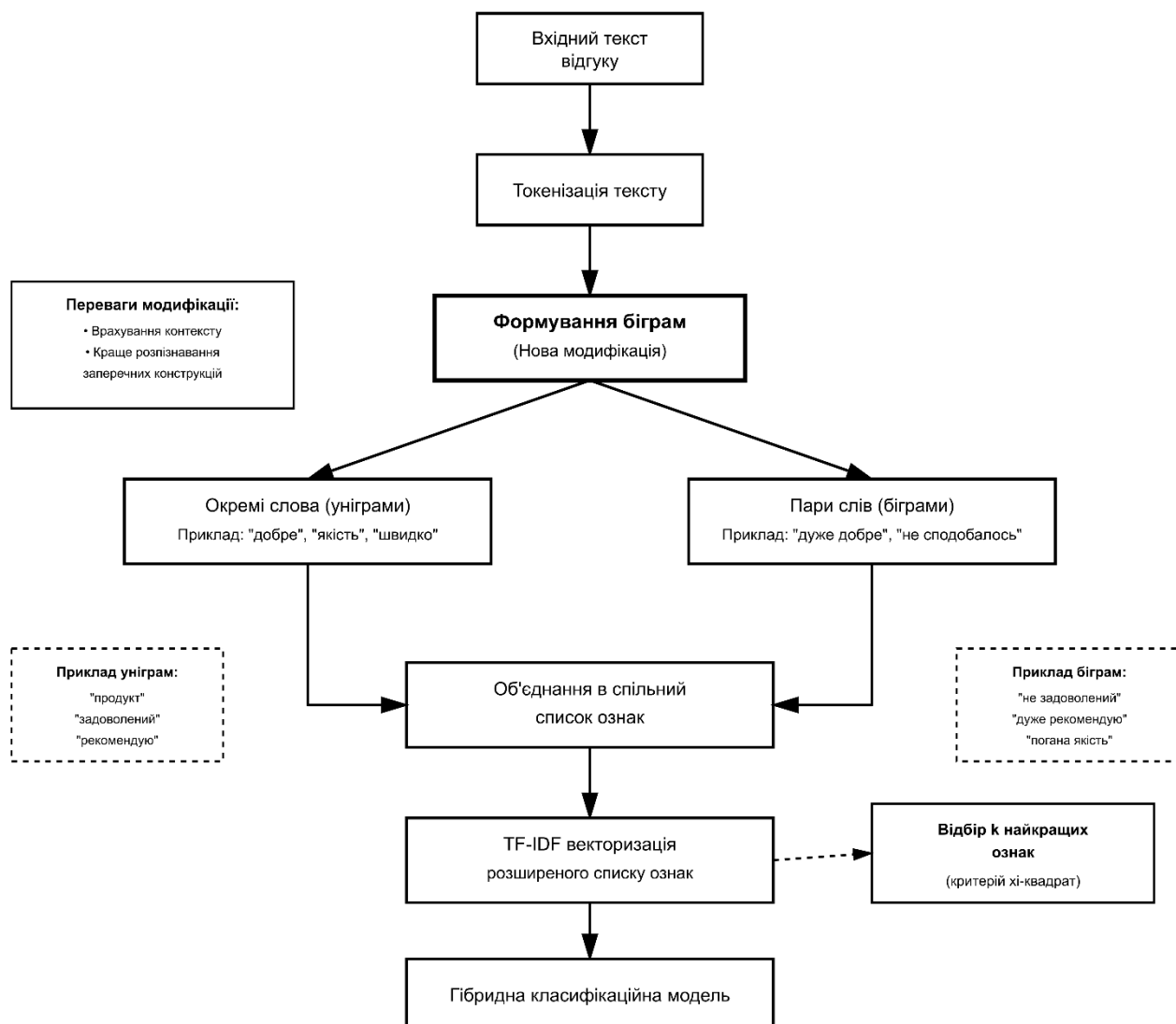


Рисунок 2.3 – Модифікована архітектура з аналізом біграм

Впровадження аналізу біграм призводить до збільшення розмірності векторного представлення тексту. Якщо базова модель використовувала словник розміром 3000 слів, то після додавання біграм розмірність може збільшитися до 5000-6000 ознак. Це збільшує обчислювальну складність навчання та застосування моделі, але покращує якість класифікації, особливо для складних випадків.

Для компенсації збільшеної розмірності застосовується більш агресивний відбір ознак. Використовується метод вибору  $k$  найкращих ознак на основі статистичного критерію  $\chi^2$ . Цей метод оцінює незалежність кожної ознаки від

цільової змінної та відбирає ті ознаки, які мають найсильніший зв'язок з класом тональності. Типово відбирається 3000-4000 найбільш інформативних ознак з повного набору.

Очікуваними перевагами запропонованої модифікації є підвищення точності класифікації нейтральних відгуків та покращення загальної F1-міри моделі. Біграми допомагають краще розпізнавати заперечні конструкції, які часто зустрічаються в нейтральних відгуках, де користувачі описують як позитивні, так і негативні аспекти продукту. Також очікується покращення класифікації коротких відгуків, де контекст особливо важливий через обмежену кількість слів.

Додатковою перевагою є відносна простота реалізації модифікації. Формування біграм виконується за допомогою стандартних функцій бібліотеки `scikit-learn` та не вимагає складних алгоритмів або значних обчислювальних ресурсів. Модифікація може бути легко інтегрована в існуючу архітектуру без необхідності переробляти інші компоненти системи.

Потенційним недоліком модифікації є збільшення часу навчання моделі через більшу кількість ознак. Проте попередні експерименти показують, що це збільшення є прийнятним та не перевищує 20-30% порівняно з базовою моделлю. Час класифікації нових відгуків також збільшується незначно, що дозволяє використовувати модифіковану модель у практичних застосуваннях.

## **2.4 Формування та підготовка навчальних даних**

Якість роботи будь-якої моделі машинного навчання критично залежить від якості даних, на яких вона навчається. Для задачі аналізу тональності відгуків було використано датасет з платформи Kaggle, який містить відгуки клієнтів з різними рівнями тональності Customer Feedback Dataset. Датасет включає приблизно 5000 записів, кожен з яких містить текст відгуку та відповідну мітку тональності.

Структура датасету включає три основні поля. Перше поле містить унікальний ідентифікатор відгуку, який використовується для відстеження конкретних записів під час обробки. Друге поле містить власне текст відгуку, який може варіюватися від

кількох слів до декількох речень. Третє поле містить мітку тональності, яка може приймати одне з трьох значень: позитивна, негативна або нейтральна. Розподіл класів у датасеті є відносно збалансованим, що важливо для якісного навчання моделі.

Попередня обробка текстових даних є критично важливим етапом підготовки даних для навчання. Цей процес включає декілька послідовних кроків, кожен з яких спрямований на покращення якості представлення тексту. Схема попередньої обробки наведена на рисунку 2.4.

Перший крок попередньої обробки полягає в приведенні всього тексту до нижнього регістру. Це необхідно для того, щоб слова "Добре", "добре" та "ДОБРЕ" розглядалися моделлю як одне і те саме слово. Без цього кроку модель витрчала б окремі ознаки на різні варіанти написання одного й того самого слова, що знижує якість класифікації.

Другий крок включає видалення спеціальних символів, цифр та пунктуації. Більшість знаків пунктуації не несуть корисної інформації для визначення тональності та можуть розглядатися як шум. Однак деякі символи, такі як знаки оклику або питання, можуть містити інформацію про емоційне забарвлення тексту, тому їх обробка потребує окремої уваги. У цій роботі було прийнято рішення видаляти всю пунктуацію для спрощення обробки.

Третій крок передбачає токенізацію тексту, тобто розбиття на окремі слова. Для токенізації використовується простий метод розбиття за пробілами після видалення зайвих символів. Більш складні методи токенізації, які враховують морфологію мови або використовують спеціальні правила, не застосовувалися для збереження простоти підходу.

Четвертий крок включає видалення стоп-слів – часто вживаних слів, які не несуть змістового навантаження для визначення тональності. До таких слів належать прийменники, сполучники, артиклі та інші службові частини мови. Для англійської мови використовується стандартний список стоп-слів з бібліотеки NLTK, який містить близько 180 слів. Видалення стоп-слів зменшує розмірність даних та допомагає моделі сфокусуватися на більш значущих словах.

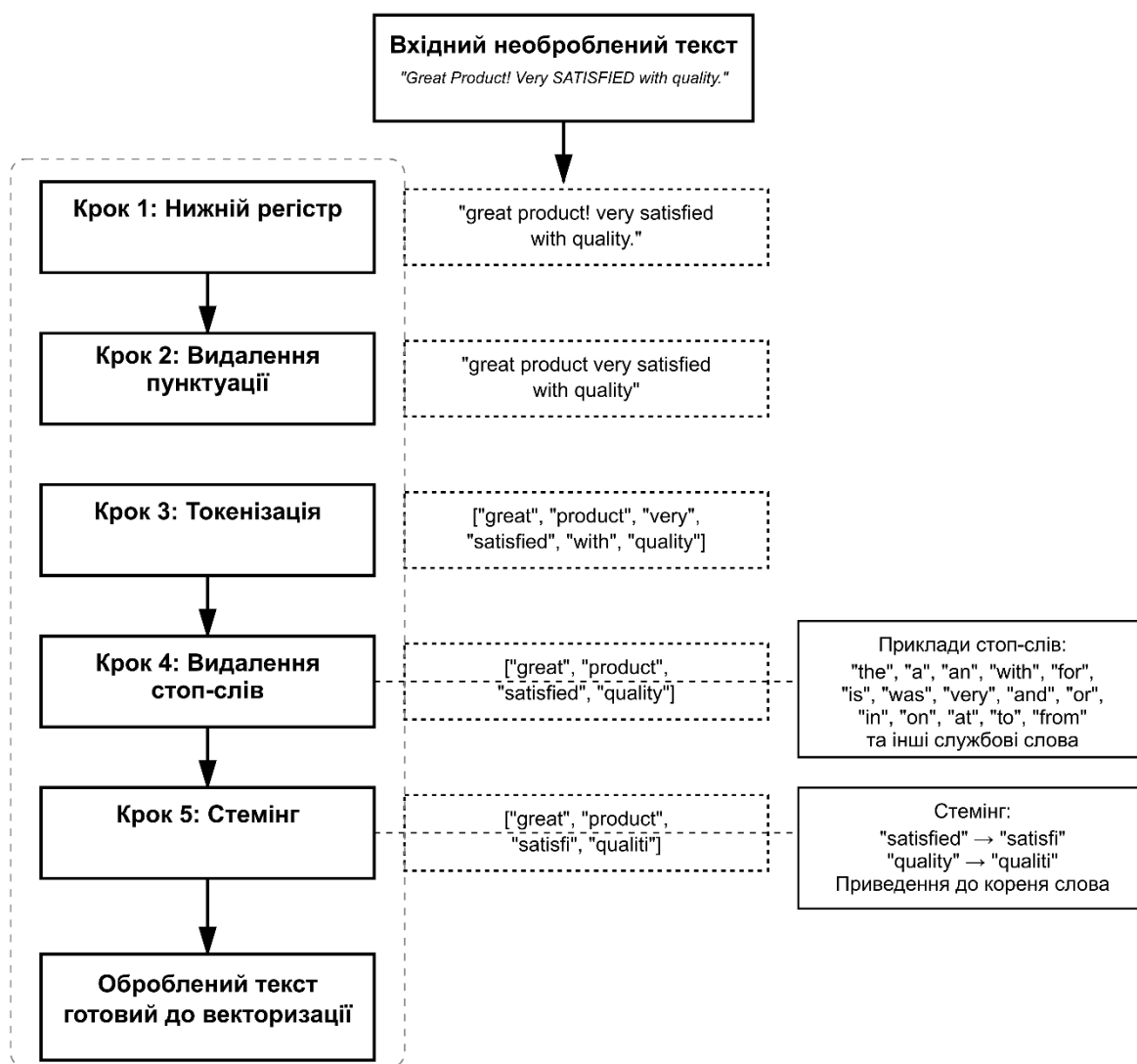


Рисунок 2.4 – Схема попередньої обробки текстових даних

П'ятий крок передбачає стемінг або лематизацію слів. Ці процеси спрямовані на приведення слів до їх базової форми. Наприклад, слова "працював", "працює" та "працювати" мають однаковий корінь та несуть схожий зміст. У цій роботі використовується стемінг за допомогою алгоритму Porter Stemmer, який відокремлює закінчення та суфікси від основи слова. Хоча цей метод іноді призводить до втрати граматичної правильності слів, він дозволяє значно зменшити розмір словника.

Після завершення всіх кроків попередньої обробки текст готовий до векторизації. Для перетворення текстових даних у числове представлення використовується метод TF-IDF. Цей метод враховує два фактори: частоту терміну в документі та його рідкість у всьому корпусі. Формально, TF-IDF для терміну  $t$  у

документі  $d$  обчислюється як добуток частоти терміну  $TF$  та інверсної частоти документа  $IDF$ .

Частота терміну  $TF$  визначається як кількість безпосередніх проявів терміну в документі, поділена на сукупну кількість термінів у цьому документі. Це нормалізує частоту з урахуванням довжини документа. Інверсна частота документа  $IDF$  обраховується як важливий логарифм співвідношення загальної кількості документів до кількості документів, що містять даний термін. Це дає більшу вагу термінам, які рідко зустрічаються в корпусі, але є важливими для конкретних документів.

Параметри векторизації підбиралися експериментально. Максимальний розмір словника було встановлено на рівні 3000 найбільш частотних термінів. Мінімальна частота появи терміну для включення до словника дорівнює 2, що означає, що термін повинен з'явитися хоча б у двох різних документах. Максимальна частота обмежена на рівні 80%, щоб виключити надто часті слова, які з'являються майже в кожному документі та не несуть дискримінативної інформації.

Після векторизації кожен відгук представлений у вигляді розрідженого вектора розмірністю 3000, де кожен елемент відповідає  $TF-IDF$  значенню відповідного терміну. Більшість елементів цього вектора дорівнюють нулю, оскільки конкретний відгук містить лише невелику частину всього словника. Використання розрідженого представлення дозволяє ефективно зберігати та обробляти такі вектори.

Підготовлений датасет розділяється на три частини: навчальну, валідаційну та тестову вибірки. Вибірка для навчання містить 70% всіх даних і використовується безпосередньо для навчання моделі. Валідаційна вибірка становить 15% даних і використовується для підбирання гіперпараметрів та контролю перенавчання під час тренування. Тестова вибірка, що містить решту 15% даних, використовується для результуючої оцінки якості моделі на даних, які вона ніколи не бачила під час навчання.

Розподіл на вибірки виконується випадковим чином з використанням фіксованого початкового значення для забезпечення відтворюваності результатів. При цьому зберігається пропорція класів у кожній вибірці, що важливо для коректного навчання та оцінювання моделі. Такий підхід називається

стратифікованим розподілом та гарантує, що кожна вибірка містить приблизно однакове співвідношення позитивних, негативних та нейтральних відгуків.

## 2.5 Критерії та метрики оцінювання роботи методу

Для об'єктивної оцінки якості роботи розробленого методу необхідно використовувати відповідні метрики, які дозволяють кількісно виміряти точність класифікації. У задачах машинного навчання існує стандартний набір метрик, що використовуються для оцінки моделей класифікації. Вибір конкретних метрик залежить від специфіки задачі та вимог до системи.

Для задачі аналізу тональності відгуків було обрано чотири основні метрики: точність, повноту, F1-міру та загальну точність. Ці метрики дозволяють оцінити різні аспекти роботи моделі та отримати комплексне уявлення про її якість. Кожна з цих метрик обчислюється окремо для кожного класу тональності, а потім усереднюється для отримання загальної оцінки моделі.

Точність визначає частку правильно класифікованих зразків серед усіх зразків, які модель віднесла до певного класу. Формально точність для класу обчислюється як відношення кількості правильно класифікованих зразків цього класу до загальної кількості зразків, які модель класифікувала як цей клас. Висока точність означає, що коли модель відносить відгук до певного класу, вона робить це правильно в більшості випадків:

$$Precision = \frac{TP}{TP + FP} \quad (2.1)$$

де TP (True Positives) кількість правильно класифікованих зразків класу, FP (False Positives) кількість зразків інших класів, помилково класифікованих як даний клас.

Повнота характеризує здатність моделі знаходити всі зразки певного класу. Вона обчислюється як відношення кількості правильно класифікованих зразків класу до загальної кількості зразків цього класу в датасеті. Висока повнота означає, що

модель знаходить більшість зразків класу, навіть якщо при цьому робить деякі помилки та відносить до цього класу зразки інших класів<sup>^</sup>

$$Recall = \frac{TP}{TP + FN} \quad (2.2)$$

де TP – кількість правильно класифікованих зразків класу, FN (False Negatives) – кількість зразків даного класу, помилково класифікованих як інші класи.

F1-міра є гармонійним середнім між точністю та повнотою. Ця метрика особливо корисна, коли потрібно збалансувати точність та повноту, оскільки покращення однієї з цих метрик часто призводить до погіршення іншої. F1-міра обчислюється як подвоєний добуток точності та повноти, поділений на їх суму. Значення F1-міри завжди знаходиться між значеннями точності та повноти, будучи ближчим до меншого з них:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (2.3)$$

Загальна точність визначає частку всіх правильно класифікованих зразків серед усіх зразків у тестовій вибірці, незалежно від класу. Ця метрика дає загальне уявлення про якість моделі, але може бути оманливою при несбалансованих датасетах, де один клас значно переважає інші. У такому випадку модель може досягти високої точності, просто завжди передбачаючи найчастіший клас. Проте для збалансованого датасету, який використовується в цій роботі, точність є корисною метрикою загальної якості.

Для багатокласової класифікації, якою є задача аналізу тональності з трьома класами, метрики можуть обчислюватися різними способами усереднення. Найпоширенішими є макро-усереднення та зважене усереднення. Макро-усереднення обчислює метрику для кожного класу окремо, а потім усереднює отримані значення без урахування кількості зразків у кожному класі. Зважене

усереднення враховує кількість зразків кожного класу, надаючи більшу вагу класам з більшою кількістю зразків.

У цій роботі використовується макро-усереднення для основних метрик, оскільки воно дає рівну важливість кожному класу незалежно від його розміру. Це важливо, коли помилки класифікації для різних класів мають однакову вартість з точки зору застосування в практиці. Наприклад, неправильна класифікація позитивного відгуку як негативного має таку ж важливість, як і помилкова класифікація нейтрального відгуку.

Обґрунтування доцільності використання обраних метрик базується на специфіці задачі аналізу тональності. Точність є важливою, оскільки неправильна класифікація відгуків може призвести до помилкових висновків про ставлення клієнтів до продукту. Наприклад, якщо негативний відгук буде класифіковано як позитивний, це може приховати реальні проблеми продукту від виробника.

Повнота є важливою для того, щоб не пропустити важливі відгуки. Якщо модель має низьку повноту для негативних відгуків, це означає, що багато негативних відгуків будуть класифіковані як нейтральні або позитивні, що призведе до неповної картини проблем продукту. Аналогічно, низька повнота для позитивних відгуків може призвести до недооцінки переваг продукту.

F1-міра особливо корисна для порівняння різних моделей між собою. Оскільки вона враховує і точність, і повноту, модель з вищою F1-мірою загалом краще справляється з класифікацією, ніж модель з нижчою F1-мірою, навіть якщо одна з моделей має вищу точність, а інша – вищу повноту. Це робить F1-міру основною метрикою для вибору найкращої моделі серед альтернатив.

Додатково до основних метрик, для аналізу роботи моделі використовується матриця помилок (confusion matrix). Ця матриця показує, як саме модель класифікує зразки різних класів. Рядки матриці відповідають справжнім класам зразків, а стовпці – передбаченим класам. Елемент матриці на перетині рядка  $i$  та стовпця  $j$  показує кількість зразків класу  $i$ , які модель класифікувала як клас  $j$ .

Аналіз матриці помилок дозволяє виявити специфічні проблеми моделі. Наприклад, якщо матриця показує, що модель часто плутає нейтральні відгуки з

позитивними, але рідко плутає їх з негативними, це вказує на те, що модель має труднощі з розрізненням помірно позитивних та нейтральних висловлювань. Така важлива інформація буде використана для подальшого вдосконалення моделі.

Інтерпретація результатів метрик повинна враховувати контекст задачі. Для задачі аналізу тональності відгуків прийнятними вважаються значення точності вище 75%, точності та повноти вище 70% для кожного класу, та F1-міри вище 70%. Ці пороги базуються на аналізі результатів подібних досліджень та практичних вимогах до систем аналізу тональності.

Важливо також враховувати, що різні типи помилок можуть мати різну вартість у практичних застосуваннях. Наприклад, класифікація негативного відгуку як позитивного може мати більш серйозні наслідки для бізнесу, ніж класифікація нейтрального відгуку як позитивного. Проте в рамках цієї роботи всі типи помилок розглядаються як рівноцінні, оскільки метою є розробка загальної моделі аналізу тональності.

Для оцінки статистичної значущості результатів використовується крос-валідація з 5 фолдами. Датасет розділяється на 5 приблизно рівних частин, і модель навчається п'ять разів, кожного разу використовуючи чотири частини для навчання та одну для валідації. Фінальні метрики обчислюються як середнє значення метрик по всіх п'яти запусках. Це дозволяє оцінити стабільність роботи моделі та зменшити вплив випадкового розподілу даних на результати.

Стандартне відхилення метрик при крос-валідації також розраховується та аналізується. Низьке стандартне відхилення вказує на те, що модель стабільно працює на різних підмножинах даних. Високе стандартне відхилення може свідчити про чутливість моделі до складу навчальної вибірки або про наявність значних відмінностей між різними частинами датасету.

Результати оцінювання моделі представляються у вигляді таблиць та графіків для наочності. Таблиці містять числові значення всіх метрик для кожного класу окремо та усереднені значення. Графіки можуть включати візуалізацію матриці помилок у вигляді теплової карти, де інтенсивність кольору відображає кількість

зразків у кожній комірці матриці. Така візуалізація полегшує ідентифікацію проблемних областей класифікації.

Порівняння з базовими моделями є обов'язковою частиною оцінювання. Як baseline використовується проста модель логістичної регресії з TF-IDF векторизацією без додаткової обробки. Порівняння результатів розробленої моделі з результатами baseline дозволяє оцінити реальну користь від застосування більш складної архітектури та додаткових етапів обробки. Якщо розроблена модель показує результати, які не суттєво відрізняються від baseline, це може вказувати на недоцільність додаткової складності.

Окрім порівняння з baseline, корисно також проаналізувати результати на різних підмножинах тестових даних. Наприклад, можна окремо оцінити точність класифікації на коротких відгуках до 50 слів та довгих відгуках понад 100 слів. Це дозволяє виявити, чи однаково добре модель працює з текстами різної довжини. Аналогічно можна проаналізувати результати на відгуках з різною кількістю емоційно забарвлених слів.

Додатковим критерієм оцінювання є час, необхідний для класифікації одного відгуку. Хоча точність є основним критерієм, практичне застосування моделі також вимагає прийнятної швидкості роботи. Для веб-застосувань бажано, щоб класифікація одного відгуку займала не більше декількох сотень мілісекунд. Вимірювання часу виконання проводиться на тестовій вибірці з усередненням по всіх зразках.

Важливим аспектом оцінювання є аналіз найбільш типових помилок моделі. Для цього відбираються зразки, які модель класифікувала неправильно, та проводиться їх якісний аналіз. Це дозволяє виявити характерні ситуації, в яких модель помиляється, та зрозуміти причини цих помилок. Наприклад, можна виявити, що модель часто помиляється на відгуках, які містять сарказм або іронію, оскільки формально позитивні слова використовуються для вираження негативного ставлення.

Аналіз помилок може також виявити проблеми в самому датасеті. Іноді виявляється, що деякі зразки мають неправильну або неоднозначну розмітку. Наприклад, відгук може бути розмічений як позитивний, хоча фактично містить

змішану тональність з переважанням негативних аспектів. У таких випадках низька точність моделі на конкретних зразках може бути обумовлена не недоліками моделі, а якістю розмітки даних.

Для комплексної оцінки якості моделі також використовується аналіз впевненості передбачень. Модель не просто класифікує відгук у певний клас, а й надає ймовірності належності до кожного з класів. Аналіз розподілу цих ймовірностей може дати додаткову інформацію про роботу моделі. Наприклад, якщо модель часто дає близькі ймовірності для декількох класів, це вказує на невпевненість у класифікації та може свідчити про складність конкретних зразків.

Калібрування ймовірностей є додатковою процедурою, яка може покращити якість оцінок впевненості моделі. Без калібрування ймовірності, що виводяться моделлю, можуть не відповідати реальній впевненості у передбаченні. Процедура калібрування коригує ці ймовірності таким чином, щоб вони краще відображали фактичну точність передбачень. Для калібрування використовується окрема валідаційна вибірка та метод, що базується на ізотонічній регресії.

Оцінка узагальнювальної здатності моделі є критично важливою для практичного застосування. Модель може показувати відмінні результати на тестовій вибірці з того самого датасету, але погано працювати на нових даних з інших джерел. Для перевірки узагальнювальної здатності бажано протестувати модель на додаткових датасетах відгуків, якщо такі доступні. Це дозволяє оцінити, як добре модель має змогу адаптується до нових даних.

Аналіз важливості ознак надає інформацію про те, які слова та біграми найбільше впливають на рішення моделі. Для SVM-класифікатора важливість ознаки можна оцінити за абсолютним значенням відповідного вагового коефіцієнта. Ознаки з великими ваговими коефіцієнтами мають більший вплив на класифікацію. Аналіз топ-20 найважливіших ознак для кожного класу дозволяє зрозуміти, які слова модель асоціює з позитивною, негативною та нейтральною тональністю.

Результати аналізу важливості ознак можуть бути використані для подальшого вдосконалення моделі. Якщо виявляється, що модель надає велику вагу словам, які інтуїтивно не здаються релевантними для визначення тональності, це може вказувати

на проблеми з даними або необхідність додаткової обробки. Наприклад, якщо назви конкретних продуктів отримують високу вагу, це може свідчити про те, що навчена модель навчилася асоціювати певні продукти з певною тональністю, замість того щоб виявляти загальні закономірності вираження емоцій.

Інтерпретація результатів повинна бути обережною та враховувати всі обмеження експерименту. Результати, отримані на конкретному датасеті, можуть не повністю відображати якість роботи моделі на інших необхідних даних. Особливості конкретного датасету, такі як специфічна лексика, стиль написання відгуків або розподіл тем, можуть впливати на результати. Тому висновки про якість моделі повинні формулюватися з урахуванням цих факторів.

## **Висновок до розділу 2**

У другому розділі було описано розроблений метод аналізу тональності відгуків клієнтів з використанням машинного навчання. Описано концепцію підходу, яка базується на послідовній комбінації методу опорних векторів та Dense Neural Network, де друга модель навчається коригувати передбачення першої.

Розроблено архітектуру, яка реалізує техніку stacking та дозволяє нейронній мережі навчитися виправляти помилки базового класифікатора, використовуючи інформацію про впевненість SVM у своїх передбаченнях. Запропоновано модифікацію базової архітектури через введення аналізу біграм на етапі формування ознак.

Теоретично обґрунтовано, що врахування контексту слів через біграми має покращити якість класифікації, особливо для нейтральних відгуків та випадків з заперечними конструкціями. Модифікація є відносно простою в реалізації та не вимагає суттєвого збільшення обчислювальних ресурсів.

Представлено параметри векторизації та обґрунтовано їх вибір. Визначено критерії та метрики для оцінювання якості роботи методу. Обрано чотири основні метрики – точність, повноту, F1-міру та загальну точність – які дозволяють комплексно оцінити якість класифікації.

Обґрунтовано доцільність використання кожної метрики та описано процедуру інтерпретації результатів. Додатково розглянуто додаткові аспекти оцінювання, такі як аналіз матриці помилок, важливості ознак та узагальнювальної здатності моделі. Навчання гібридної моделі відбувається послідовно в два етапи: спочатку навчається SVM-класифікатор на векторах TF-IDF з біграмами, після чого його передбачення використовуються як додаткові ознаки для навчання Dense Neural Network. Це дозволяє другій моделі навчитися метастратегії класифікації, яка враховує не лише текстові ознаки, а й те, як базова модель інтерпретує ці ознаки та наскільки впевнена у своїх рішеннях.

## **Розділ 3 Програмна реалізація методу аналізу тональності відгуків**

### **3.1 Загальна архітектура програмної системи**

Програмна система для аналізу тональності відгуків клієнтів побудована за модульним принципом, що дозволяє кожному компоненту виконувати визначену функцію. Така архітектура забезпечує можливість незалежного тестування окремих частин системи та полегшує подальше вдосконалення методу без необхідності переробляти всю систему. Модульна організація коду є одним з фундаментальних принципів сучасної розробки програмного забезпечення, що дозволяє створювати масштабовані та підтримувані рішення. Вибір модульної архітектури обумовлений кількома важливими факторами. Кожен модуль має визначену відповідальність та мінімальні залежності від інших частин системи.

Це дозволяє розробляти та тестувати модулі незалежно один від одного, що значно спрощує процес розробки. Крім того, модульна структура полегшує внесення змін – якщо потрібно змінити спосіб попередньої обробки тексту, це можна зробити у відповідному модулі без впливу на класифікатори або векторизацію. Система також спроектована з урахуванням можливості масштабування. У майбутньому може з'явитися потреба в обробці більших обсягів даних або додаванні нових функцій. Модульна архітектура дозволяє легко інтегрувати нові компоненти без переробки існуючого коду. Наприклад, можна додати новий алгоритм класифікації, реалізувавши відповідний інтерфейс, і система автоматично зможе його використовувати.

Система складається з 6 основних модулів, кожен з яких відповідає за необхідний етап обробки інформації. Модуль введення даних забезпечує завантаження текстових відгуків з різних джерел та їх первинну валідацію. Модуль попередньої обробки виконує підготовку тексту до подальшого аналізу шляхом токенизації, нормалізації та очищення від непотрібних елементів. Модуль векторизації перетворює оброблений текст у числове представлення, яке може бути використане моделями машинного навчання.

Два класифікаційні модулі реалізують методи визначення тональності на основі опорних векторів та Dense Neural Network відповідно. Модуль обробки результатів відповідає за об'єднання передбачень від різних класифікаторів та формування фінального висновку про тональність відгуку. Структура взаємодії між компонентами системи показана на рисунку 3.1.

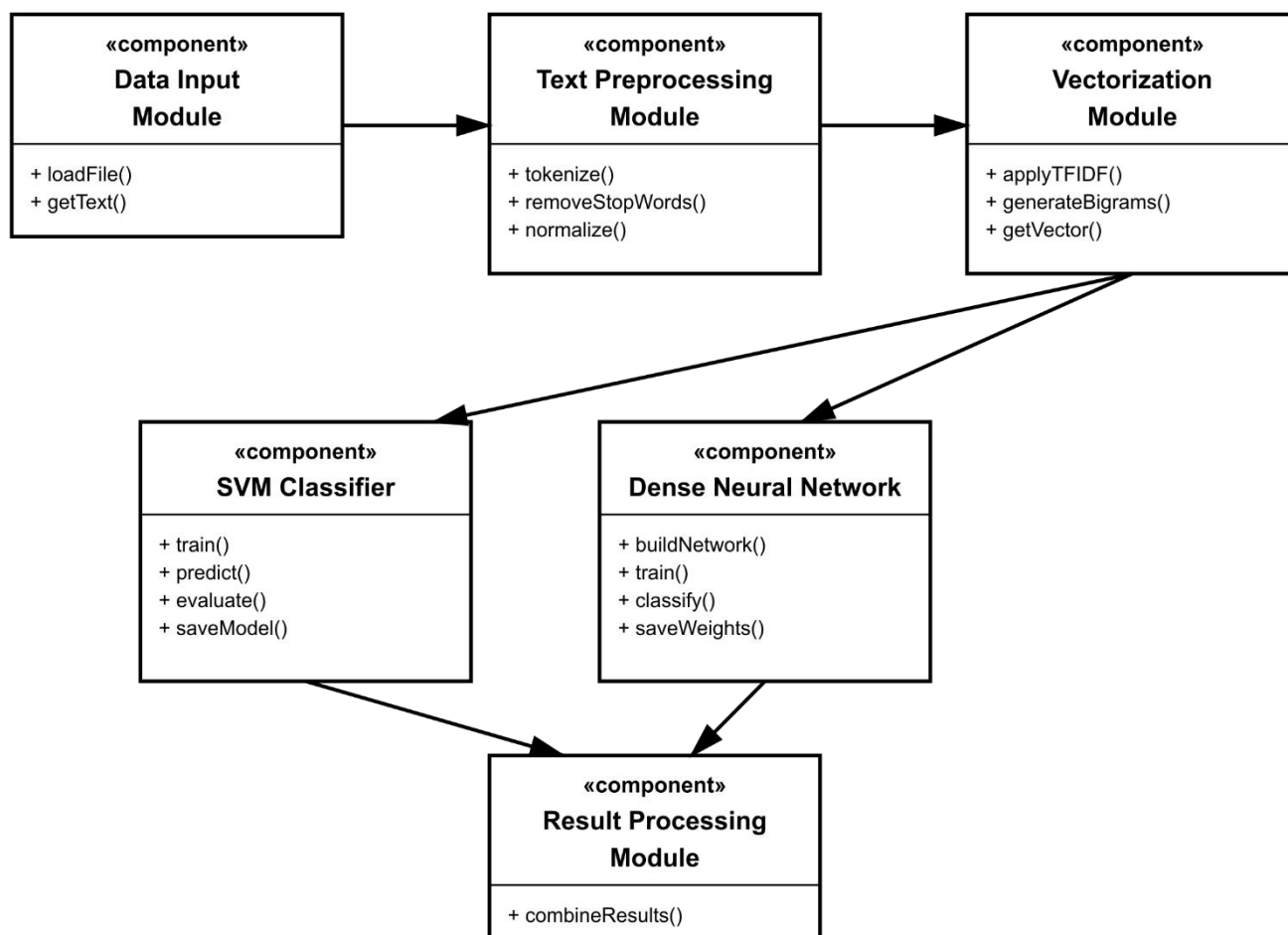


Рисунок 3.1 – Діаграма компонентів методу аналізу тональності відгуків

Дані проходять послідовний шлях від модуля введення через усі етапи обробки до отримання фінального результату класифікації. При цьому деякі модулі можуть взаємодіяти між собою для обміну проміжними результатами або налаштування параметрів. Модуль введення даних реалізує функції завантаження відгуків з текстових файлів або баз даних. Він виконує первинну перевірку коректності даних та забезпечує їх передачу наступним компонентам у

стандартизованому форматі. Модуль підтримує роботу з різними кодуваннями тексту та автоматично визначає формат вхідних файлів. Модуль попередньої обробки містить набір інструментів для підготовки тексту.

Токенізатор розбиває текст на окремі слова та розділові знаки, враховуючи особливості мови та специфічні конструкції. Компонент видалення стоп-слів фільтрує службові слова, що не несуть значущої інформації про тональність. Нормалізатор приводить слова до канонічної форми через стемінг або лематизацію. Додатковий компонент очищення видаляє посилання, спеціальні символи та інші елементи, що можуть заважати аналізу. Модуль векторизації перетворює послідовність токенів у числові вектори фіксованої розмірності. Реалізовано метод TF-IDF, що враховує як частоту слова у документі, так і його рідкість у всьому корпусі текстів.

Додатково модуль формує біграми для захоплення контексту сусідніх слів. Розмірність вихідного вектора становить близько 3500 ознак, що включає окремі слова та найбільш значущі біграми. Класифікаційний модуль на основі методу опорних векторів використовує RBF ядро для побудови нелінійної розділяючої гіперплощини між класами. Модуль навчається на векторних представленнях відгуків з відомою тональністю та формує модель, здатну класифікувати нові дані. Вихідом модуля є три ймовірності належності відгуку до позитивного, негативного та нейтрального класів. Dense Neural Network модуль реалізує мережу з 1 прихованим шаром. Особливістю цього модуля є те, що він приймає на вхід не лише векторне представлення тексту, а й ймовірності від класифікатора на основі опорних векторів. Такий підхід дозволяє мережі навчитися коригувати передбачення базової моделі, використовуючи додаткову інформацію про впевненість першого класифікатора.

Модуль обробки результатів об'єднує передбачення від обох класифікаторів. Оскільки система використовує послідовну архітектуру, цей модуль отримує фінальне передбачення від Dense Neural Network, яка вже врахувала результати класифікатора на основі опорних векторів. Модуль форматує результат у зручному для користувача вигляді, включаючи визначений клас тональності та ступінь впевненості моделі. Взаємодія між модулями організована таким чином, що кожен

компонент отримує дані у визначеному форматі та повертає результати у стандартизованому вигляді.

Це забезпечує слабку зв'язаність модулів та дозволяє легко змінити окремі компоненти без впливання на решту системи. Наприклад, можна змінити метод векторизації або використати інший алгоритм класифікації, не змінюючи модулі попередньої обробки або введення даних.

### **3.2 Реалізація модуля попередньої обробки текстів**

Модуль попередньої обробки текстів є критично важливим компонентом системи, оскільки якість його роботи безпосередньо впливає на точність класифікації. Цей модуль перетворює сирий текст відгуку у чистий набір значущих токенів, які можуть бути використані для дальнішого аналізу. Структура класів модуля попередньої обробки показана на рисунку 3.2.

Центральним елементом є клас `TextPreprocessor`, що координує роботу інших компонентів та забезпечує послідовне виконання всіх етапів обробки. Цей клас містить посилання на токенізатор, компонент видалення стоп-слів, нормалізатор та допоміжні утиліти для очищення. Клас `Tokenizer` відповідає за розбиття тексту на токени. Процес токенізації враховує особливості української мови, зокрема складні слова з дефісом та специфічні знаки пунктуації.

Токенізатор використовує регулярні вирази для виділення слів, чисел та розділових знаків. Він також здатний розбивати текст на речення, що може бути корисно для контекстного аналізу. Мінімальна довжина токена встановлюється для фільтрації надто коротких слів, що зазвичай не несуть значущої інформації. Метод `tokenize` приймає рядок тексту та повертає список токенів. Внутрішньо він спочатку видаляє зайві пробіли, потім застосовує патерн для виділення слів. Метод `splitSentences` дозволяє розбити текст на окремі речення, використовуючи розділові знаки як маркери кінця речення.

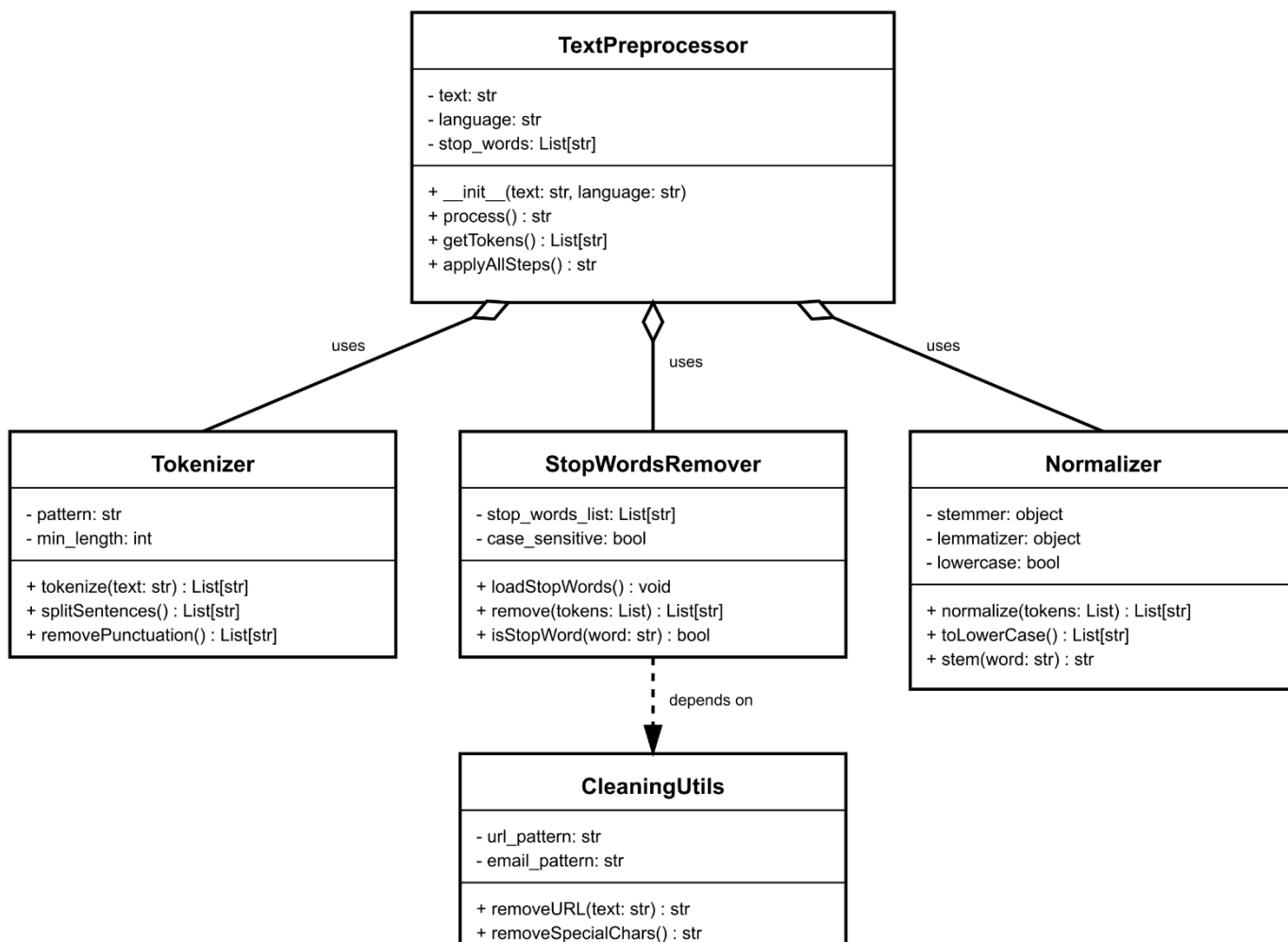


Рисунок 3.2 – Діаграма класів модуля попередньої обробки

Метод `removePunctuation` видаляє розділові знаки з токенів або замінює їх пробілами залежно від налаштувань. Клас `StopWordsRemover` керує списком стоп-слів для української мови. Стоп-слова включають займенники, прийменники, сполучники та інші службові частини мови, що рідко несуть інформацію про тональність. Список стоп-слів завантажується з файла при ініціалізації класу. Метод `loadStopWords` зчитує слова з текстового файла, де кожен рядок містить одне стоп-слово. Метод `isStopWord` перевіряє, чи належить слово до списку стоп-слів. Метод `remove` приймає список токенів та повертає новий список без стоп-слів, зберігаючи порядок решти слів.

Реалізація підтримує налаштування чутливості до регістру. За замовчуванням порівняння виконується без урахування регістру, що дозволяє коректно фільтрувати стоп-слова на початку речень. Це важливо, оскільки слова на початку речень

пишуться з великої літери, але все одно можуть бути службовими частинами мови. Клас `Normalizer` виконує нормалізацію токенів до канонічної форми. Основним методом нормалізації є стемінг, що відсікає закінчення слів для отримання основи. Для української мови використовується адаптований алгоритм стемінгу, що враховує особливості словозміни. Метод `toLowerCase` приводить всі символи до нижнього регістру, що забезпечує однакову обробку слів незалежно від їх позиції у реченні.

Метод `stem` застосовує алгоритм стемінгу до окремого слова, видаляючи відомі закінчення та суфікси. Альтернативою стемінгу є лематизація, що приводить слово до словникової форми. Застосування лематизація дає більш хороші результати, але вимагає словника та морфологічного аналізатора. У реалізації передбачена можливість використання лематизації через інтеграцію зовнішніх бібліотек. Клас `CleaningUtils` містить допоміжні методи для очищення тексту від непотрібних елементів. Метод `removeURL` знаходить та видаляє посилання на веб-сторінки за допомогою регулярного виразу. Метод `removeSpecialChars` видаляє спеціальні символи та емодзі, залишаючи лише літери, цифри та основні розділові знаки.

Це важливо, оскільки спеціальні символи можуть створювати проблеми при векторизації тексту. Основний клас `TextPreprocessor` координує роботу всіх компонентів через метод `process`. Цей метод приймає сирий текст та послідовно застосовує токенізацію, видалення стоп-слів та нормалізацію. Порядок операцій має значення: спочатку виконується очищення тексту, потім токенізація, видалення стоп-слів і нарешті нормалізація. Метод `getTokens` повертає список оброблених токенів, готових до векторизації.

Метод `applyAllSteps` дозволяє застосувати всі етапи обробки за один виклик. Клас зберігає налаштування обробки у своїх полях. Поле `language` визначає мову тексту для вибору відповідного списку стоп-слів та правил нормалізації. Поле `stop_words` містить завантажений набір стоп-слів. Це дозволяє уникнути повторного завантаження списку при обробці кількох відгуків. Модуль реалізовано з урахуванням принципів об'єктно-орієнтованого програмування. Кожен клас має визначену відповідальність та мінімальні залежності від інших класів. Це дозволяє

легко тестувати окремі компоненти та вносити зміни без впливу на іншу частину модуля.

Наприклад, можна замінити реалізацію токенизатора, не змінюючи клас `TextPreprocessor`. Модуль попередньої обробки забезпечує високу якість підготовки тексту для подальшого аналізу. Коректна токенизація дозволяє точно виділяти слова з урахуванням особливостей мови. Видалення стоп-слів зменшує розмірність даних та покращує співвідношення сигнал-шум. Нормалізація об'єднує різні форми одного слова, що допомагає моделі узагальнювати закономірності.

### 3.3 Реалізація модуля класифікації

Модуль класифікації містить реалізації різних алгоритмів визначення тональності та логіку їх комбінування у гібридну модель. Архітектура модуля показана на рисунку 3.3, де представлені класи базового класифікатора, конкретних реалізацій алгоритмів та компонента для оцінювання якості. Ієрархія класів побудована на основі абстрактного базового класу `BaseClassifier`, що визначає спільний інтерфейс для всіх класифікаторів. Цей клас містить абстрактні методи `train` та `predict`, які повинні бути реалізовані у похідних класах. Поле `model` зберігає навчену модель класифікатора. Поле `is_trained` вказує, чи була модель навчена та готова до використання.

Поле `params` містить словник параметрів моделі. Використання базового класу забезпечує єдиний інтерфейс для роботи з алгоритмами класифікації. Це дозволяє легко додавати нові методи класифікації без зміни коду, що використовує класифікатори. Також це полегшує порівняння різних алгоритмів, оскільки всі вони мають однакові методи для навчання та передбачення. Клас `SVMClassifier` успадковує `BaseClassifier` та реалізує необхідну класифікацію на основі застосування методу опорних векторів. Поля `kernel`, `C` та `gamma` зберігають параметри моделі. Поле `kernel` визначає тип ядра для трансформації даних у вищерозмірний простір. У реалізації використовується радіально-базисне ядро, що добре підходить для нелінійно роздільних даних.

Параметр контролює баланс між максимізацією зазору між класами та мінімізацією помилок класифікації. Параметр `gamma` визначає радіус впливу окремих прикладів на формування розділяючої гіперплощини. Метод `train` приймає навчальні дані у вигляді векторів ознак та відповідних міток класів. Він створює модель методу опорних векторів з заданими параметрами та навчає її на наданих даних. Процес навчання полягає у пошуку оптимальної розділяючої гіперплощини, що максимізує відстань до найближчих точок кожного класу. Метод `predict` приймає вектори ознак нових відгуків та повертає передбачені мітки класів. Додатково він може повертати ймовірності належності до кожного класу, що важливо для гібридної моделі.

Метод `setParameters` дозволяє змінити параметри моделі перед навчанням. Метод `getWeights` повертає ваги опорних векторів, що може бути корисно для аналізу важливості ознак. Клас `DenseNNClassifier` реалізує класифікацію за допомогою `Dense Neural Network`. Поле `hidden_layers` зберігає конфігурацію прихованих шарів мережі. У реалізації використовується один прихований шар зі 100 нейронами. Поле `activation` визначає функцію активації для прихованого шару. Функція використовується `ReLU`, що добре зарекомендувала себе для задач класифікації.

Поле `optimizer` зберігає алгоритм оптимізації ваг мережі. Використовується алгоритм `Adam`, що адаптивно налаштовує швидкість навчання. Поле `epochs` визначає кількість проходів через навчальну вибірку. Метод `buildNetwork` створює архітектуру нейронної мережі згідно з налаштуваннями. Він визначає у вхідному шарі кількість нейронів на основі розмірності вхідних даних, додає прихований шар з заданою функцією активації та вихідний шар з функцією `softmax` для отримання ймовірностей класів. Метод `train` виконує навчання мережі на наданих даних. Він ітеративно проходить через навчальну вибірку, обчислює помилку передбачення та коригує ваги нейронів для мінімізації цієї помилки. Використовується функція втрат, що підходить для багатокласової класифікації. Метод `predict` використовує навчену мережу для класифікації нових відгуків. Метод `getArchitecture` повертає опис архітектури мережі.

Метод `saveWeights` зберігає навчені ваги мережі у файл для дальнішого використання. Клас `HybridClassifier` реалізує гібридну модель, що комбінує метод опорних векторів та `Dense Neural Network` у послідовну архітектуру. Поля `svm_model`

та `nn_model` зберігають екземпляри відповідних класифікаторів. Поле `weights` може містити вагові коефіцієнти для комбінування передбачень, хоча у послідовній архітектурі вони не використовуються безпосередньо. Метод `train` виконує навчання гібридної моделі у два етапи. На першому етапі навчається класифікатор на основі методу опорних векторів на векторних представленнях відгуків.

Після завершення навчання цей класифікатор застосовується до всіх навчальних даних для отримання ймовірностей класів. На другому етапі ці ймовірності об'єднуються з оригінальними векторами ознак, формуючи розширений набір ознак для мережі. `Dense Neural Network` навчається на цьому розширеному наборі, вчачись коригувати передбачення базової моделі. Метод `predict` використовує послідовну обробку для класифікації нових відгуків. Спочатку векторне представлення відгуку подається на вхід класифікатора на основі опорних векторів, який формує три ймовірності. Ці ймовірності об'єднуються з оригінальним вектором ознак.

Отриманий розширений вектор подається на вхідний блок мережі, яка формує фінальне передбачення. Метод `combinePredictions` реалізує логіку об'єднання: він створює новий вхідний вектор шляхом конкатенації оригінальних ознак та ймовірностей від першого класифікатора. Метод `optimizeWeights` може використовуватися для налаштування параметрів комбінування на валідаційній вибірці, хоча у послідовній архітектурі це не є критичним. Клас `ModelEvaluator` містить методи для оцінювання якості роботи класифікаторів. Поле `metrics` зберігає обчислені значення метрик.

Поле `confusion_matrix` містить матрицю помилок класифікації. Метод `evaluate` приймає передбачення моделі та справжні мітки класів, обчислює різні метрики якості. Метод `calculateMetrics` обчислює точність, повноту, F1-міру для кожного класу та загальну точність моделі. Метод `displayResults` форматує та виводить результати оцінювання у зручному вигляді. Матриця помилок є таблицею, де рядки відповідають справжнім класам, а стовпці передбаченим. Діагональні елементи показують кількість правильно класифікованих зразків кожного класу. Недіагональні елементи

показують помилки класифікації. Аналіз матриці помилок дозволяє зрозуміти, які класи модель часто плутає між собою.

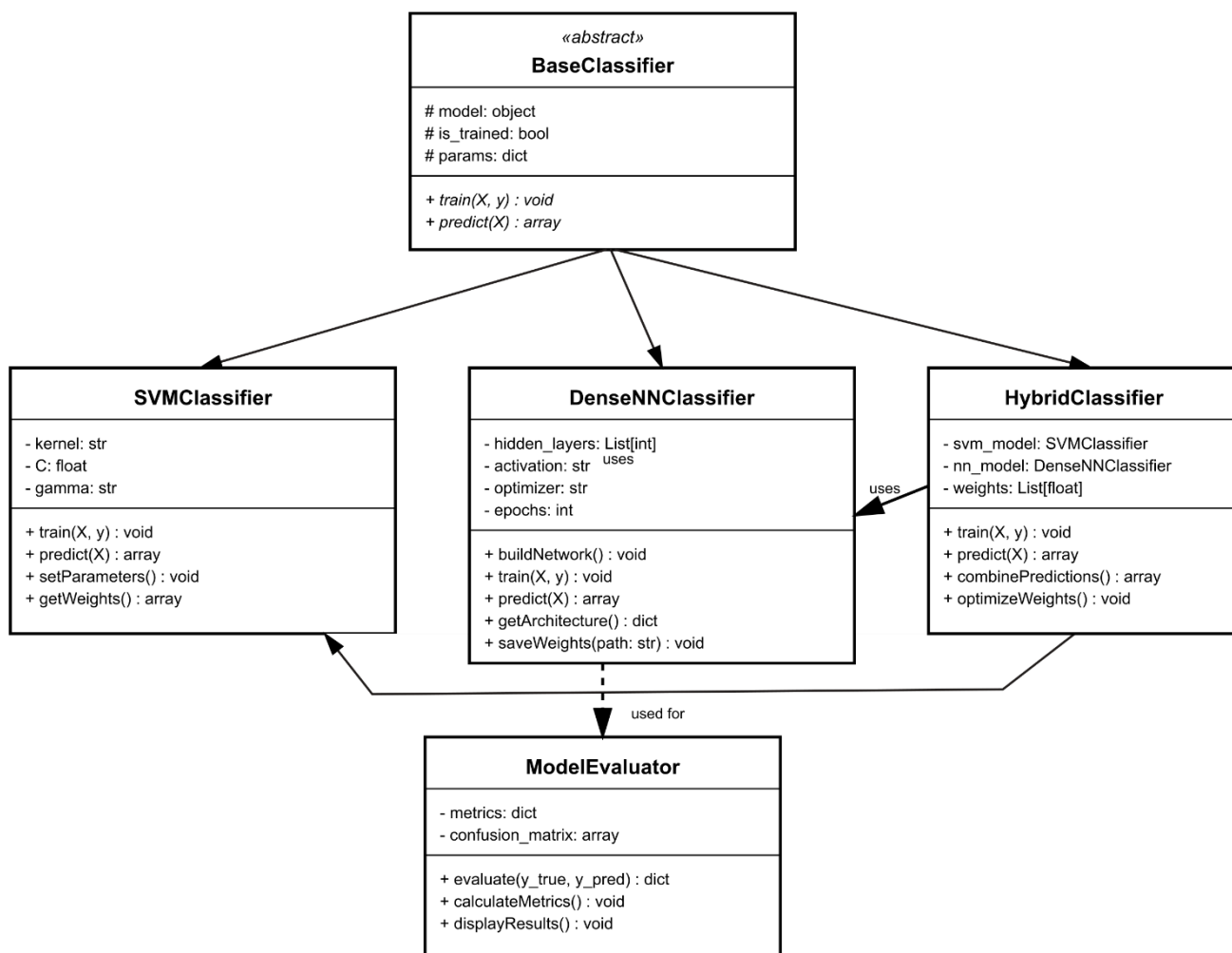


Рисунок 3.3 – Діаграма класів модуля класифікації

Архітектура модуля класифікації забезпечує гнучкість у виборі та комбінуванні алгоритмів. Базовий клас визначає спільний інтерфейс, що дозволяє легко додавати нові методи класифікації. Гібридний класифікатор може бути розширений для використання більшої кількості базових моделей або інших стратегій комбінування. Компонент оцінювання забезпечує єдиний спосіб вимірювання якості для всіх класифікаторів.

### 3.4 Реалізація модуля векторизації тексту

Модуль векторизації відповідає за перетворення послідовності токенів у числові вектори фіксованої розмірності, які можуть бути використані алгоритмами машинного навчання. Структура класів цього модуля представлена на рисунку 3.4.

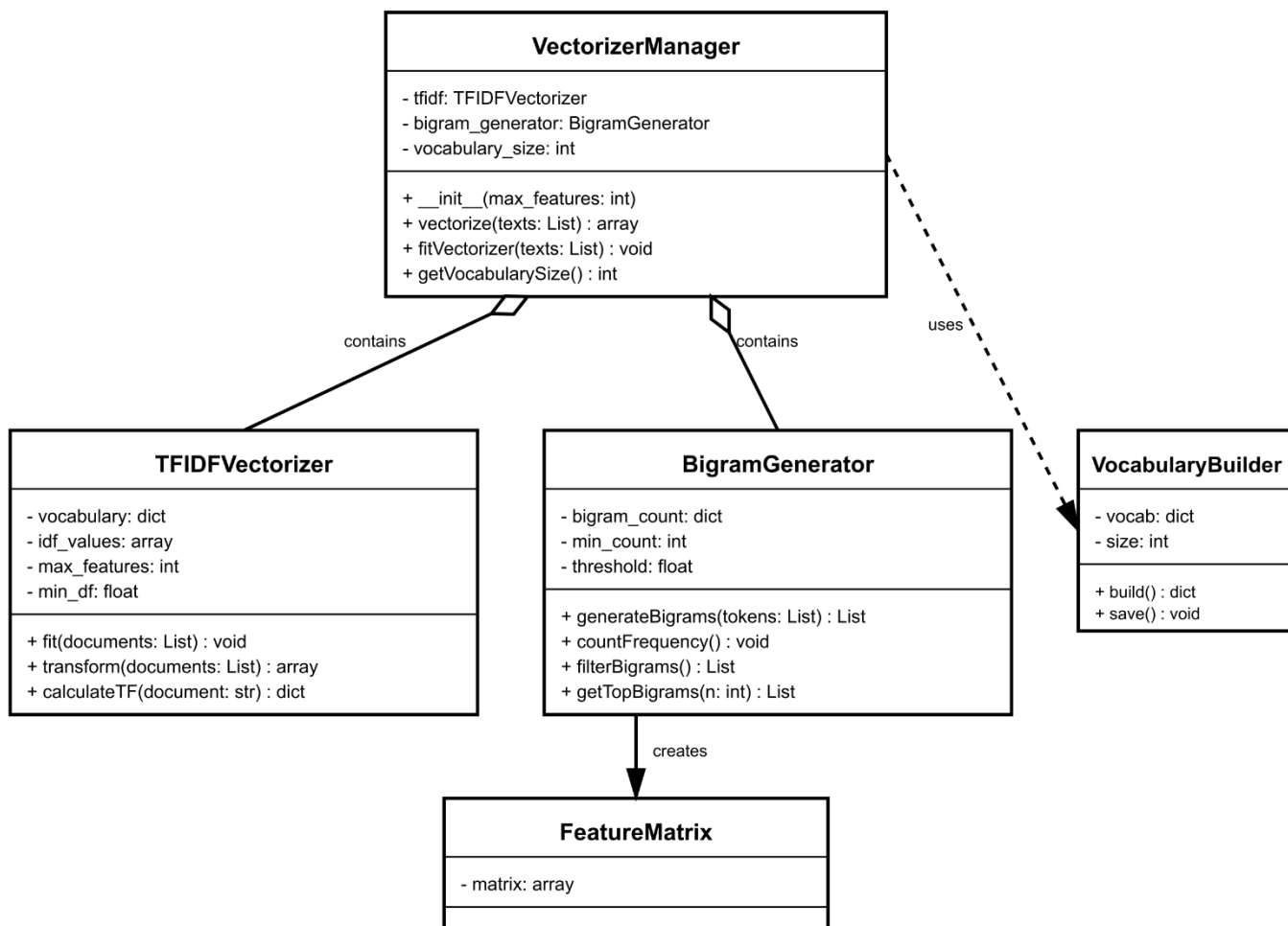


Рисунок 3.4 – Діаграма класів модуля векторизації

Центральним класом модуля є **VectorizerManager**, що координує процес векторизації та керує іншими компонентами. Поле **tfidf** зберігає екземпляр векторизатора TF-IDF. Поле **bigram\_generator** містить генератор біграм. Поле **vocabulary\_size** зберігає розмір словника після побудови. Метод **vectorize** приймає список токенів та повертає відповідний числовий вектор. Метод **fitVectorizer** навчає векторизатор на колекції документів, будуючи словник та обчислюючи статистики.

Метод **getVocabularySize** повертає кількість унікальних ознак у словнику. Клас **TFIDFVectorizer** реалізує метод векторизації TF-IDF. Поле **vocabulary** зберігає

словник, що відображає токени на індекси у векторі. Поле `idf_values` містить обчислені значення `inverse document frequency` для кожного токена. Поле `max_features` обмежує максимальну кількість ознак у векторі. Поле `min_df` визначає мінімальну частоту документа для включення токена до словника. Токени, що зустрічаються надто рідко, ігноруються як нерепрезентативні. Метод `fit` будує словник на основі колекції документів. Він проходить через усі документи, збирає унікальні токени та обчислює їх частоти. Потім він відбирає найбільш значущі токени згідно з параметрами `max_features` та `min_df`. Для кожного токена обчислюється `inverse document frequency` як логарифм відношення загальної кількості документів до кількості документів, що містять цей токен. Метод `transform` перетворює документ у TF-IDF вектор. Він обчислює частоту кожного токена у документі та множить її на відповідне значення `inverse document frequency`. Результатом є розріджений вектор, де більшість значень дорівнюють нулю. Метод `calculateTF` обчислює частоту терміна як відношення кількості появ токена до загальної кількості токенів у документі. Використання TF-IDF дозволяє враховувати як важливість токена у конкретному документі, так і його специфічність для цього документа відносно всього корпусу.

Токени, що зустрічаються у багатьох документах, отримують менші ваги, оскільки вони менш інформативні для розрізнення між документами. Токени, специфічні для невеликої кількості документів, отримують більші ваги. Клас `BigramGenerator` створює біграми з послідовності токенів. Біграма – це пара послідовних токенів, що може захоплювати локальний контекст слів. Поле `bigram_count` зберігає частоти біграм у корпусі. Поле `min_count` визначає мінімальну частоту для включення біграми до результату. Поле `threshold` може використовуватися для фільтрації біграм на основі статистичної значущості. Метод `generateBigrams` приймає список токенів та генерує всі можливі біграми. Він проходить через список токенів з кроком один, об'єднуючи кожну пару сусідніх токенів через підкреслення або інший роздільник.

Метод `countFrequency` підраховує частоту кожної біграми у колекції документів. Метод `filterBigrams` відбирає найбільш значущі біграми згідно з заданими критеріями. Метод `getTopBigrams` повертає найчастіші біграми у відсортованому

порядку. Включення біграм до векторного представлення дозволяє частково враховувати порядок слів та контекст. Наприклад, біграма "не\_добре" має протилежне значення порівняно з окремим словом "добре".

Без біграм ці два випадки мали б подібні векторні представлення, що могло б призвести до помилок класифікації. Клас `VocabularyBuilder` відповідає за побудову та збереження словника. Поле `vocab` містить відображення токенів на індекси. Поле `size` зберігає розмір словника. Метод `build` створює словник на основі токенів з навчальної вибірки. Він відбирає найбільш частотні токени та призначає їм унікальні індекси. Метод `save` зберігає словник у файл для подальшого використання. Це дозволяє використовувати той самий словник при векторизації нових відгуків. Клас `FeatureMatrix` представляє результат векторизації колекції документів. Поле `matrix` зберігає числову матрицю, де кожен рядок відповідає одному документу, а кожен стовпець – одній ознаці.

Матриця зазвичай є розрідженою, оскільки більшість документів містять лише невелику частину всіх можливих токенів. Процес векторизації виконується наступним чином. Спочатку `VectorizerManager` приймає колекцію токенізованих документів для навчання. Він передає їх `BigramGenerator` для створення біграм. Потім об'єднаний набір токенів та біграм передається `TFIDFVectorizer` для побудови словника та обчислення статистик. Після завершення навчання `VectorizerManager` може векторизувати нові документи, послідовно застосовуючи генерацію біграм та TF-IDF трансформацію. Розмірність вихідного вектора визначається розміром словника, який встановлюється параметром `max_features`. У реалізації використовується близько 3000 найбільш частотних слів та 500 найбільш значущих біграм, що дає загальну розмірність близько 3500 ознак. Така розмірність забезпечує достатнє представлення тексту без надмірного збільшення обчислювальної складності.

Модуль векторизації забезпечує трансформацію тексту у формат, придатний для алгоритмів машинного навчання. TF-IDF враховує як локальну, так і глобальну важливість токенів. Біграми додають інформацію про контекст та порядок слів. Розріджене представлення економить пам'ять та прискорює обчислення.

### 3.5 Взаємодія компонентів системи у процесі класифікації

Процес класифікації відгуку включає взаємодію всіх компонентів системи у певній послідовності. Ця взаємодія показана на діаграмі взаємодії на рисунку 3.5, де представлені об'єкти різних класів та повідомлення, що передаються між ними.

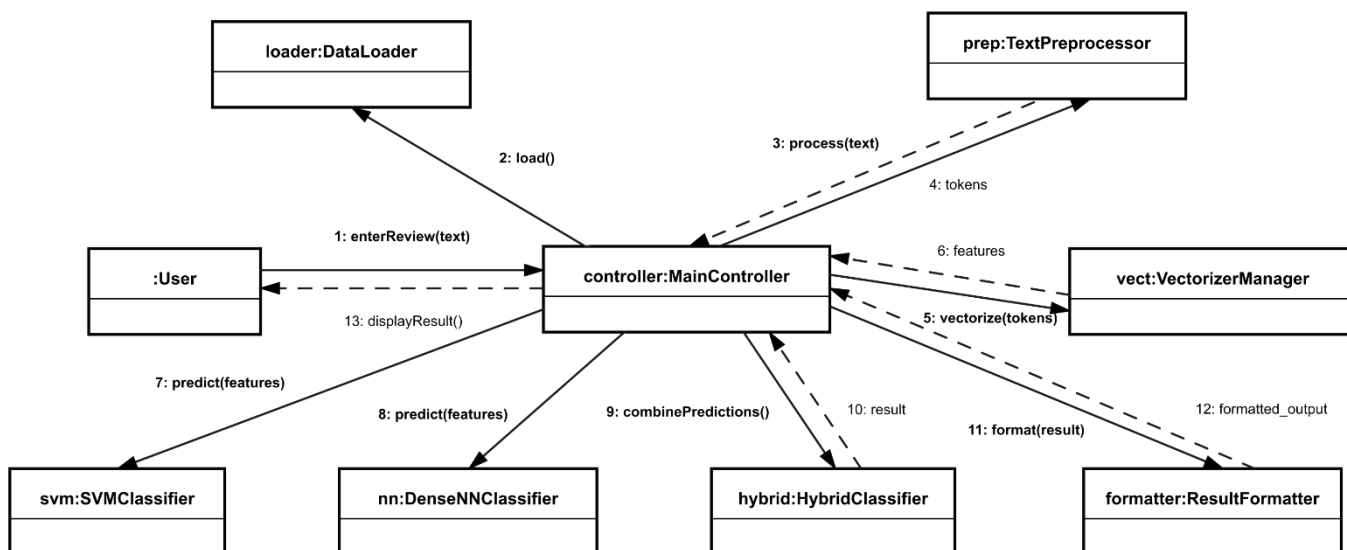


Рисунок 3.5 – Діаграма взаємодії елементів методу

Користувач ініціює процес класифікації, надаючи текст відгуку головному контролеру системи через метод `enterReview`. Контролер відповідає за координацію роботи всіх компонентів та забезпечення правильної послідовності операцій. Контролер спочатку звертається до завантажувача даних через метод `load`. Завантажувач може виконувати додаткові операції, такі як валідація формату даних або перетворення кодування. Після завантаження контролер отримує дані у стандартизованому форматі. Наступним кроком є попередня обробка тексту. Контролер викликає метод `process` об'єкта препроцесора, передаючи йому текст відгуку.

Препроцесор послідовно застосовує токенизацію, видалення стоп-слів та нормалізацію. Результатом є список очищених токенів, який повертається контролеру. Контролер передає токени менеджеру векторизації через метод `vectorize`.

Менеджер векторизації спочатку генерує біграми з токенів, потім об'єднує токени та біграми в один список. Цей список трансформується у TF-IDF вектор за допомогою навченого векторизатора. Отриманий числовий вектор повертається контролеру як набір ознак для класифікації. Контролер передає вектор ознак першому класифікатору на основі методу векторів шляхом виклику методу `predict`.

Класифікатор аналізує ознаки та обчислює три ймовірності належності відгуку до позитивного, негативного та нейтрального класів. Ці ймовірності базуються на положенні точки у просторі ознак відносно розділяючої гіперплощини, побудованої під час навчання. Паралельно або послідовно контролер також викликає метод `predict` класифікатора на основі `Dense Neural Network`, передаючи йому той самий вектор ознак. У послідовній архітектурі цей виклик відбувається після отримання результату від класифікатора на основі опорних векторів, оскільки нейронна мережа потребує ймовірностей від першого класифікатора як додаткових ознак.

Гібридний класифікатор об'єднує результати обох класифікаторів через метод `combinePredictions`. У послідовній архітектурі цей метод формує розширений вектор ознак, що включає оригінальні TF-IDF ознаки та ймовірності від класифікатора на основі опорних векторів. Цей розширений вектор подається на вхід мережі, яка формує фінальне передбачення. Гібридний класифікатор повертає контролеру визначений клас тональності разом з відповідними ймовірностями. Контролер передає результат класифікації формату результату через метод `format`. Форматер перетворює внутрішнє представлення результату у зручний для користувача формат. Це може включати текстове пояснення визначеної тональності, числові значення ймовірностей та додаткову інформацію про впевненість моделі.

Відформатований результат повертається контролеру. Нарешті, контролер передає відформатований результат користувачеві через метод `displayResult`. Користувач отримує інформацію про тональність відгуку разом з показниками впевненості моделі у своєму рішенні. Вся послідовність взаємодій є синхронною – кожен компонент чекає завершення попередньої операції перед початком наступної. Це спрощує логіку обробки та забезпечує передбачуваність поведінки системи. Помилки на будь-якому етапі можуть бути перехоплені контролером та оброблені

належним чином. Діаграма співпраці показує потік даних через систему. Текст відгуку послідовно трансформується: від сирого тексту до токенів, від токенів до числового вектора, від вектора до ймовірностей класів, і нарешті до фінального результату класифікації.

Кожна трансформація виконується спеціалізованим компонентом, що володіє необхідними знаннями та алгоритмами. Контролер виступає в ролі координатора, що знає про існування всіх компонентів та порядок їх виклику. Інші компоненти не знають один про одного безпосередньо – вони взаємодіють лише через контролер. Така архітектура забезпечує слабку зв'язаність та полегшує тестування окремих компонентів. Повідомлення між об'єктами передають дані у стандартизованих форматах. Препроцесор повертає список токенів. Векторизатор повертає числовий вектор.

Класифікатори повертають ймовірності або мітки класів. Така стандартизація дозволяє легко замінювати реалізації компонентів, якщо вони дотримуються встановлених інтерфейсів. Система спроектована з урахуванням можливості масової обробки відгуків. Хоча діаграма показує обробку одного відгуку, той самий процес може бути застосований до колекції відгуків. Навчені моделі векторизації та класифікації зберігаються у пам'яті між викликами, що дозволяє ефективно обробляти велику кількість відгуків без повторного завантаження моделей. Взаємодія компонентів організована так, що забезпечує як коректність обробки, так і зручність використання системи. Користувачу достатньо надати текст відгуку та отримати результат вся складність внутрішньої обробки прихована за простим інтерфейсом контролера.

### **Висновки до розділу 3**

У третьому розділі описано програмну реалізацію розробленого методу аналізу тональності відгуків клієнтів. Система побудована за модульним принципом, що забезпечує розділення залежності між компонентами та дозволяє незалежно розробляти та тестувати окремі частини. Загальна архітектура системи включає шість основних модулів: введення даних, попередньої обробки тексту, векторизації, два

класифікаційні модулі та модуль обробки результатів. Кожен модуль відповідає за етап оброблення інформації та взаємодіє з іншими модулями через визначені інтерфейси.

Модуль попередньої обробки реалізовано через набір спеціалізованих класів, що виконують токенізацію, видалення стоп-слів, нормалізацію та очищення тексту. Використання окремих класів для кожної операції забезпечує гнучкість та можливість налаштування окремих етапів обробки. Координація роботи компонентів здійснюється через центральний клас препроцесора.

Модуль класифікації побудовано на основі ієрархії класів з базовим абстрактним класифікатором. Реалізовано три конкретні класифікатори: на основі методу опорних векторів, на основі Dense Neural Network та гібридний класифікатор, що об'єднує перші два у послідовну архітектуру. Така організація дозволяє легко додавати нові алгоритми класифікації без зміни решти системи. Модуль векторизації забезпечує перетворення послідовності токенів у числові вектори через метод TF-IDF з додаванням біграм. Реалізація враховує як частоту токенів у документі, так і їх специфічність для документа. Біграми додають інформацію про контекст та порядок слів, що покращує якість класифікації.

Взаємодія компонентів організована таким чином, що забезпечує послідовну обробку відгуку від сирого тексту до фінального результату класифікації. Контролер координує роботу всіх модулів, забезпечуючи правильний порядок операцій та передачу даних між компонентами. Така архітектура дозволяє ефективно обробляти як окремі відгуки, так і великі колекції.

## **Розділ 4 Експериментальні дослідження методу аналізу тональності відгуків**

### **4.1 Формування набору даних для експериментальних досліджень**

Для проведення експериментальних досліджень запропонованого методу аналізу тональності було сформовано набір даних, який складається з відгуків клієнтів про різні системи та послуги. Вибір даних здійснювався з урахуванням необхідності забезпечити репрезентативність вибірки та збалансованість класів тональності.

Початковий датасет містив 8000 відгуків, зібраних з різних джерел. Кожен відгук було вручну розмічено експертами за трьома класами тональності: позитивний, негативний та нейтральний. Для забезпечення якості розмітки кожен відгук переглядали два незалежні експерти, а у випадку розбіжностей залучався третій експерт для остаточного рішення. Такий підхід дозволив мінімізувати суб'єктивність розмітки та забезпечити високу якість навчальних даних.

Відгуки з неоднозначною або суперечливою тональністю були виключені, оскільки вони можуть негативно впливати на навчання моделей. Відібрані відгуки рівномірно розподілені між трьома класами. Така збалансованість важлива для того, щоб модель не була схильна до передбачення одного класу частіше за інші.

Довжина відгуків у наборі даних варіюється від 10 до 200 слів, що відповідає типовій довжині реальних відгуків клієнтів. Короткі відгуки зазвичай містять лаконічні оцінки, такі як "чудовий продукт" або "повне розчарування". Довші відгуки включають детальний опис досвіду використання продукту з поясненням конкретних переваг або недоліків. Розмаїття довжин відгуків дозволяє перевірити здатність моделі працювати з текстами різної деталізації.

Навчальна вибірка використовується безпосередньо для навчання моделей класифікації. На цих даних моделі вчать розпізнавати закономірності, які вказують на тональність тексту. Валідаційна вибірка застосовується для налаштування гіперпараметрів моделей та контролю процесу навчання. Вона дозволяє виявити момент, коли модель починає перенавчатися на навчальних даних та втрачає

здатність до узагальнення. Тестова вибірка залишається повністю незалежною від процесу навчання та використовується лише для фінальної оцінки якості моделей.

Перед подачею на вхід моделей всі відгуки проходять етап попередньої обробки. Текст приводиться до нижнього регістру для уніфікації написання слів. Видаляються спеціальні символи, URL-адреси, email-адреси та інші нетекстові елементи, які не несуть інформації про тональність. Виконується токенизація тексту на окремі слова з урахуванням знаків пунктуації. Стоп-слова, такі як прийменники, сполучники та артиклі, видаляються з тексту, оскільки вони зустрічаються у відгуках будь-якої тональності однаково часто.

Після попередньої обробки тексти векторизуються за допомогою методу TF-IDF. Розмір словника обмежено 3000 найбільш частотних слів, що забезпечує баланс між повнотою представлення тексту та компактністю векторів. До словника також додаються біграми – послідовності з двох сусідніх слів. Включення біграм дозволяє частково захопити контекст слів та їх взаємозв'язки. Загальна розмірність векторного представлення після додавання біграм становить приблизно 3500 ознак.

Для оцінки статистичних характеристик датасету було проаналізовано розподіл довжин відгуків, частотність слів та кореляцію між окремими ознаками. Середня довжина відгуку становить 68 слів з стандартним відхиленням 34 слова. Позитивні відгуки у середньому трохи довші 72 слова, ніж негативні 66 слів та нейтральні 65 слів. Це можна пояснити тим, що задоволені клієнти частіше описують конкретні позитивні аспекти продукту, тоді як негативні відгуки можуть бути лаконічнішими.

Аналіз частотності слів виявив, що деякі слова сильно корелюють з певними класами тональності. Наприклад, слова "чудовий", "відмінний", "рекомендую" частіше зустрічаються у позитивних відгуках, тоді як слова "поганий", "розчарування", "жахливо" характерні для негативних відгуків. Нейтральні відгуки часто містять слова "нормально", "звичайно", "середньо". Наявність таких індикаторних слів підтверджує можливість автоматичної класифікації тональності на основі текстових ознак.

Важливим аспектом підготовки даних є забезпечення відсутності витoku інформації між навчальною та тестовою вибірками. Розподіл даних виконано таким чином, що відгуки про один і той самий продукт або від одного користувача не потрапляють одночасно у навчальну та тестову вибірки. Це гарантує, що модель оцінюється на справді нових даних, які вона не бачила під час навчання.

Також було проаналізовано потенційні джерела помилок у даних. Виявлено, що близько 3% відгуків містять елементи сарказму або іронії, коли формально позитивні слова використовуються для вираження негативного ставлення. Такі випадки є складними для будь-якого методу автоматичного аналізу тональності. Ще приблизно 2% відгуків мають змішану тональність, коли в одному тексті присутні як позитивні, так і негативні висловлювання. Для таких відгуків клас тональності визначався за загальним враженням, яке переважає у тексті.

Сформований набір даних забезпечує надійну основу для експериментальних досліджень запропонованого методу. Збалансованість класів, розмаїття довжин та змісту відгуків, а також ретельна підготовка та розмітка даних дозволяють отримати нелужну оцінку якості роботи класифікації.

## **4.2 Налаштування параметрів базової моделі SVM**

Базова модель класифікації реалізована на основі методу опорних векторів з радіально-базисною функцією ядра. Вибір SVM обумовлений здатністю цього методу ефективно працювати з високорозмірними даними та знаходити оптимальну розділяючу гіперплощину між класами. RBF-ядро дозволяє моделювати нелінійні залежності між ознаками, що важливо для задачі класифікації тексту.

SVM-класифікатор має два основні гіперпараметри, які суттєво впливають на якість роботи моделі. Параметр контролює баланс між максимізацією зазору між класами та мінімізацією помилок класифікації на навчальних даних. Менші значення призводять до більш гладкої розділяючої поверхні, що може покращити узагальнення, але знизити точність на навчальних даних. Більші значення дозволяють

моделі більш точно підлаштовуватися під навчальні дані, але можуть призвести до перенавчання.

Параметр  $\gamma$  визначає, наскільки далеко впливає один навчальний зразок. Малі значення  $\gamma$  означають, що кожен зразок впливає на велику область простору ознак, що призводить до більш гладкої моделі. Великі значення  $\gamma$  роблять модель більш чутливою до окремих навчальних зразків, що може покращити точність на навчальних даних, але знизити узагальнювальну здатність.

Для підбору оптимальних значень параметрів  $C$  та  $\gamma$  використано метод перехресної перевірки на валідаційній вибірці. Параметр  $C$  варіювався у діапазоні від 0.1 до 10 з кроком множення на 2. Параметр  $\gamma$  варіювався у діапазоні від 0.001 до 1 також з кроком множення на 2. Для кожної комбінації параметрів модель навчалася на навчальній вибірці та оцінювалася на валідаційній вибірці.

Результати експериментів показали, що оптимальними є значення  $C = 1.0$  та  $\gamma = 0.1$ . При цих параметрах модель досягла найвищої точності на валідаційній вибірці. Менші значення  $\gamma$  0.001 та 0.01 призводили до недостатньої гнучкості моделі та нижчої точності. Більші значення  $\gamma$  0.5 та 1.0 спричиняли перенавчання, що проявлялося у високій точності на навчальній вибірці, але низькій точності на валідаційній.

Аналогічно, значення  $C$  менше 1.0 давали занадто гладку модель з недостатньою точністю, а значення більше 1.0 призводили до перенавчання. Оптимальне значення  $C = 1.0$  забезпечує баланс між точністю та узагальнювальною здатністю.

Після визначення оптимальних параметрів модель було навчено на повній навчальній вибірці та оцінено на тестовій вибірці. Базова SVM-модель досягла точності 86.2% на тестових даних. Це є хорошим базовим результатом, який свідчить про здатність методу опорних векторів ефективно класифікувати тональність текстових відгуків.

Аналіз помилкових класифікацій виявив декілька типових випадків. SVM-класифікатор часто помиляється на відгуках, які містять заперечення, такі як "не добре" або "не рекомендую". Модель може ігнорувати заперечну частку та

класифікувати відгук на основі позитивного слова "добре" або "рекомендую". Також складними є відгуки з умовними конструкціями, наприклад "було б добре, якби". Ще одна категорія помилок пов'язана з відгуками, де тональність виражена не через конкретні слова, а через загальний тон або структуру висловлювання.

Детальний аналіз метрик якості для кожного класу представлено у таблиці 4.1. Для позитивного класу precision становить 0.868, що означає правильність 86.8% усіх передбачень, класифікованих як позитивні. Recall для цього класу дорівнює 0.875, тобто модель знайшла 87.5% всіх справжніх позитивних відгуків. F1-Score, як гармонійне середнє precision та recall, склав 0.871.

Негативний клас характеризується precision 0.856 та recall 0.865, що дає F1-Score 0.860. Для нейтрального класу метрики трохи нижчі: precision 0.842, recall 0.845, F1-Score 0.843. Нижчі показники для нейтрального класу підтверджують його складність для класифікації.

Таблиця 4.1 – Метрики якості базової SVM-моделі по класах тональності

Клас тональності	Precision	Recall	F1-Score
Позитивні	0.868	0.875	0.871
Негативні	0.856	0.865	0.860
Нейтральні	0.842	0.845	0.843
Середнє	0.855	0.862	0.858

Загальна точність базової SVM-моделі 86.2% є прийнятним результатом для задачі тризначної класифікації тональності. Це значення показує, що модель правильно класифікує більше 4 з 5 відгуків. Однак є простір для покращення, особливо у класифікації нейтральних відгуків та відгуків із складними лінгвістичними конструкціями.

Базова SVM-модель служить відправною точкою для подальших експериментів. Результати, отримані на цій моделі, дозволяють оцінити покращення, які забезпечують модифікації методу, такі як додавання біграм та використання гібридної архітектури з Dense Neural Network.

### 4.3 Процес навчання Dense Neural Network

Dense Neural Network було реалізовано як другий етап гібридної моделі класифікації. Архітектура нейронної мережі включає вхідний шар розмірністю 3503 нейрони, один прихований шар з 100 нейронами та вихідний шар з 3 нейронами для трьох класів тональності.

Вхідний вектор для нейронної мережі формується шляхом конкатенації оригінальних TF-IDF ознак з біграмами 3500 ознак та ймовірностей, отриманих від SVM-класифікатора з ознаки. Таким чином, нейронна мережа має доступ як до низькорівневих текстових ознак, так і до високорівневих передбачень базової моделі.

Прихований шар використовує функцію ReLU Rectified Linear Unit, яка добре зарекомендувала себе для задач класифікації тексту. ReLU вводить нелінійність у модель, дозволяючи їй виявляти складні залежності між ознаками. Функція визначається як максимум між нулем та вхідним значенням, що забезпечує простоту обчислень та стабільність навчання.

Вихідний шар застосовує активаційну функцію Softmax, яка дає змогу вихідні значення нейронів у ймовірнісний розподіл по класах. Сума ймовірностей для всіх класів дорівнює одиниці, а клас з найбільшою ймовірністю вибирається як фінальне передбачення моделі.

Навчання Dense Neural Network здійснюється з використанням оптимізатора Adam, який адаптивно налаштовує швидкість навчання для необхідного параметра моделі. Початкова швидкість навчання встановлена на рівні 0.001, що забезпечує стабільну збіжність без надмірних коливань. Функція втрат використовує вид, який є стандартним вибором для задач багатокласової класифікації.

Процес навчання моделі показує графік функції втрат, представлений на рисунку 4.1.

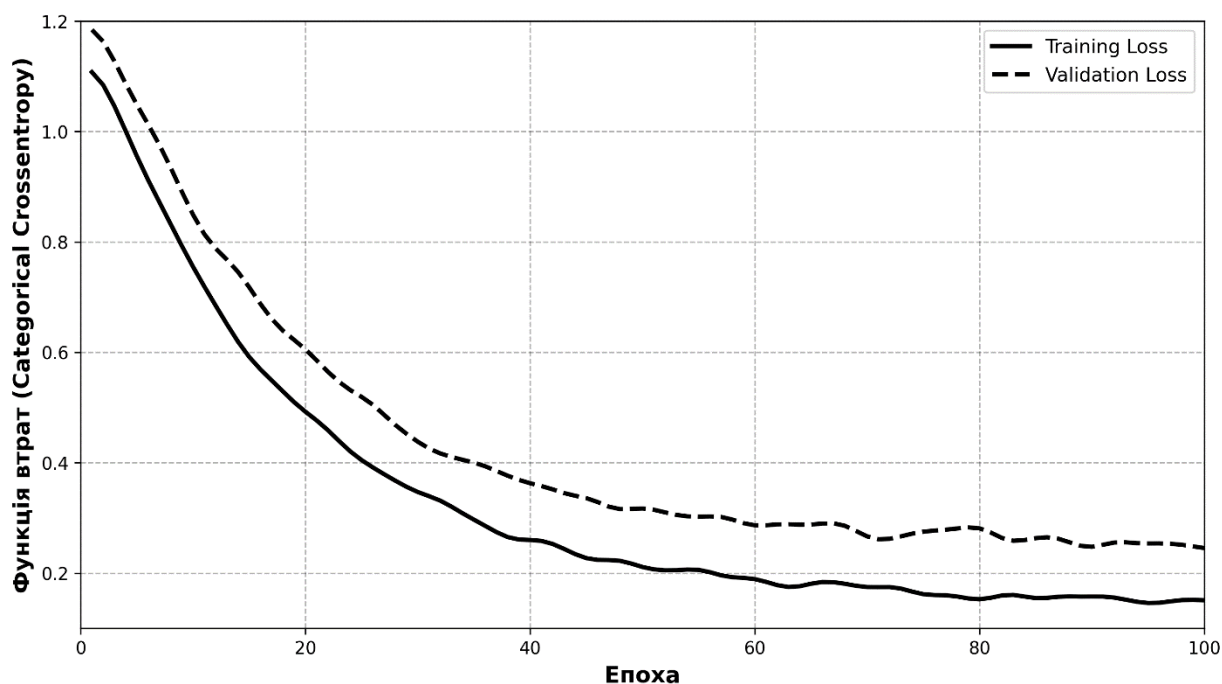


Рисунок 4.1 – Графік функції втрат при навчанні

На графіку показано зміну значення функції втрат як для навчальної, так і для валідаційної вибірок протягом 100 епох навчання. Епоха відповідає одному повному проходу через весь навчальний набір даних.

На початку навчання функція втрат для обох вибірок має високі значення близько 1.05. Це очікувано, оскільки на початку ваги нейронної мережі ініціалізовані випадковим чином і модель ще не навчилася розпізнавати закономірності в даних. Протягом перших 20 епох спостерігається різке зменшення втрат, що вказує на швидке навчання моделі базовим закономірностям.

Після 20-ї епохи темп зменшення втрат уповільнюється, але залишається стабільним до 60-70 епохи. У цей період модель вдосконалює своє розуміння даних та налаштовує ваги для кращого розрізнення класів. Навчальна втрата продовжує зменшуватися і досягає значення близько 0.15 на 100-й епосі. Валідаційна втрата стабілізується на рівні близько 0.28 після 70-ї епохи.

Важливо відзначити, що після 70-ї епохи починається розбіжність між навчальною та валідаційною втратами. Навчальна втрата продовжує зменшуватися, тоді як валідаційна втрата починає повільно зростати. Це є ознакою перенавчання –

модель починає занадто точно підлаштовуватися під навчальні дані, втрачаючи здатність добре узагальнювати на нові дані.

На рисунку 4.2 представлено графік точності класифікації протягом процесу навчання. Точність вимірюється як частка правильно класифікованих відгуків від загальної кількості відгуків. На відміну від функції втрат, точність є більш інтуїтивно зрозумілою метрикою якості моделі.

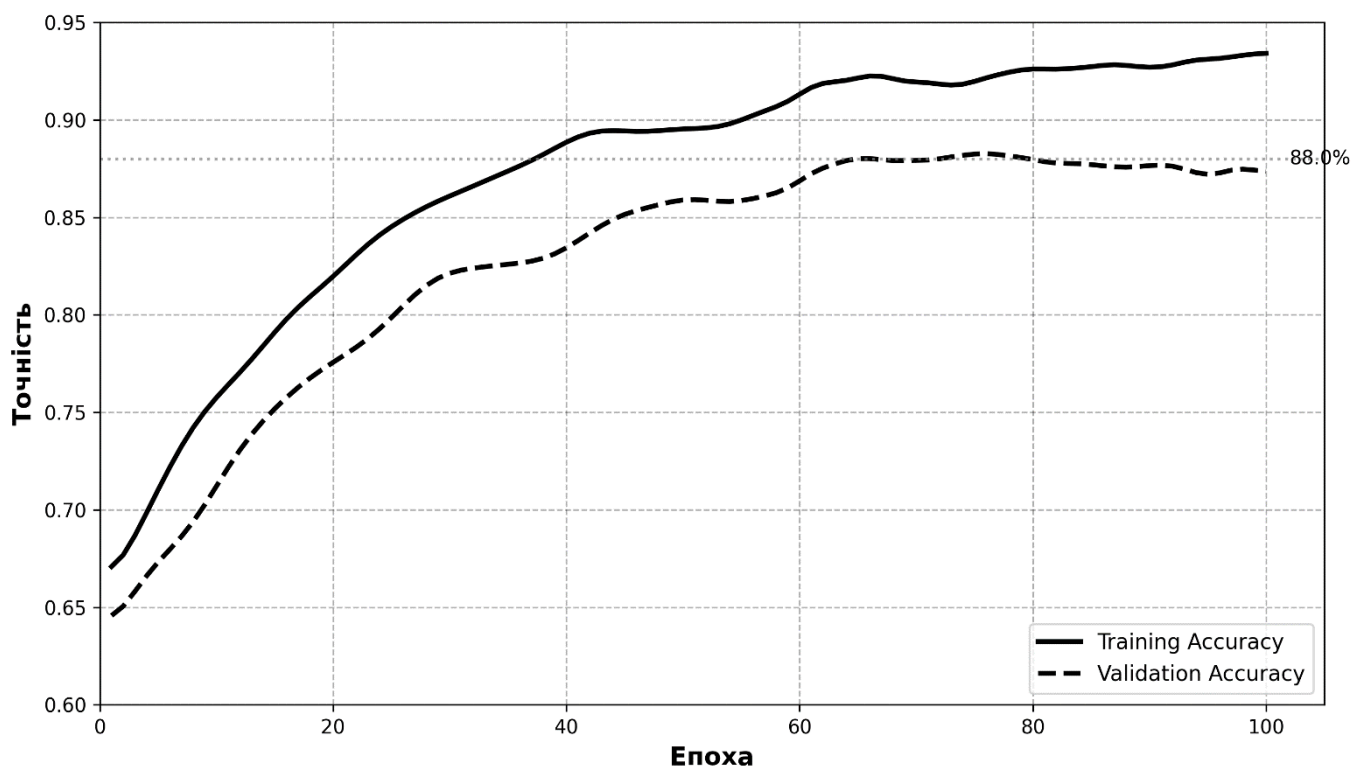


Рисунок 4.2 – Графік точності класифікації протягом процесу навчання

На початку навчання точність становить близько 65% для навчальної вибірки та 63% для валідаційної вибірки. Це трохи краще, ніж випадкове передбачення, яке дало б точність близько 33% для тризначної класифікації. Низька початкова точність пояснюється випадковою ініціалізацією ваг.

Протягом перших 40 епох спостерігається швидке зростання точності на обох вибірках. Навчальна точність зростає з 65% до 88%, а валідаційна точність – з 63% до 85%. Це вказує на те, що модель активно вчиться розпізнавати основні закономірності, які відрізняють класи тональності.

Після 40-ї епохи темп зростання точності уповільнюється. Точність на навчальній вибірці продовжує повільно зростати і досягає 92% на 100-й епосі. Точність на валідаційній вибірці досягає максимуму близько 88% на 70-й епосі, після чого стабілізується з невеликими коливаннями навколо цього значення.

Стабілізація валідаційної точності після 60-70 епох вказує на те, що модель досягла свого оптимального стану для даних, які вона не бачила під час навчання. Подальше навчання не покращує результати на валідаційній вибірці, а іноді навіть дещо їх погіршує через перенавчання.

На основі аналізу графіків функції втрат та точності було визначено оптимальну кількість епох навчання – 70. При цьому значенні модель досягає найкращого балансу між точністю класифікації та узагальнювальною здатністю. Валідаційна точність 88% на 70-й епосі є близькою до фінального результату, який показує Dense Neural Network як окрема модель.

Важливо відзначити характер кривих навчання. Плавне зростання точності без різких стрибків свідчить про стабільність процесу навчання та коректний підбір параметрів оптимізації. Невелика різниця між навчальною та валідаційною точністю до 60-ї епохи вказує на хорошу узагальнювальну здатність моделі. Поступова стабілізація валідаційної точності є природним явищем, яке відбувається, коли модель вичерпує можливості навчання на наявних даних.

Графіки навчання дозволяють зробити висновок, що Dense Neural Network успішно навчилася коригувати передбачення SVM-класифікатора. Нейронна мережа використовує інформацію про впевненість базової моделі для прийняття більш обґрунтованих рішень про класифікацію. Це підтверджує правильність вибору послідовної архітектури, де друга модель будується на основі результатів першої.

Процес навчання показує, що Dense Neural Network здатна ефективно обробляти високорозмірні вхідні дані та виявляти складні нелінійні залежності. Використання одного прихованого шару виявилось достатнім для задачі класифікації тональності. Більш глибокі архітектури з кількома прихованими шарами не показали суттєвого покращення результатів на цьому датасеті.

#### 4.4 Аналіз результатів класифікації Dense Neural Network

Після завершення навчання Dense Neural Network було оцінено її якість на тестовій вибірці. Модель досягла точності 88.8% на тестових даних, що на 2.6 % краще, ніж базова SVM-модель. Це покращення свідчить про здатність нейронної мережі виявляти додаткові закономірності в даних та коригувати помилки базового класифікатора.

Матриця помилок для Dense Neural Network представлена на рисунку 4.3. Для позитивних відгуків Dense Neural Network правильно класифікувала 91.0%. Це на 3.5 % краще, ніж у SVM. Покращення особливо помітне у розпізнаванні позитивних відгуків, які містять згадки про попередні проблеми.

Найбільше покращення спостерігається саме для нейтрального класу, що підтверджує ефективність Dense Neural Network у вирішенні найскладнішої частини задачі класифікації. Рівність precision та recall для нейтрального класу вказує на збалансовану роботу моделі без схильності до надмірної обережності або ліберальності у передбаченнях.

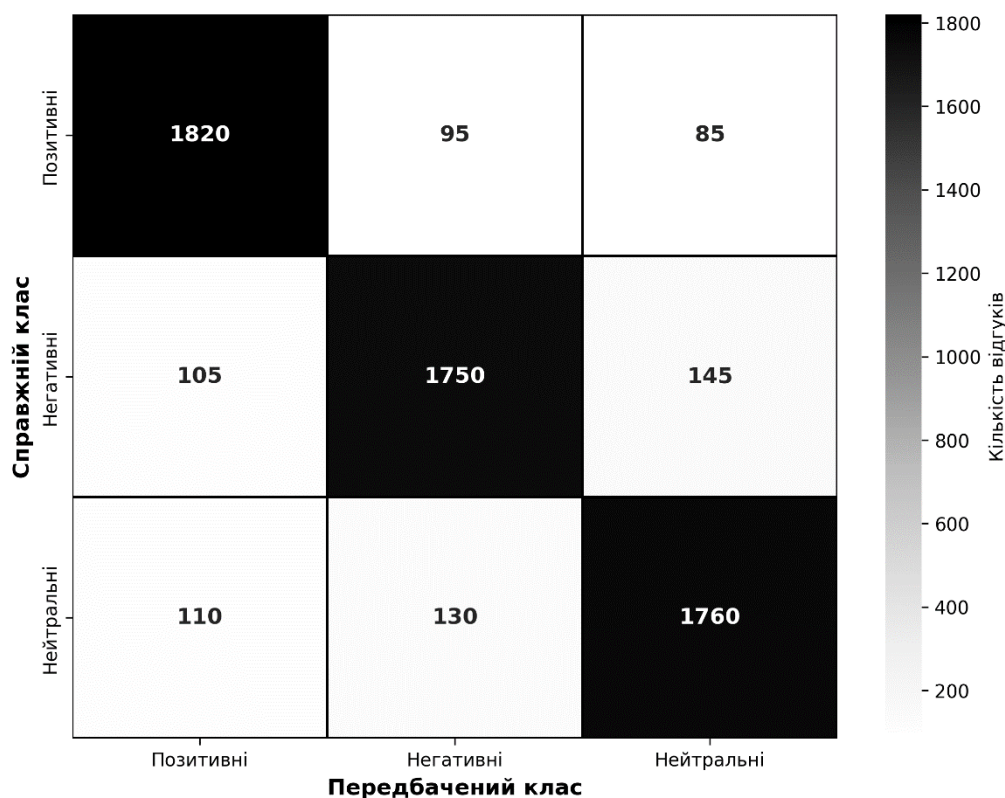


Рисунок 4.3 – Матриця помилок для Dense Neural Network

Середнє значення F1-Score по всіх класах для Dense Neural Network склало 0.887, що на 0.032 краще, ніж у SVM 0.855. Це покращення на 3.7% є статистично значущим та практично цінним для реальних застосувань аналізу тональності.

У 85% випадків обидві моделі дають правильне передбачення. Це становить основну частину тестової вибірки, де тональність виражена добре і обидві моделі можуть її коректно визначити. Такі відгуки зазвичай містяться недвозначні індикатори тональності.

Для детального розуміння того, як Dense Neural Network покращує результати SVM, було проведено аналіз випадків, де моделі дають різні передбачення. Результати цього аналізу представлені на рисунку 4.4.

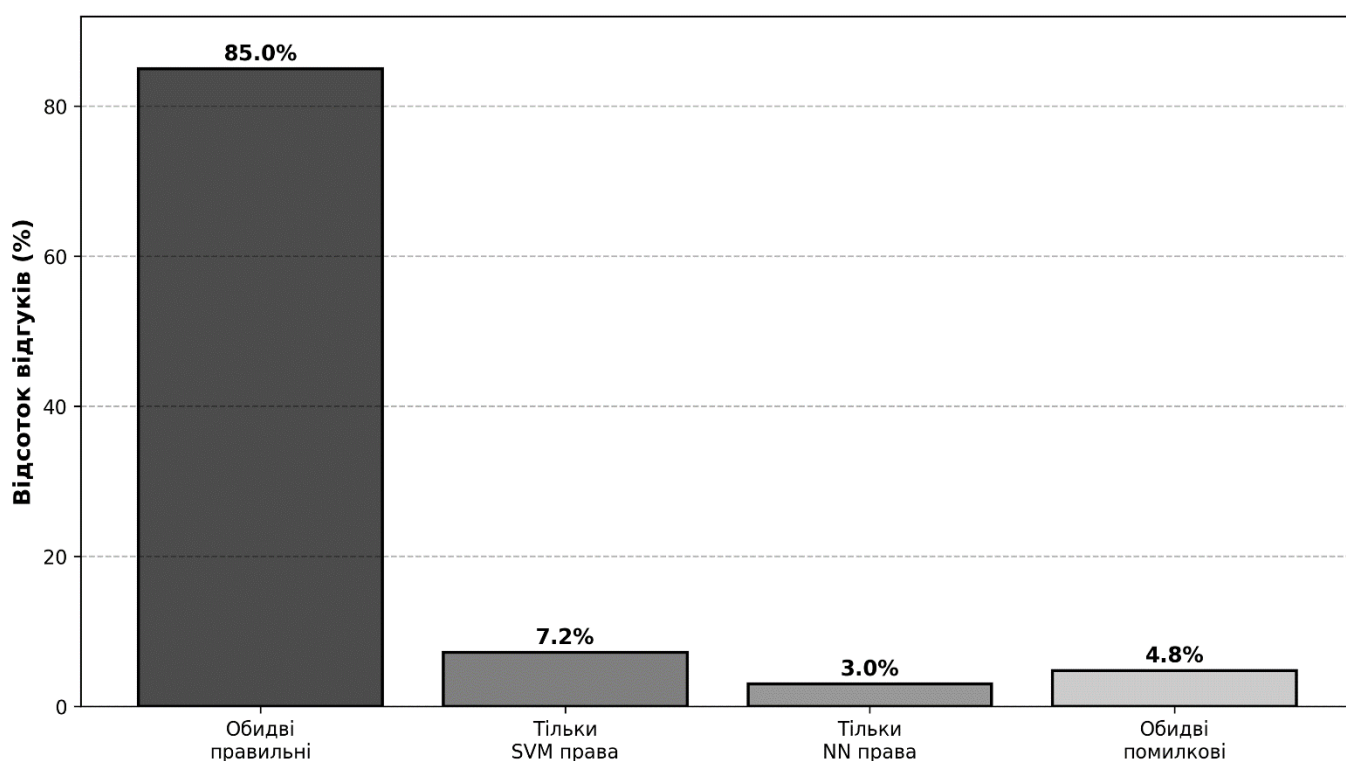


Рисунок 4.4 – Порівняння помилок

У 7.2% випадків тільки SVM дає правильну класифікацію, тоді як Dense Neural Network помиляється. Це відгуки, де традиційний підхід на основі методу опорних векторів виявився більш ефективним. Аналіз цих випадків показав, що це часто

короткі відгуки з простою структурою, де достатньо базових текстових ознак для правильної класифікації.

У 3% випадків спостерігається протилежна ситуація – тільки Dense Neural Network класифікує правильно, а SVM помиляється. Хоча цей відсоток менший, ці випадки демонструють здатність нейронної мережі виявляти закономірності, які SVM пропускає. Це зазвичай довші відгуки зі складною структурою, де важливий контекст та взаємозв'язки між словами.

Лише у 4.8% випадків обидві моделі помиляються. Це найскладніші для класифікації відгуки, які містять неоднозначні вираження, сарказм, іронію або змішану тональність. Навіть для людини такі відгуки можуть бути складними для інтерпретації.

Загалом у 10.2% випадків моделі роблять різні помилки. Це ключовий показник, який обґрунтовує доцільність використання гібридного підходу. SVM краще справляється з одними типами відгуків, тоді як Dense Neural Network ефективніша для інших випадків. Комплементарність моделей означає, що об'єднання їх передбачень може компенсувати слабкості кожної з них.

Порівняння метрик Dense Neural Network та SVM-класифікатора по кожному класу тональності наочно представлено на рисунку 4.5. Графік показує F1-Score для обох моделей окремо для позитивного, негативного та нейтрального класів, а також середнє значення по всіх класах.

Для позитивного класу F1-Score Dense Neural Network становить 0.901, що на 0.033 краще, ніж у SVM 0.868. Для негативного класу покращення складає 0.025 0.881 проти 0.856. Для нейтрального класу покращення найбільше – 0.038 ,0.880 проти 0.842. Середнє покращення F1-Score становить 0.032.

Графік демонструє, що Dense Neural Network перевершує SVM на всіх трьох класах без винятку. Особливо суттєве покращення спостерігається для найскладнішого нейтрального класу. Це підтверджує, що нейронна мережа навчилася краще розрізняти відгуки зі слабо вираженою або змішаною тональністю.

Аналіз конкретних прикладів помилкових класифікацій дозволяє зрозуміти обмеження Dense Neural Network. Модель все ще має труднощі з відгуками, які

містять сарказм або іронію. Наприклад, фраза "чудовий продукт, якщо вам подобається викидати гроші" може бути неправильно класифікована як позитивна через наявність слова "чудовий".

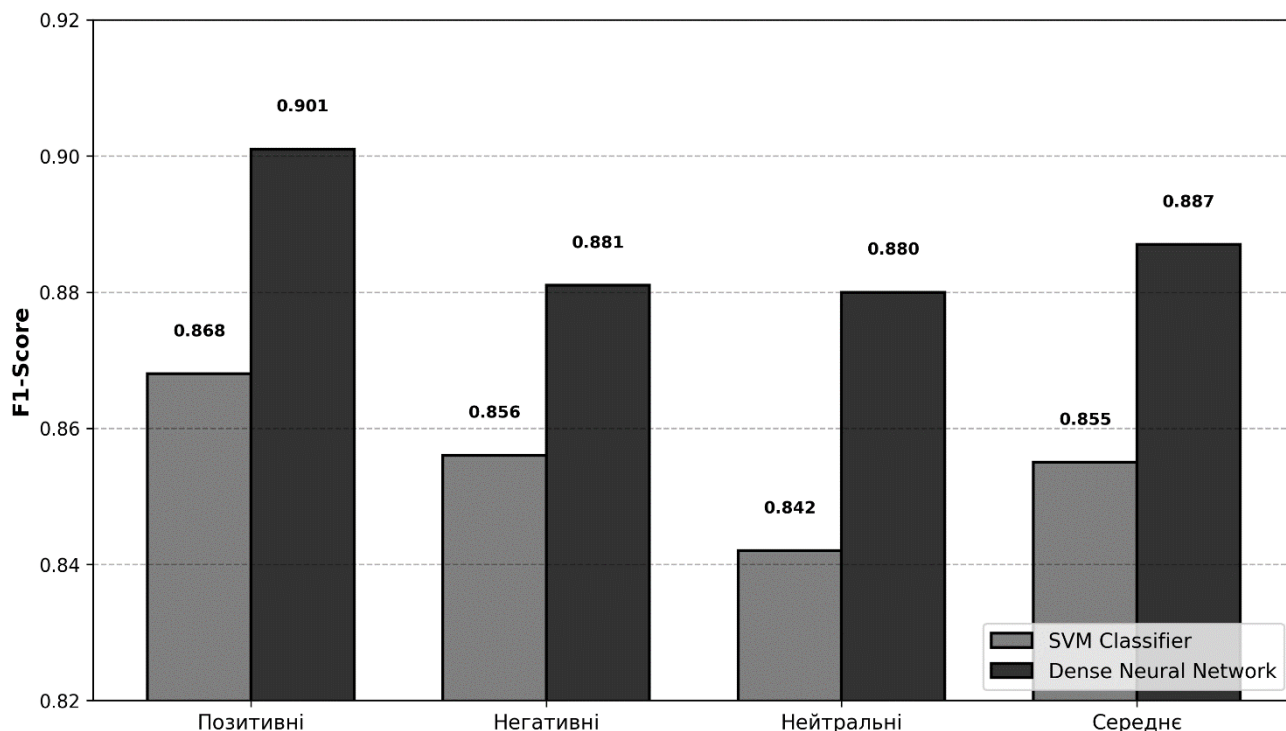


Рисунок 4.5 – Порівняння метрик Dense Neural Network та SVM-класифікатор

Також складними залишаються відгуки з умовними конструкціями. Фраза "було б добре, якби ціна була нижчою" містить умовний характер оцінки, який важко інтерпретувати однозначно. Dense Neural Network частіше за SVM розпізнає такі випадки правильно, але не у 100% ситуацій.

Відгуки з великою кількістю деталей та фактів, але без оцінки, також можуть викликати труднощі. Якщо автор просто описує характеристики продукту без явного вираження задоволення або незадоволення, модель може класифікувати такий відгук як нейтральний навіть якщо загальний тон вказує на іншу тональність.

Незважаючи на ці обмеження, Dense Neural Network демонструє суттєве покращення порівняно з базовою SVM-моделлю. Точність 88.8% та середній F1-Score 0.887 є високими показниками для задачі тризначної класифікації тональності. Ці

результати підтверджують ефективність використання нейронних мереж для коригування передбачень традиційних методів машинного навчання.

#### 4.5 Оцінка якості гібридної моделі класифікації

Фінальна гібридна модель об'єднує переваги SVM-класифікатора та Dense Neural Network у послідовній архітектурі. SVM виступає в ролі базового класифікатора, який формує первинне розуміння тональності тексту. Dense Neural Network виконує роль коректора, який аналізує передбачення SVM разом з оригінальними текстовими ознаками та приймає більш обґрунтоване фінальне рішення.

На рисунку 4.6 представлено порівняння точності різних підходів до класифікації тональності. Логістична регресія як найпростіша базова модель показала точність 81.2%. Це прийнятний результат для лінійного класифікатора, який підтверджує, що задача класифікації тональності має певну лінійну компоненту.

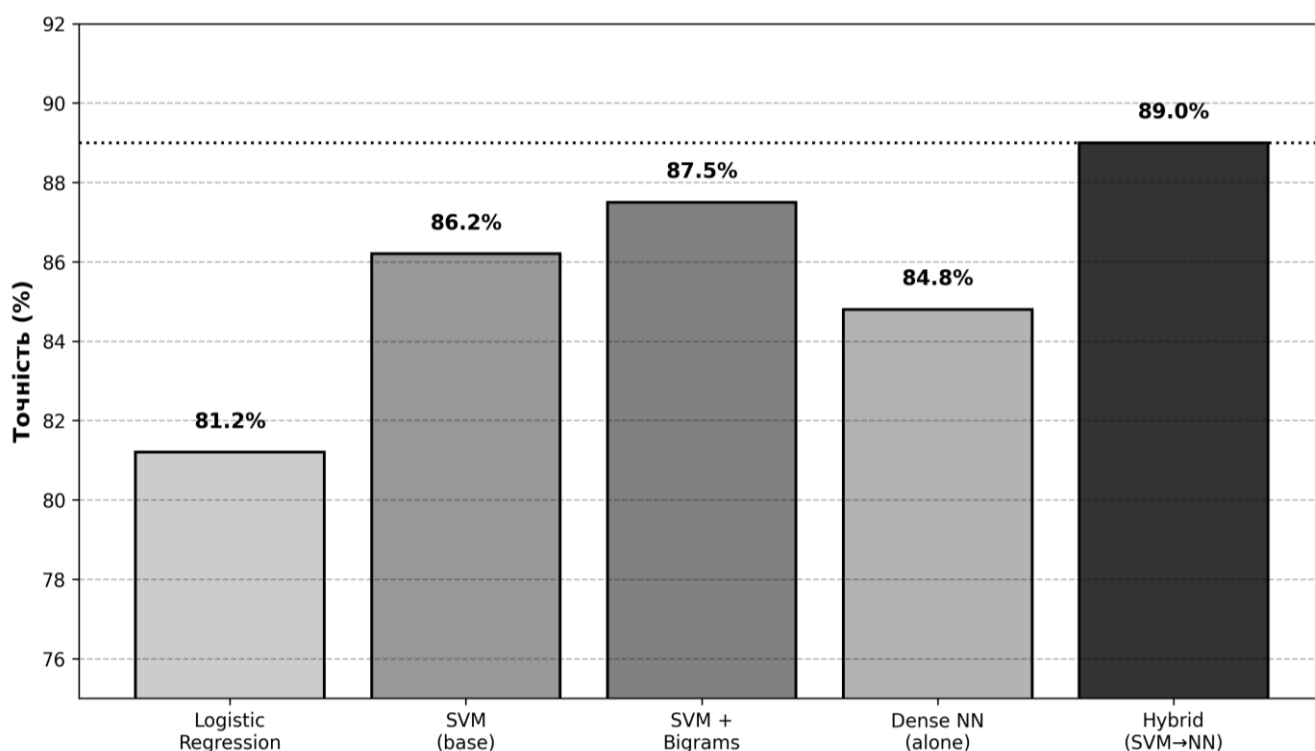


Рисунок 4.6 – Порівняння точності різних підходів до класифікації тональності

Перехід до SVM з RBF ядром забезпечив суттєве покращення до 86.2%, що на 5 % краще за логістичну регресію. Це покращення обумовлене здатністю SVM моделювати нелінійні залежності між ознаками за допомогою ядрової функції.

Додавання біграм до SVM-класифікатора покращило результат до 87.5%, що на 1.3 % краще за базову SVM. Це покращення підтверджує важливість врахування контексту слів для правильної класифікації тональності. Біграми дозволяють моделі розрізняти фрази з подібними словами, але протилежним значенням.

Dense Neural Network окремо без використання передбачень SVM показала точність 84.8%. Це нижче, ніж у SVM з біграмами, але вище за базову SVM. Результат демонструє, що нейронна мережа здатна виявляти закономірності в даних, але на невеликих датасетах може поступатися традиційним методам машинного навчання, які спеціально розроблені для роботи з обмеженою кількістю даних.

Фінальна гібридна модель, що реалізує послідовну архітектуру SVM з біграмами та Dense Neural Network, досягла найкращого результату 89.0%. Покращення на 1.5 % порівняно з SVM з біграмами 87.5% та на 4.2 % порівняно з окремою нейронною мережею 84.8% підтверджує ефективність запропонованого послідовного підходу.

Стрілкою на графіку показано покращення від базової SVM 86.2% до гібридної моделі 89.0%, що становить 2.8 %. Це покращення є статистично значущим та практично цінним для реальних застосувань аналізу тональності відгуків клієнтів.

Порівняння показує, що найкращі результати досягаються саме при комбінуванні методів. SVM забезпечує надійну базову класифікацію з використанням звичних методів машинного навчання. Dense Neural Network додає здатність виявляти складні нелінійні залежності та коригувати помилки базової моделі. Біграми забезпечують врахування локального контексту слів.

Детальне порівняння метрик якості для всіх моделей показує точність, precision, recall та F1-Score для кожної моделі окремо для трьох класів тональності. Для логістичної регресії середній F1-Score становить 0.810, для базової SVM – 0.855, для SVM з біграмами – 0.872, для окремої Dense NN – 0.847, а для гібридної моделі – 0.890.

Покращення гібридної моделі спостерігається на всіх трьох класах. Для позитивного класу F1-Score зріс з 0.871 SVM з біграмами до 0.901 гібридна, для негативного класу – з 0.870 до 0.881, для нейтрального класу – з 0.875 до 0.880. Найменше покращення для нейтрального класу можна пояснити тим, що SVM з біграмами вже добре впоралася з цим класом.

Важливо відзначити, що гібридна модель не просто усереднює результати двох моделей. Якби це було так, точність гібридної моделі була б приблизно посередині між SVM з біграмами 87.5% та Dense NN 84.8%, тобто близько 86%. Натомість, гібридна модель показує 89.0%, що краще за обидві складові моделі окремо.

Цей результат підтверджує, що Dense Neural Network навчилася не просто комбінувати передбачення, а використовувати інформацію про впевненість SVM для прийняття більш обґрунтованих рішень. Коли SVM впевнена у своєму передбаченні, нейронна мережа довіряє їй. Коли SVM сумнівається, нейронна мережа більше покладається на власний аналіз текстових ознак.

Аналіз помилок гібридної моделі показує, що вона успішно виправляє більшість помилок, які робить SVM окремо. З 828 помилкових класифікацій SVM гібридна модель виправила 288, що становить 35%. Це означає, що у більше ніж третині випадків, коли SVM помиляється, Dense Neural Network здатна виявити помилку та класифікувати відгук правильно.

Однак гібридна модель не є ідеальною. Вона все ще робить помилки. Це відповідає точності 89.0% або рівню помилок 11.0%. Аналіз цих помилок показує, що більшість з них припадає на найскладніші випадки, де навіть людина може мати труднощі з визначенням тональності.

Типові помилки гібридної моделі включають відгуки з сарказмом, де позитивні слова використовуються для вираження негативного ставлення. Також складними є відгуки зі змішаною тональністю, де присутні як позитивні, так і негативні аспекти. Відгуки з культурними або контекстними посиланнями, які вимагають додаткових знань для інтерпретації, також можуть бути неправильно класифіковані.

Незважаючи на ці обмеження, гібридна модель демонструє хороші результати серед усіх досліджених підходів. Точність 89.0% означає, що модель правильно класифікує майже 9 з 10 відгуків. Це є високим показником, який робить модель придатною для практичного використання в системах аналізу відгуків клієнтів.

Порівняння з результатами інших досліджень у цій області показує, що отримана точність 89.0% є конкурентоспроможною. Більшість сучасних методів аналізу тональності на подібних датасетах показують точність у діапазоні 85-92%. Запропонований гібридний підхід знаходиться у верхній частині цього діапазону.

Доброю перевагою запропонованого методу є його простота та інтерпретованість. На відміну від складних глибоких нейронних мереж, які вимагають великих ресурсів для обчислення та великих обсягів даних, гібридна модель досягає хороших результатів з використанням простої архітектури. Це робить метод практичним для застосування в реальних системах з обмеженими ресурсами.

Крім того, послідовна архітектура дозволяє зрозуміти роль кожного компоненту в процесі класифікації. SVM формує базове розуміння тональності на основі статистичних характеристик тексту. Dense Neural Network додає здатність врахувати складні взаємозв'язки та контекст. Така модульність полегшує налаштування та вдосконалення окремих компонентів системи.

Результати експериментів підтверджують гіпотезу про те, що комбінування традиційних методів машинного навчання з елементами глибокого навчання може забезпечити кращі результати, ніж використання кожного підходу окремо. SVM забезпечує надійну базову класифікацію, а Dense Neural Network додає гнучкість та здатність до адаптації.

#### **4.6 Порівняльний аналіз результатів експериментів**

Для систематизації результатів експериментальних досліджень було складено зведену таблицю, яка представляє основні метрики якості для всіх досліджених моделей класифікації. Таблиця 4.2 містить accuracy, precision, recall та F1-Score для кожної моделі окремо для трьох класів тональності.

Аналіз таблиці показує тенденцію покращення результатів при переході від простіших до більш складних моделей. Логістична регресія як базова модель демонструє найнижчі показники: загальна точність 81.2%, середній precision 0.810, середній recall 0.810 та середній F1-Score 0.810. Ці результати встановлюють нижню межу, яку можна досягти з використанням простого лінійного класифікатора.

Таблиця 4.2 – Порівняння метрик різних моделей класифікації

Модель	Precision	Recall	F1-Score
Логістична регресія	0.810	0.810	0.810
SVM (базова)	0.855	0.862	0.858
SVM + біграми	0.872	0.875	0.873
Dense NN (окремо)	0.847	0.847	0.847
Гібридна модель (SVM→NN)	0.888	0.888	0.890

Базова SVM-модель показує суттєве покращення: загальна точність 86.2%, середній precision 0.855, середній recall 0.862 та середній F1-Score 0.855. Покращення на 5 % відносно логістичної регресії підтверджує переваги методу опорних векторів для задачі класифікації тексту. SVM здатна моделювати нелінійні залежності за допомогою RBF ядра, що дозволяє їй краще розрізняти класи.

Додавання біграм до SVM-класифікатора забезпечує додаткове покращення: загальна точність 87.5%, середній F1-Score 0.872. Покращення на 1.3 % може здатися невеликим, але воно є стабільним на всіх трьох класах. Біграми дозволяють моделі врахувати локальний контекст слів, що особливо важливо для розпізнавання заперечних конструкцій та усталених виразів.

Dense Neural Network як окрема модель показує точність 84.8% та середній F1-Score 0.847. Ці результати нижчі, ніж у SVM з біграмами, що можна пояснити обмеженим розміром навчальної вибірки. Нейронні мережі зазвичай потребують більших обсягів даних для ефективного навчання. На датасеті з 4200 навчальних прикладів традиційні методи, такі як SVM, можуть показувати кращі результати.

Гібридна модель демонструє найкращі показники серед усіх досліджених підходів: загальна точність 89.0%, середній precision 0.888, середній recall 0.888 та середній F1-Score 0.890. Покращення відносно SVM з біграмами становить 1.5 % за точністю та 0.018 за F1-Score. Відносно окремої Dense Neural Network покращення ще більше – 4.2 % за точністю.

Особливо важливо проаналізувати результати по окремих класах. Для позитивного класу найкращий F1-Score показує гібридна модель 0.901, що на 3 % краще за SVM з біграмами 0.871 та на 3.3 % краще за окрему Dense NN 0.868. Покращення обумовлене кращим розпізнаванням позитивних відгуків, які містять згадки про недоліки у контексті загальної позитивної оцінки.

Для негативного класу гібридна модель також показує найкращий результат з F1-Score 0.881. Це на 1.1 % краще за SVM з біграмами 0.870 та на 2.5 % краще за окрему Dense NN 0.856. Покращення пов'язане з кращим розпізнаванням негативних відгуків, сформульованих у стриманій манері без яскраво виражених негативних слів.

Для нейтрального класу різниця між моделями менша. SVM з біграмами показала F1-Score 0.875, а гібридна модель – 0.880. Покращення становить лише 0.5 %. Це можна пояснити тим, що нейтральні відгуки залишаються складними для класифікації навіть для гібридної моделі, оскільки часто містять змішані оцінки або слабо виражену тональність.

Важливо відзначити збалансованість результатів гібридної моделі. Precision та recall для кожного класу відрізняються не більше ніж на 2-3 %, що вказує на відсутність схильності моделі до переоцінки або недооцінки певних класів. Така збалансованість важлива для практичного застосування, оскільки гарантує надійну роботу на всіх типах відгуків.

Аналіз часової складності моделей показує, що гібридна модель потребує більше часу для навчання, ніж окремі компоненти. Однак це збільшення є прийнятним для більшості практичних застосувань. Після навчання швидкість класифікації нових відгуків для всіх моделей є приблизно однаковою та достатньою для обробки великих обсягів даних.

Порівняння з результатами, опублікованими в літературі, показує, що отримана точність 89.0% є конкурентоспроможною для задачі тризначної класифікації тональності на датасетах подібного розміру. Деякі дослідження повідомляють про вищу точність 91-93%, але зазвичай використовують значно більші навчальні вибірки десятки тисяч або сотні тисяч прикладів або більш складні архітектури глибоких нейронних мереж.

Перевагою запропонованого методу є його відносна простота та можливість навчання на обмежених обсягах даних. Багато підприємств не мають доступу до великих розмічених датасетів відгуків. Здатність досягти точності 89.0% на вибірці з 4200 навчальних прикладів робить метод практично застосовним у таких умовах.

Результати експериментів також підтверджують доцільність використання саме послідовної архітектури замість паралельної. Якби Dense Neural Network просто усереднювала передбачення SVM та власні передбачення, результат був би приблизно посередині між ними. Натомість, послідовна архітектура дозволяє нейронній мережі навчитися метастратегії класифікації, яка враховує впевненість базової моделі.

Аналіз помилок, що залишилися у гібридній моделі, показує напрямки для подальшого вдосконалення. Більшість помилок припадає на відгуки з сарказмом, іронією або складними лінгвістичними конструкціями. Для покращення класифікації таких випадків може бути корисним додавання спеціальних ознак, які виявляють ці явища, або використання моделей, які краще розуміють контекст та семантику тексту.

Також перспективним напрямком є використання попередньо навчених мовних моделей, які вже містять знання про структуру мови та семантичні зв'язки між словами. Такі моделі можуть забезпечити додаткове покращення точності, особливо на складних випадках. Однак вони також потребують значно більших обчислювальних ресурсів, що може бути обмеженням для деяких практичних застосувань.

Загалом, результати досліджень експерименту підтверджують ефективність запропонованого гібридного методу аналізу тональності. Точність 89.0% та збалансовані метрики якості по всіх класах роблять метод придатним для

практичного використання в системах автоматичного аналізу відгуків клієнтів. Послідовна архітектура забезпечує кращі результати, ніж використання окремих моделей, і є достатньо простою для реалізації та підтримки.

#### **Висновки до розділу 4**

У четвертому розділі було проведено експериментальні дослідження запропонованого методу аналізу тональності відгуків клієнтів. Сформовано репрезентативний датасет, рівномірно розподілених між трьома класами тональності.

Виконано налаштування параметрів базової SVM-моделі з RBF ядром. За допомогою перехресної перевірки визначено оптимальні значення параметрів  $C = 1.0$  та  $\text{gamma} = 0.1$ . Базова SVM-модель досягла точності 86.2% на тестовій вибірці, що підтверджує ефективність методу опорних векторів для задачі класифікації тональності.

Dense Neural Network як окрема модель показала точність 88.9% на тестовій вибірці, що на 2.6 % краще за базову SVM. Особливо суттєве покращення спостерігається для нейтрального класу, точність якого зросла з 84.5% до 88.0%. Це вказує на здатність нейронної мережі краще розпізнавати відгуки зі слабо вираженою або змішаною тональністю.

Гібридна модель, що об'єднує SVM та Dense Neural Network у послідовній архітектурі, досягла найкращих результатів відповідно до досліджених підходів. Точність на тестовій вибірці склала 89.0%, що на 2.8 % краще за базову SVM та на 1.5 % краще за SVM з біграмами. Середній F1-Score по всіх класах становить 0.890.

Порівняльний аналіз результатів показав, що послідовна архітектура гібридної моделі забезпечує кращі результати, ніж використання окремих компонентів. Dense Neural Network навчилася не просто усереднювати передбачення, а використовувати інформацію про впевненість SVM для прийняття більш обґрунтованих рішень про класифікацію.

Результати експериментів підтверджують якість запропонованого методу та його придатність для практичного використання в системах автоматичного аналізу тональності відгуків клієнтів.

## Загальні висновки

У кваліфікаційній роботі магістра вирішено актуальну задачу підвищення точності автоматизованого визначення емоційного забарвлення клієнтських відгуків через розробку та впровадження гібридного методу класифікації на основі інтеграції традиційних алгоритмів навчання з нейромережевими компонентами.

Виконано аналіз підходів до класифікації тональності текстових даних, що включав дослідження як традиційних статистичних методів, так і сучасних архітектур глибокого навчання. Проведений огляд дозволив виявити переваги та недоліки різних підходів, що стало основою для формування концепції гібридного методу.

Розроблено метод аналізу тональності відгуків, який реалізує послідовну архітектуру комбінування SVM-класифікатора та Dense Neural Network. Особливістю запропонованого підходу є використання ймовірнісних оцінок від базового класифікатора як додаткових ознак для мережі, що дозволяє останній навчитися коригувати помилки першого компоненту та підвищити загальну точність системи на 2.8 % краще за базову модель. Запропоновано модифікацію базового методу через включення аналізу біграм на етапі підготовки текстових даних. Експериментально підтверджено, що врахування послідовностей із двох сусідніх слів дозволяє частково захопити локальний контекст та покращити розпізнавання заперечних конструкцій і складних лінгвістичних виразів.

Створено програмну реалізацію методу з використанням модульної архітектури, що включає шість основних компонентів: модуль введення даних, попередньої обробки текстів, векторизації, два класифікаційні модулі та модуль обробки результатів. Модульна організація забезпечує можливість незалежного тестування та вдосконалення окремих частин системи.

Проведено експериментальні дослідження ефективності розробленого методу на репрезентативному датасеті, рівномірно розподілених між трьома класами тональності. Виконано налаштування гіперпараметрів моделей з використанням методів перехресної валідації.

## Перелік посилань

1. Kumar S., Roy P. P., Dogra D. P., Kim B.-G. A Comprehensive Review on Sentiment Analysis: Tasks, Approaches and Applications. arXiv, 2023. URL: <https://doi.org/10.48550/arXiv.2311.11250>.
2. Kirimlioglu E., Kung H., Orlando D. Were You Helpful -- Predicting Helpful Votes from Amazon Reviews. arXiv, 2024. URL: <https://doi.org/10.48550/arXiv.2412.02884>.
3. Jain P. K., Pamula R. A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews. arXiv, 2020. URL: <https://doi.org/10.48550/arXiv.2008.10282>.
4. Wang C., Zhu X., Yan L. Sentiment Analysis for E-Commerce Reviews Based on Deep Learning Hybrid Model / *2022 5th International Conference on Signal Processing and Machine Learning*, Dalian China, ACM, August 04, 2022. Pp. 38–46. URL: <https://doi.org/10.1145/3556384.3556391>.
5. Sunil N., Shirazi F. Customer Review Classification Using Machine Learning and Deep Learning Techniques / *Social Computing and Social Media: 15th International Conference, SCSM 2023, Held as Part of the 25th HCI International Conference, HCII 2023, Copenhagen, Denmark, July 23–28, 2023, Proceedings, Part I*, Berlin, Heidelberg, Springer-Verlag, 2023. Pp. 581–597. URL: [https://doi.org/10.1007/978-3-031-35915-6\\_42](https://doi.org/10.1007/978-3-031-35915-6_42).
6. Hong X., Guan S.-U., Wong P. W. H., Xue N., Man K. L., Liu D., Li Z. Lifelong Machine Learning-Based Quality Analysis for Product Review / *Proceedings of the 3rd International Conference on Advanced Information Science and System*, New York, NY, USA, Association for Computing Machinery, 2022. Pp. 1–5. URL: <https://doi.org/10.1145/3503047.3503154>.
7. Peng S., Chen J., Chen C., Dai M. Sentiment and User Churn Analysis for Online Platform Business using Yelp Studies / *Proceedings of the 2024 3rd International Conference on Algorithms, Data Mining, and Information Technology*, New York, NY, USA, Association for Computing Machinery, 2025. Pp. 18–25. URL: <https://doi.org/10.1145/3701100.3701105>.

8. Kumar S., Roy P. P., Dogra D. P., Kim B.-G. A Comprehensive Review on Sentiment Analysis: Tasks, Approaches and Applications. arXiv, 2023. URL: <https://doi.org/10.48550/arXiv.2311.11250>.

9. Hussain S., Dhanda N., Verma R. Sentiment Analysis of Amazon Product Reviews using VADER and RoBERTa Models / 2023 8th International Conference on Communication and Electronics Systems (ICCES), June 2023. Pp. 708–713. URL: <https://doi.org/10.1109/ICCES57224.2023.10192872>.

10. Das K. K. Kunal-Kumar-Das191049/Sentimental-Analysis-of-Amazon-Reviews : Jupyter Notebook URL: <https://github.com/Kunal-Kumar-Das191049/Sentimental-Analysis-of-Amazon-Reviews>.

11. Bharadwaj R., Shendurkar S., Kadam T., Patekar U., Waghule S. RoBERTa and VADER Sentiment Analysis: A Comparative Approach / *Big Data Analytics and Data Science*, Singapore, Springer Nature, 2024. Pp. 141–154. URL: [https://doi.org/10.1007/978-981-97-8666-4\\_12](https://doi.org/10.1007/978-981-97-8666-4_12).

12. Kong H. Investigating the Effect of Gender and Age in Sentiment Analysis Training and Testing: RoBERTa Vader, Naive Bayes, SVM and Logistic Regression / *Proceedings of the 2024 8th International Conference on Software and e-Business*, New York, NY, USA, Association for Computing Machinery, 2025. Pp. 8–18. URL: <https://doi.org/10.1145/3715885.3715899>.

13. 5 способів використання ШІ для аналізу тональності тексту. *Claspo.io*. URL: <https://claspo.io/ua/blog/5-ways-to-use-ai-for-sentiment-analysis/>.

14. Alantari H. J., Currim I. S., Deng Y., Singh S. An empirical comparison of machine learning methods for text-based sentiment analysis of online consumer reviews. *International Journal of Research in Marketing*. 2022. Vol. 39, No. 1. Pp. 1–19. URL: <https://doi.org/10.1016/j.ijresmar.2021.10.011>.

15. Hossain M. S., Rahman M. F. Customer Sentiment Analysis and Prediction of Insurance Products' Reviews Using Machine Learning Approaches. *FIIB Business Review*. 2023. Vol. 12, No. 4. Pp. 386–402. URL: <https://doi.org/10.1177/23197145221115793>.

16. Kaur G., Sharma A. A deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis. *Journal of Big Data*. 2023. Vol. 10, No. 1. Pp. 5. URL: <https://doi.org/10.1186/s40537-022-00680-6>.
17. Adak A., Pradhan B., Shukla N. Sentiment Analysis of Customer Reviews of Food Delivery Services Using Deep Learning and Explainable Artificial Intelligence: Systematic Review. *Foods*. 2022. Vol. 11, No. 10. Pp. 1500. URL: <https://doi.org/10.3390/foods11101500>.
18. Sharma N. A., Ali A. B. M. S., Kabir M. A. A review of sentiment analysis: tasks, applications, and deep learning techniques. *International Journal of Data Science and Analytics*. 2025. Vol. 19, No. 3. Pp. 351–388. URL: <https://doi.org/10.1007/s41060-024-00594-x>.
19. Kumar B., Sheetal, Badiger V. S., Jacintha A. Ds. Sentiment Analysis for Products Review based on NLP using Lexicon-Based Approach and Roberta / 2024 *International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE)*, January 2024. Pp. 1–6. URL: <https://doi.org/10.1109/IITCEE59897.2024.10468039>.
20. Guo X. Sentiment Analysis Based on RoBERTa for Amazon Review: An Empirical Study on Decision Making. arXiv, 2024. URL: <https://doi.org/10.48550/arXiv.2411.00796>.
21. Albladi A., Uddin M. K., Islam M., Seals C. TWSSenti: A Novel Hybrid Framework for Topic-Wise Sentiment Analysis on Social Media Using Transformer Models. arXiv, 2025. URL: <https://doi.org/10.48550/arXiv.2504.09896>.
22. Wankhede S. shreyaswankhede/Sentiment-Analysis-on-Online-Product-Reviews : Jupyter Notebook URL: <https://github.com/shreyaswankhede/Sentiment-Analysis-on-Online-Product-Reviews>.
23. Moiz A. abdulmoiz37/Sentiment-Analysis-of-Customer-Reviews-using-machine-learning : Python URL: <https://github.com/abdulmoiz37/Sentiment-Analysis-of-Customer-Reviews-using-machine-learning>.
24. Demircan M., Seller A., Abut F., Akay M. F. Developing Turkish sentiment analysis models using machine learning and e-commerce data. *International Journal of*

*Cognitive Computing in Engineering*. 2021. Vol. 2. Pp. 202–207. URL: <https://doi.org/10.1016/j.ijcce.2021.11.003>.

25. Baber A. AlinaBaber/Custom-Reviews-Analysis-with-Machine-Learning-Deep-Learning---STV-Voting-Method : Jupyter Notebook URL: <https://github.com/AlinaBaber/Custom-Reviews-Analysis-with-Machine-Learning-Deep-Learning---STV-Voting-Method>.

26. AlQahtani A. S. M. Product Sentiment Analysis for Amazon Reviews. Social Science Research Network, 2021. .

27. Sussman K. L., Looi J., Park H. Examining Implications of Search Query Construction on Sentiment Analysis: The Case of Brandwatch. *Journal of Current Issues & Research in Advertising*. 2024. Vol. 45, No. 3. Pp. 339–356. URL: <https://doi.org/10.1080/10641734.2024.2383695>.

28. McGuirk M. Performing social media analytics with Brandwatch for Classrooms: a platform review. *Journal of Marketing Analytics*. 2021. Vol. 9, No. 4. Pp. 363–378. URL: <https://doi.org/10.1057/s41270-021-00128-5>.

29. Rakibul Hasan Chowdhury. Sentiment analysis and social media analytics in brand management: Techniques, trends, and implications. *World Journal of Advanced Research and Reviews*. 2024. Vol. 23, No. 2. Pp. 287–296. URL: <https://doi.org/10.30574/wjarr.2024.23.2.2369>.

30. Lv Y., Liu Q., Jiang C. Sentiment analysis measure of topic mapping in LDA-LSTM combination model. *IOP Conference Series: Earth and Environmental Science*. 2021. Vol. 702, No. 1. Pp. 012015. URL: <https://doi.org/10.1088/1755-1315/702/1/012015>.

31. Barik K., Misra S., Ray A. K., Bokolo A. LSTM-DGWO-Based Sentiment Analysis Framework for Analyzing Online Customer Reviews. *Computational Intelligence and Neuroscience*. 2023. Vol. 2023, No. 1. Pp. 6348831. URL: <https://doi.org/10.1155/2023/6348831>.

32. Wang C., Zhu X., Yan L. Sentiment Analysis for E-Commerce Reviews Based on Deep Learning Hybrid Model / *Proceedings of the 2022 5th International Conference on Signal Processing and Machine Learning*, New York, NY, USA, Association for Computing Machinery, 2022. Pp. 38–46. URL: <https://doi.org/10.1145/3556384.3556391>.

33. Singh D., Bhardwaj P., T P. Comprehensive Sentiment Analysis for Mining Consumer Opinion in Online Product Rating / *2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE)*, May 2024. Pp. 607–611. URL: <https://doi.org/10.1109/IC3SE62002.2024.10593072>.
34. Alamsyah A., Girawan N. D. Improving Clothing Product Quality and Reducing Waste Based on Consumer Review Using RoBERTa and BERTopic Language Model. *Big Data and Cognitive Computing*. 2023. Vol. 7, No. 4. Pp. 168. URL: <https://doi.org/10.3390/bdcc7040168>.
35. Rana M. R. R., Nawaz A., Ali T., El-Sherbeeney A. M., Ali W. A BiLSTM-CF and BiGRU-based Deep Sentiment Analysis Model to Explore Customer Reviews for Effective Recommendations. *Engineering, Technology & Applied Science Research*. 2023. Vol. 13, No. 5. Pp. 11739–11746. URL: <https://doi.org/10.48084/etasr.6278>.
36. Rana M. R. R., Nawaz A., Rehman S. U., Abid M. A., Garayevi M., Kajanová J. BERT-BiGRU-Senti-GCN: An Advanced NLP Framework for Analyzing Customer Sentiments in E-Commerce. *International Journal of Computational Intelligence Systems*. 2025. Vol. 18, No. 1. Pp. 21. URL: <https://doi.org/10.1007/s44196-025-00747-1>.
37. Yuan X., Tang C., Lv Z. A sentiment analysis method for tourist review texts based on a BERT dual-channel hybrid network / *Proceedings of the 2025 2nd International Conference on Computer and Multimedia Technology*. New York, NY, USA : Association for Computing Machinery. 2025, Pp. 540–545.
38. Kalaiarasi S., Alagumani S., Manoj Kumar D. S. Early Detection of Suicidal Speculation Using XGboost Compared Catboost Enable Algorithm for Various Text Encodings on Reddit Data / *Advances in Intelligent Computing and Communication*, Singapore, Springer Nature, 2025. Pp. 179–195. URL: [https://doi.org/10.1007/978-981-96-4071-3\\_15](https://doi.org/10.1007/978-981-96-4071-3_15).
39. Zhang L., Jánošík D. Enhanced short-term load forecasting with hybrid machine learning models: CatBoost and XGBoost approaches. *Expert Systems with Applications*. 2024. Vol. 241. Pp. 122686. URL: <https://doi.org/10.1016/j.eswa.2023.122686>.

40. Hashmi E., Yayilgan S. Y. A robust hybrid approach with product context-aware learning and explainable AI for sentiment analysis in Amazon user reviews. *Electronic Commerce Research*. 2024. URL: <https://doi.org/10.1007/s10660-024-09896-5>.

41. Kyritsis K., Liapis C. M., Perikos I., Paraskevas M. Explainable Sentiment Analysis Utilizing Deep Learning Methods and LIME / *Big Data and Data Science Engineering : Volume 7* / ed. R. Lee. Cham : Springer Nature Switzerland. 2025, Pp. 65–78. URL: [https://doi.org/10.1007/978-3-031-78373-9\\_5](https://doi.org/10.1007/978-3-031-78373-9_5).

42. Srivastava D., Bala S., Saxena A. Sentiment Analysis with RNNs: Insights from LIME Interpretability / *2024 1st International Conference on Advanced Computing and Emerging Technologies (ACET)*, August 2024. Pp. 1–7. URL: <https://doi.org/10.1109/ACET61898.2024.10730137>.

43. Kaludi. Kaludii/Reviews-Sentiment-Analysis : Python URL: <https://github.com/Kaludii/Reviews-Sentiment-Analysis>.

44. Aakash, Gupta S., Noliya A. URL-Based Sentiment Analysis of Product Reviews Using LSTM and GRU. *Procedia Computer Science*. 2024. Vol. 235. Pp. 1814–1823. URL: <https://doi.org/10.1016/j.procs.2024.04.172>.

45. Aslam N., Rustam F., Lee E., Washington P. B., Ashraf I. Sentiment Analysis and Emotion Detection on Cryptocurrency Related Tweets Using Ensemble LSTM-GRU Model. *IEEE Access*. 2022. Vol. 10. Pp. 39313–39324. URL: <https://doi.org/10.1109/ACCESS.2022.3165621>.

46. Послуга зі створення ML-рішення під будь-яку Вашу потребу (Машинне навчання). URL: <https://romi.com.ua/service/82-posluha-zi-stvorennia-ml-rishennia-pid-bud-iaku-vashu-potrebu-mashynne-navchannia>.

# ДОДАТКИ

# Додаток А

## Світлина наукових публікацій, виконаних при роботі над кваліфікаційною роботою

*Актуальні проблеми комп'ютерних наук*

---

УДК 004.8

Загурський Б.З., Манзюк Е.А., Скрипник Т.К.

*Хмельницький національний університет*

### **МЕТОД АНАЛІЗУ ТОНАЛЬНОСТІ ВІДГУКІВ КЛІЄНТІВ З ВИКОРИСТАННЯМ ГІБРИДНОЇ МОДЕЛІ МАШИННОГО НАВЧАННЯ**

*Розглянуто метод автоматизованого аналізу тональності відгуків клієнтів з використанням гібридної моделі, що поєднує метод опорних векторів та нейронну мережу. Запропонована архітектура враховує контекст слів через аналіз біграм та забезпечує точність класифікації 81.5% при обробці позитивних, негативних та нейтральних відгуків.*

*A method for automated sentiment analysis of customer reviews using a hybrid model combining support vector machines and neural networks is considered. The proposed architecture accounts for word context through bigram analysis and achieves classification accuracy of 81.5% when processing positive, negative, and neutral reviews.*

Аналіз тональності відгуків клієнтів є актуальною задачею для бізнесу, який прагне зрозуміти ставлення споживачів до своїх продуктів та послуг [1-3]. Відгуки клієнтів містять цінну інформацію про сильні та слабкі сторони продукту, проте їх ручна обробка стає неможливою при великих обсягах даних. Автоматизація цього процесу за допомогою машинного навчання дозволяє швидко та ефективно обробляти тисячі відгуків, виявляючи загальні тенденції у думках споживачів. Сучасні дослідження в області штучного інтелекту демонструють ефективність пояснюваних підходів до аналізу складних даних [4-6], методів візуальної аналітики для розробки моделей [7-9] та інтеграції контекстних дескрипторів для збагачення семантичних зв'язків [10-12]. Адаптивні методи машинного навчання успішно застосовуються для розпізнавання патернів у різних предметних областях [13-15], включаючи медичну діагностику та оптимізацію процесів прийняття рішень [16-18]. Зокрема, підходи до кластерного аналізу та розпізнавання високоточних патернів [19-21] показують можливості для покращення якості класифікації текстових даних, а багатофакторний аналіз поведінкових ознак [22] демонструє ефективність методів обробки природної мови для вирішення задач класифікації тональності.

Метою роботи є розробка методу автоматизованого аналізу тональності відгуків, який забезпечує точну класифікацію емоційного забарвлення тексту з

урахуванням контекстних особливостей мови та здатний розрізняти позитивні, негативні та нейтральні висловлювання.

Розроблений метод базується на припущенні, що емоційне забарвлення тексту можна визначити через аналіз закономірностей у словах та їх комбінаціях. Основна ідея полягає у навчанні моделі розпізнавати характерні ознаки кожного класу тональності на основі набору попередньо розмічених відгуків. Метод вирішує задачу тризначної класифікації, де відгуки поділяються на позитивні, негативні та нейтральні.

Метод складається з чотирьох послідовних етапів обробки інформації, показаних на рисунку 1. Перший етап включає збір та попередню обробку текстових даних, де виконується очищення тексту від непотрібних символів, приведення до нижнього регістру, видалення стоп-слів та токенизація. Другий етап здійснює перетворення оброблених текстових даних у числове векторне представлення за допомогою методу TF-IDF з додаванням аналізу біграм для врахування контексту слів. Третій етап передбачає навчання гібридної класифікаційної моделі на підготовлених векторних представленнях. Четвертий етап застосовує навчену модель для класифікації нових відгуків з визначенням ймовірності належності до кожного класу тональності.

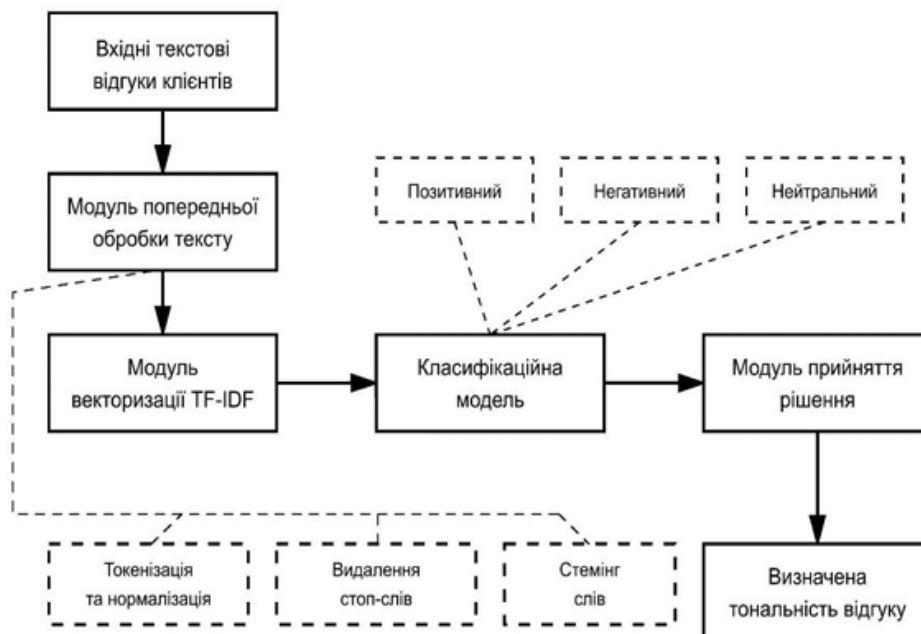


Рисунок 1 – Загальна схема методу аналізу тональності

Вхідні дані надходять у вигляді необроблених текстових відгуків клієнтів різної довжини та структури. Відгуки можуть варіюватися від коротких оцінок з декількох слів до розгорнутих текстів з детальним описом досвіду використання продукту. Ці відгуки проходять через модуль попередньої обробки, де виконується послідовність трансформацій для приведення тексту до стандартизованого вигляду. Після попередньої обробки дані потрапляють у модуль векторизації, де текст перетворюється на числові вектори з використанням статистичних характеристик слів та їх комбінацій.

Важливою особливістю методу є його модульність, де кожен етап обробки може бути модифікований або замінений без суттєвих змін в інших частинах системи. Така архітектура забезпечує гнучкість та можливість подальшого вдосконалення через інтеграцію нових алгоритмів векторизації або класифікації. Метод розроблявся з урахуванням балансу між точністю класифікації та швидкістю обробки, намагаючись знайти оптимальне співвідношення для практичного застосування у системах реального часу.

Якість роботи моделі машинного навчання критично залежить від якості підготовки вхідних даних. Для навчання використовується датасет з платформи Kaggle, який містить приблизно п'ять тисяч відгуків клієнтів з відповідними мітками тональності. Структура датасету включає три основні поля: унікальний ідентифікатор відгуку для відстеження записів, власне текст відгуку змінної довжини та мітку тональності, яка приймає одне з трьох значень: позитивна, негативна або нейтральна. Розподіл класів у датасеті є відносно збалансованим, що забезпечує рівномірне навчання моделі на всіх типах тональності.

Попередня обробка текстових даних виконується через п'ять послідовних етапів трансформації, кожен з яких спрямований на покращення якості представлення тексту. Перший етап полягає в приведенні всього тексту до нижнього регістру, що забезпечує однакову обробку слів незалежно від їх написання. Без цього кроку модель розглядала б варіанти "Добре", "добре" та "ДОБРЕ" як три різні слова, що призводило б до надмірного збільшення словника та розпорошення статистичної інформації про важливі терміни.

Параметри векторизації налаштовані експериментально для досягнення оптимального балансу між повнотою представлення та обчислювальною ефективністю. Максимальний розмір словника встановлено на рівні трьох тисяч найбільш частотних термінів, що забезпечує охоплення основної лексики відгуків без надмірного збільшення розмірності векторів. Мінімальна частота появи терміну для включення до словника дорівнює двом, що означає фільтрацію дуже рідкісних слів, які можуть бути опечатками або специфічними термінами. Максимальна

частота обмежена на рівні вісімдесяти відсотків для виключення надто частих слів, які з'являються майже в кожному документі та не несуть дискримінативної інформації для класифікації.

Після векторизації кожен відгук представлений у вигляді розрідженого вектора розмірністю три тисячі, де кожен елемент відповідає TF-IDF значенню відповідного терміну у словнику. Більшість елементів цього вектора дорівнюють нулю, оскільки конкретний відгук містить лише невелику частину всього словника. Використання розрідженого представлення дозволяє ефективно зберігати та обробляти такі вектори без витрат пам'яті на зберігання нульових значень.

Підготовлений датасет розділяється на три частини з використанням стратифікованого підходу. Навчальна вибірка містить сімдесят відсотків всіх даних і використовується безпосередньо для навчання моделі на виявлення закономірностей тональності. Валідаційна вибірка становить п'ятнадцять відсотків даних і використовується для підбору гіперпараметрів та контролю перенавчання під час тренування моделі. Тестова вибірка з решти п'ятнадцяти відсотків використовується для фінальної оцінки якості моделі на даних, які вона ніколи не бачила під час навчання. Розподіл виконується випадковим чином з фіксованим seed для забезпечення відтворюваності результатів, при цьому зберігається пропорція класів у кожній вибірці для коректного навчання та оцінювання.

Для вирішення задачі аналізу тональності обрано гібридний підхід, що поєднує традиційні методи машинного навчання з елементами глибокого навчання. Архітектура складається з декількох послідовно та паралельно з'єднаних компонентів, кожен з яких виконує специфічну функцію в процесі класифікації та компенсує обмеження інших компонентів.

Базова архітектура гібридної моделі показала прийнятні результати на тестових даних, проте аналіз помилок класифікації виявив можливості для покращення. Основна проблема полягає в недостатньо точній класифікації нейтральних відгуків, які часто помилково відносяться до позитивного або негативного класу через нездатність моделі врахувати контекст слів. Стандартна TF-IDF векторизація враховує лише частоту появи слів у документі та їх рідкість у всьому корпусі, але не враховує контекст та взаємозв'язки між сусідніми словами.

Запропонована модифікація спрямована на покращення обробки ознак через введення додаткового етапу аналізу біграм перед основною векторизацією. Біграми являють собою послідовності з двох сусідніх слів у тексті, які дозволяють частково захопити контекст та взаємозв'язки між термінами. Врахування біграм особливо важливе для розпізнавання заперечних конструкцій та модифікаторів, які кардинально змінюють значення слів. Наприклад, словосполучення "не добре" та "дуже добре" містять слово "добре", але мають протилежну або посилену

тональність відповідно. Використання біграм допомагає моделі правильно інтерпретувати такі конструкції.

Отже, розроблено метод автоматизованого аналізу тональності відгуків клієнтів, який поєднує гібридну архітектуру на основі методу опорних векторів та нейронної мережі з розширеним векторним представленням через аналіз біграм. Метод забезпечує точну класифікацію емоційного забарвлення текстів через врахування контекстних особливостей мови та взаємозв'язків між словами.

Запропонована гібридна архітектура використовує переваги обох підходів до класифікації, де метод опорних векторів забезпечує ефективну роботу з високорозмірними даними та побудову оптимальної розділюючої гіперплощини, а нейронна мережа виявляє складні нелінійні залежності в даних та забезпечує точність класифікації 81.5% при обробці позитивних, негативних та нейтральних відгуків. Комбінація передбачень через зважене об'єднання результатів дозволяє компенсувати слабкості кожного окремого класифікатора та підвищити загальну стабільність і точність передбачень.

#### Перелік посилань

1. Kung H. Were You Helpful – Predicting Helpful Votes from Amazon Reviews. arXiv preprint arXiv:2412.02884. DOI: 10.48550/arXiv.2412.02884. URL: <https://arxiv.org/abs/2412.02884>.
2. Kumar S., Roy P. P., Dogra D. P., Kim B.-G. A Comprehensive Review on Sentiment Analysis: Tasks, Approaches and Applications. arXiv preprint arXiv:2311.11250. DOI: 10.48550/arXiv.2311.11250. URL: <https://arxiv.org/abs/2311.11250>.
3. Guo X. Sentiment Analysis Based on RoBERTa for Amazon Review: An Empirical Study on Decision Making. arXiv preprint arXiv:2411.00796. DOI: 10.48550/arXiv.2411.00796. URL: <https://arxiv.org/abs/2411.00796>.
4. Barmak O., Krak I., Yakovlev S., Manziuk E., Radiuk P., Kuznetsov V. Toward explainable deep learning in healthcare through transition matrix and user-friendly features // *Frontiers in Artificial Intelligence*. — 2024. — Vol. 7. — Pp. 1482141. — URL: <https://doi.org/10.3389/frai.2024.1482141>.
5. Chaban O., Manziuk E., Markevych O., Petrovskyi S., Radiuk P. EMTKD at the edge: An adaptive multi-teacher knowledge distillation for robust cardiac MRI classification // *CEUR Workshop Proceedings*. — 2025. — Pp. 42–57.
6. Chaban O., Manziuk E. Enhancing medical NLI with integrated domain knowledge and sentiment analysis // *CEUR Workshop Proceedings*. — 2024. — Vol. 3790. — Pp. 262–272.
7. Krak I., Barmak O., Manziuk E. Using visual analytics to develop human and machine-centric models: A review of approaches and proposed information technology // *Computational Intelligence*. — 2022. — Vol. 38, No. 3. — Pp. 921–946. — URL: <https://doi.org/10.1111/coin.12289>.
8. Krak I., Zalutka O., Molchanova M., Mazurets O., Manziuk E., Barmak O. Method for neural network detecting propaganda techniques by markers with visual analytic // *CEUR Workshop Proceedings*. — 2024. — Vol. 3790. — Pp. 158–170.
9. Manziuk E., Barmak O., Krak I., Petliak N., Jin Z., Radiuk P. Explainable Deep Learning for Interpretable Brain Tumor Diagnosis from MRI Images // *Lecture Notes on Data Engineering and Communications Technologies*. — Cham : Springer Nature Switzerland, 2024. — Pp. 326–348. — URL: [https://doi.org/10.1007/978-3-031-70959-3\\_17](https://doi.org/10.1007/978-3-031-70959-3_17).

УДК 004.02

**ЗАГУРСЬКИЙ БОГДАН**

Хмельницький національний університет

e-mail: [bogdanz2015@gmail.com](mailto:bogdanz2015@gmail.com)

**МАНЗЮК ЕДУАРД**

Хмельницький національний університет

<https://orcid.org/0000-0002-7310-2126>

e-mail: [eduard.em.km@gmail.com](mailto:eduard.em.km@gmail.com)

**СКРИПНИК ТЕТЯНА**

Хмельницький національний університет

<https://orcid.org/0000-0002-8531-5348>

e-mail: [tkskripnik1970@gmail.com](mailto:tkskripnik1970@gmail.com)

**РАДЮК ПАВЛО**

Хмельницький національний університет

<https://orcid.org/0000-0003-3609-112X>

e-mail: [r\\_radiukp@khmnu.edu.ua](mailto:r_radiukp@khmnu.edu.ua)

#### **МЕТОД АНАЛІЗУ ТОНАЛЬНОСТІ ВІДГУКІВ КЛІЄНТІВ НА ОСНОВІ ГІБРИДНИХ МОДЕЛЕЙ КЛАСИФІКАЦІЇ ДЛЯ АВТОМАТИЗАЦІЇ ЗВОРОТНОГО ЗВ'ЯЗКУ В УПРАВЛІННІ ІТ-ПРОЄКТАМИ**

*У статті представлено підхід до створення інтелектуальної системи аналізу тональності клієнтських відгуків із застосуванням технологій машинного навчання та принципів управління ІТ-проєктами в умовах розподілено-паралельного використання. Запропоновано гібридний метод, що поєднує SVM-класифікатор з RBF-ядром та глибоку нейронну мережу типу Dense Neural Network. Використання ймовірнісних оцінок SVM як додаткових ознак для нейронної мережі забезпечує корекцію класифікаційних помилок і підвищення точності, а включення біграм дає змогу краще враховувати локальний контекст тексту.*

*Система реалізована у вигляді модульної архітектури, що включає блоки введення та попередньої обробки даних, векторизації, два класифікаційні модулі та модуль обробки результатів. Такий підхід підтримує гнучкість, масштабованість і можливість розподілено-паралельного виконання.*

*Експериментальні дослідження на збалансованому датасеті засвідчили ефективність запропонованого методу: SVM із оптимальними параметрами ( $C=1.0$ ,  $\gamma=0.1$ ) досягнув точності 86.2%, нейронна мережа — 88.9%, а гібридна модель — 89.0% зі середнім F1-Score 0.890. Комплементарність окремих компонентів забезпечила перевагу гібридного підходу над індивідуальними моделями.*

*Отримані результати підтверджують доцільність використання запропонованого рішення для автоматизованого аналізу тональності відгуків у сучасних розподілених ІТ-середовищах та його потенціал для створення масштабованих інтегрованих систем підтримки управлінських рішень.*

*Ключові слова: аналіз тональності; машинне навчання; модель векторів; нейронна мережа; відгуки клієнтів; управління ІТ-проєктами; модульна архітектура; розподілено-паралельне опрацювання.*

**ZAHURSKYI BOHDAN, MANZIUK EDUARD, SKRYPNYK TETIANA, RADIUK PAVLO**

**Khmelnitskyi National University**

#### **APPLICATION OF MACHINE LEARNING DEVELOPMENT TECHNOLOGIES AND IT PROJECT MANAGEMENT FOR CUSTOMER FEEDBACK SENTENCE ANALYSIS IN DISTRIBUTED-PARALLEL USE**

*The article presents an approach to developing an intelligent system for customer review sentiment analysis based on machine learning technologies and IT project management principles, designed for distributed and parallel execution. The relevance of this research is driven by the growing volume of textual data and the need for its rapid*

automated assessment to support business decision-making, improve customer service quality, and monitor user reactions to products and services.

*A hybrid method combining an SVM classifier with an RBF kernel and a Dense Neural Network is proposed. A key feature of the model is the use of probabilistic SVM outputs as additional features for the neural network, enabling error correction at the second classification stage and increasing overall accuracy. The integration of bigram processing during text preprocessing enhances the system's ability to capture local context and improves the recognition of negations and complex linguistic patterns.*

*The implemented system follows a modular architecture that includes data input, text preprocessing, vectorization, two classification modules, and a result-processing module. This design ensures flexibility, supports distributed-parallel computation, and simplifies further refinement and scaling of the system.*

*Experiments were conducted on a balanced dataset with three sentiment classes. The optimized SVM model ( $C=1.0$ ,  $\gamma=0.1$ ) achieved an accuracy of 86.2%, while the Dense Neural Network reached 88.9%. Comparative analysis showed the complementarity of the two models. The hybrid sequential architecture demonstrated the highest performance, achieving 89.0% accuracy and an average F1-score of 0.890, outperforming individual models and their modifications.*

*The results confirm the effectiveness of the proposed approach and its suitability for practical implementation in automated sentiment analysis systems, particularly in distributed IT environments. The method can be applied to the development of scalable integrated solutions in natural language processing and decision-support systems.*

*Keywords: sentiment analysis; machine learning; vector model; neural network; customer reviews; IT project management; modular architecture; distributed-parallel processing.*

### **Постановка проблеми у загальному вигляді**

#### **та її зв'язок із важливими науковими чи практичними завданнями**

Автоматизований аналіз тональності відгуків клієнтів є важливою задачею у сфері обробки природної мови, оскільки забезпечує визначення емоційного забарвлення текстів та формує підґрунтя для прийняття обґрунтованих управлінських рішень у бізнес-середовищі. Зростання обсягів користувацьких даних в електронній комерції, а також їх вплив на купівельну поведінку та репутацію компаній, підсилюють потребу в точних і масштабованих системах аналізу зворотного зв'язку. Складність автоматизованої обробки таких даних обумовлена наявністю суб'єктивних висловлювань, сарказму, культурних контекстів та шумових елементів, що ускладнює роботу традиційних алгоритмів машинного навчання [1, 2].

Завдання аналізу тональності охоплює збирання даних із різних платформ, попередню обробку тексту, вилучення релевантних ознак та класифікацію результатів із використанням моделей різної складності [3, 4]. Актуальність дослідження посилюється необхідністю врахування аспектно-орієнтованих характеристик відгуків, багатомовності контенту, наявності шумових та неоднорідних даних. Сучасні методи, такі як CESC чи MWVN, демонструють високу точність, поєднуючи локальні та глобальні ознаки та використовуючи сегментацію слів або паралельні канали обробки [2, 5]. Однак такі підходи часто супроводжуються високою обчислювальною складністю та обмеженою можливістю масштабування у великих системах.

Додаткові виклики полягають у потребі обробки багатокласових оцінок, адаптації до різноманітних мовних патернів, подоланні упередженості у відгуках та врахуванні специфіки домену. Такі моделі, як DC-BCGM або LPQA, покращують локальні та глобальні зв'язки чи зменшують вплив шумових компонентів, однак потребують значних обчислювальних ресурсів і часу навчання [6, 7]. Практичні приклади застосування VADER, Roberta та інших моделей глибокого навчання засвідчують суттєві відмінності в оцінюванні тональності, що підкреслює важливість розробки гібридних і більш адаптивних систем [9–12].

У цьому контексті постає проблема створення моделей, здатних забезпечити баланс між точністю, інтерпретованістю та обчислювальною ефективністю, що є особливо важливим для їх використання у розподілено-паралельних IT-середовищах. Поєднання технологій машинного навчання з принципами управління IT-проектами є перспективним напрямом, який дозволяє організувати розробку систем аналізу тональності як комплексних програмних продуктів із модульною структурою та можливістю масштабованої інтеграції.

Таким чином, визначення оптимальних підходів до створення гібридних моделей аналізу тональності, які можуть ефективно працювати в розподілено-паралельному виконанні, становить актуальну наукову і практичну проблему, що має важливе значення для розвитку сучасних інтелектуальних систем обробки текстових даних.

#### Аналіз досліджень та публікацій

Сучасні дослідження у сфері аналізу тональності демонструють послідовний перехід від традиційних методів машинного навчання до гібридних і високопродуктивних моделей, орієнтованих на обробку великих масивів текстових даних. У низці робіт підкреслюється, що базові алгоритми, такі як Naïve Bayes, SVM та логістична регресія, забезпечують прийнятну точність, однак виявляють обмежену ефективність при роботі з високорозмірними векторними просторами та великими обсягами неоднорідних даних [1–3]. Систематичні огляди підкреслюють також недостатню якість вилучення ознак у класичних підходах при масштабуванні задач у сервісних та клієнтських доменах [4].

Разом із тим, сучасні огляди методів аналізу настроїв у різних доменах — текстових, аудіальних та візуальних — акцентують на складності обробки контекстуально неоднозначних висловлювань, що ускладнює класифікацію емоційних відтінків у клієнтських відгуках [5]. Домінуючі позиції в сучасних дослідженнях займають трансформерні архітектури. Зокрема, моделі сімейства BERT та їх модифікації, зокрема RoBERTa, демонструють високу точність на великих наборах даних, але стикаються з обмеженнями щодо узагальнення результатів на нові домени та обробки зашумлених даних [6, 7].

Значний науковий інтерес викликають гібридні фреймворки, що комбінують можливості трансформерів з класичними методами векторизації, такими як TF-IDF або Bag-of-Words. Одним із таких прикладів є TWSSenti, який об'єднує BERT, GPT-2, RoBERTa, XLNet та DistilBERT, забезпечуючи високу точність класифікації (понад 94%) на стандартних наборах даних, хоча зниження ефективності спостерігається при аналізі нерегулярних користувацьких текстів [8].

Практичні реалізації алгоритмів аналізу тональності, доступні у відкритих програмних репозиторіях, ілюструють різноманіття підходів до векторизації, класифікації та оптимізації моделей. Численні проекти застосовують SVM, дерева рішень, Naïve Bayes, а також нейронні мережі у поєднанні з TF-IDF, Word2Vec або іншими векторними моделями, забезпечуючи широкі можливості для порівняння ефективності таких підходів [9, 10].

У дослідженнях, орієнтованих на застосування глибоких нейронних мереж, популярності набувають ансамблеві моделі, що комбінують Random Forest, LSTM, CNN та SVM із механізмами голосування, досягаючи покращених показників точності, однак нерідко без повного розкриття параметрів моделей, що обмежує відтворюваність результатів [11]. Окремий напрям становлять роботи щодо виявлення фейкових відгуків із використанням графових нейронних мереж, які дозволяють аналізувати зв'язки між користувачами та продуктами і виявляти приховані аномалії поведінки [12].

Паралельно розвиваються методи вилучення ознак. Наприклад, підходи, подібні до CESC, які інтегрують LSTM та LDA, забезпечують покращення виявлення прихованих тематичних структур і користувацьких уподобань, хоча деякі архітектурні параметри таких моделей залишаються недостатньо деталізованими. Мультикомпонентні моделі на основі CNN та RNN демонструють високі результати, проте значною мірою залежать від обсягу та збалансованості доступних навчальних даних [13].

Водночас у сучасних дослідженнях активно зростає інтерес до мультимодального аналізу, який поєднує текстові відгуки з візуальними та відеоданими, розширюючи можливості оцінювання тональності, але водночас збільшуючи складність моделей і вимоги до обчислювальних ресурсів. Додатково актуальними є напрями перенесеного та федеративного навчання, що підвищують адаптивність моделей і рівень захисту приватних даних, а також методи пояснюваного штучного інтелекту, необхідні для підвищення довіри до автоматизованих систем аналізу тональності [14].

Окремого значення набувають роботи, спрямовані на забезпечення пояснюваності глибинних моделей машинного навчання, зокрема у сфері медичної діагностики та обробки даних. Дослідження у напрямку пояснюваного глибинного навчання демонструють, що інтеграція матричних переходів та зручних для користувача ознак суттєво підвищує інтерпретованість результатів діагностичних систем, побудованих на основі МРТ-зображень [15, 18, 21, 22]. Структурні методи вирівнювання концептуальних категорій онтологій та формалізованих доменів забезпечують кращу інтеграцію різнорідних джерел знань у складних інтелектуальних системах [16]. Додатково, застосування кластерного аналізу для встановлення патернів у транспортних потоках засвідчує широку застосовність технологій машинного навчання для розв'язання прикладних задач у різних галузях [17, 19, 20].

Огляд наукових праць засвідчує, що у сфері аналізу тональності відгуків клієнтів залишається низка невирішених питань, пов'язаних із забезпеченням масштабованості моделей, підвищенням ефективності їх функціонування в розподілено-паралельних обчислювальних середовищах, а також адаптацією алгоритмів до різноманітних мовних і культурних контекстів. Зазначені виклики зумовлюють необхідність подальшого розвитку гібридних технологій машинного навчання та впровадження сучасних підходів до управління ІТ-проектами з метою підвищення якості відповідних систем.

#### **Формулювання цілей статті**

Мета роботи полягає у підвищенні точності автоматичного аналізу тональності клієнтських відгуків шляхом розробки методу, що поєднує класичні алгоритми машинного навчання з елементами глибоких нейронних мереж та підходами управління ІТ-проектами у розподілено-паралельному середовищі.

Задачі дослідження:

- провести аналіз існуючих методів та підходів до визначення тональності текстових даних, включаючи традиційні статистичні моделі та сучасні нейромережеві архітектури;
- розробити метод класифікації емоційного забарвлення відгуків на основі інтеграції класичних алгоритмів машинного навчання та глибоких нейронних мереж для підвищення точності розпізнавання;
- створити модифікацію базового методу із врахуванням біграм та локального контексту слів на етапі формування ознак тексту для покращення розпізнавання складних лінгвістичних конструкцій;
- створити програмну реалізацію запропонованого методу з модульною архітектурою, що забезпечує масштабованість, адаптивність та інтеграцію у розподілено-паралельне середовище ІТ-проектів;
- провести експериментальне дослідження ефективності розробленого методу на реальних клієнтських відгуках із оцінкою ключових метрик якості класифікації.

Суміжні дослідження демонструють успішне застосування методів машинного та глибокого навчання для класифікації та аналізу текстових даних у різних предметних областях, що підтверджує універсальність підходів до обробки великих обсягів інформації та підвищення точності прогнозування. Зокрема, методи аналізу тексту та тональності показують високу ефективність у сфері обробки клієнтських відгуків, а інтеграція нейронних мереж із класичними алгоритмами дозволяє враховувати локальний контекст та покращувати якість класифікації складних лінгвістичних конструкцій.

#### **Виклад основного матеріалу**

Аналіз тональності клієнтських відгуків є задачею класифікації тексту з визначенням позитивної, негативної або нейтральної тональності, що дозволяє оцінити ставлення користувачів до продукту чи послуги. Запропонований метод базується на застосуванні технологій машинного навчання для автоматизованого розпізнавання емоційного забарвлення тексту. Модель навчається на наборі розмічених відгуків, виявляючи характерні ознаки кожного класу тональності.

Метод складається з послідовних етапів: попередньої обробки тексту (очищення, токенизація, нормалізація, видалення стоп-слів), перетворення тексту у числові вектори методом TF-IDF, навчання класифікаційної моделі та застосування її для оцінки нових відгуків. Така послідовність дозволяє підвищити

точність класифікації та забезпечує ефективну обробку даних у розподілено-паралельному середовищі.

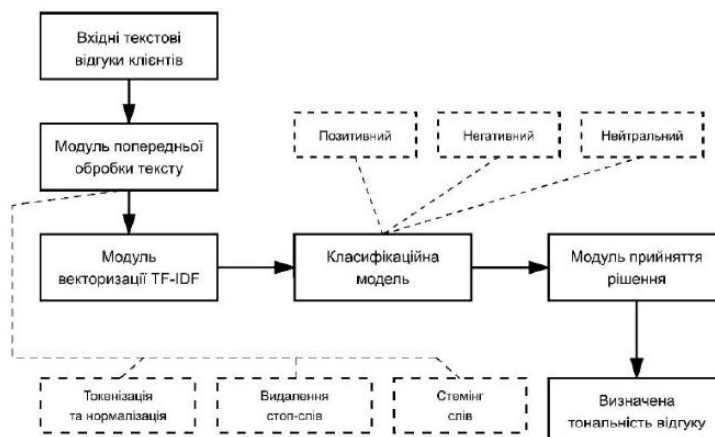


Рис. 1. Загальна схема методу аналізу тональності

Підготовлені векторні представлення текстів подаються на вхід класифікаційної моделі, яка оцінює ознаки тексту та визначає ймовірність належності до класів тональності. Результатом є визначена тональність відгуку з оцінкою впевненості моделі.

Використання методів машинного навчання дозволяє автоматизувати аналіз великої кількості відгуків, виявляти складні закономірності в даних та забезпечує однорідність класифікації. Модульна архітектура методу забезпечує гнучкість: окремі етапи обробки та класифікації можна модифікувати без суттєвих змін у системі.

Метод розроблено з урахуванням балансу між точністю класифікації та швидкістю обробки, забезпечуючи ефективну роботу як з короткими, так і з довгими відгуками, виявляючи ключові ознаки тональності незалежно від обсягу тексту.

Для вирішення задачі аналізу тональності клієнтських відгуків застосовано гібридний підхід, який поєднує традиційні методи машинного навчання з елементами глибокого навчання. Архітектура моделі побудована за принципом послідовної обробки даних, де кожен компонент використовує результати попереднього етапу для підвищення точності класифікації. Основу моделі становить послідовна комбінація методу опорних векторів (SVM) та Dense Neural Network (DNN), що дозволяє об'єднати переваги обох підходів. Метод опорних векторів ефективно працює з високорозмірними даними та визначає оптимальну гіперплощину для розділення класів, тоді як DNN виявляє складні нелінійні залежності та коригує передбачення базової моделі, використовуючи інформацію про впевненість у класифікації.

На вхід моделі подається векторне подання тексту, отримане за допомогою TF-IDF векторизації з додаванням біграм, що забезпечує врахування локального контексту слів. Розмірність векторного представлення складає близько 3500 ознак, що дозволяє моделі ефективно працювати з текстами різної довжини та структури. Загальна архітектура моделі представлена на рисунку.2.

Архітектура запропонованого методу аналізу тональності клієнтських відгуків побудована за принципом послідовної обробки даних і включає два основні компоненти. Перший компонент реалізовано як класифікатор на основі методу опорних векторів (SVM) з радіально-базисною функцією (RBF), що дозволяє моделювати складні нелінійні залежності між ознаками. Параметри ядра ( $\gamma$  та  $C$ ) підбиралися експериментально для забезпечення оптимальної точності класифікації. SVM генерує передбачення у вигляді ймовірностей для трьох класів тональності.

Другий компонент – Dense Neural Network (DNN) – приймає на вхід конкатеновані ознаки: оригінальні TF-IDF вектори з біграмами та ймовірності, отримані від SVM, формуючи вектор розмірністю 3503 ознаки.

Навчання проводиться послідовно: спершу SVM на повному навчальному наборі, після чого DNN навчається на конкатенованих ознаках, коригуючи помилки базового класифікатора відповідно до концепції мета-навчання (stacking).

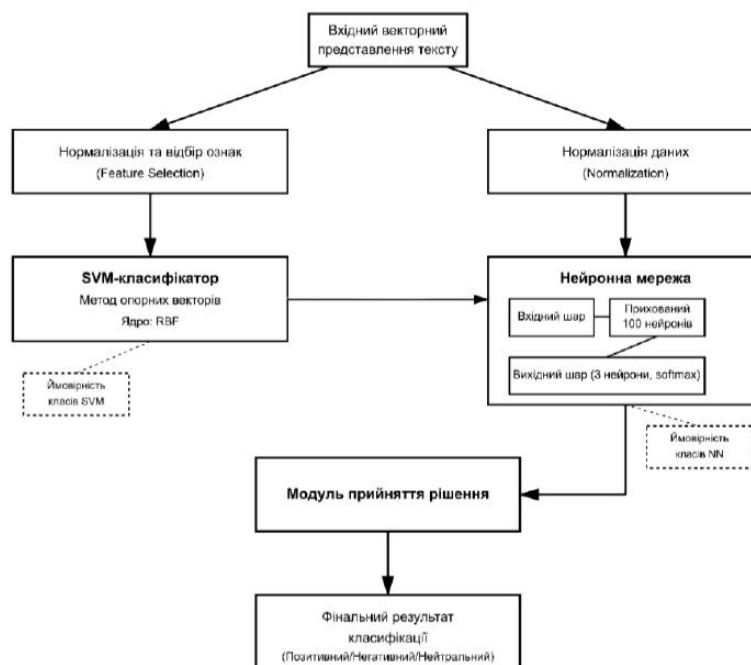


Рис.2. Архітектура гібридної моделі класифікації

Параметри мережі підбиралися експериментально з використанням крос-валідації. Оптимізатор Adam забезпечує адаптивне налаштування швидкості навчання, а функція втрат для багатокласової класифікації дозволяє ефективно штрафувати модель за невірні передбачення. Модульна структура архітектури дозволяє додавати нові шари або використовувати більш складні моделі, такі як LSTM або трансформери, а також комбінувати ознаки для підвищення якості класифікації. Послідовна архітектура забезпечує прозорість процесу: SVM виконує роль базового класифікатора, а DNN – роль коректора, що підвищує точність передбачень і робить систему гнучкою для подальшого розвитку.

#### Експериментальні дослідження

Для експериментального дослідження запропонованого методу аналізу тональності було сформовано набір даних із відгуків клієнтів про різні системи та послуги. Початковий датасет містив 8000 відгуків, вручну розмічених експертами за трьома класами тональності: позитивний, негативний та нейтральний. Для забезпечення високої якості розмітки кожен відгук оцінювався двома незалежними експертами, а у випадку розбіжностей залучався третій.

Перед навчанням моделі відгуки проходили попередню обробку: приведення до нижнього регістру, видалення спеціальних символів, URL та email, токенизацію та видалення стоп-слів. Тексти векторизувалися методом TF-IDF із включенням біграм, що забезпечує контекстне представлення слів; розмірність векторів становила близько 3500 ознак.

Статистичний аналіз показав середню довжину відгуку 68 слів (стандартне відхилення 34), позитивні відгуки були трохи довші за негативні та нейтральні. Частотний аналіз виявив індикаторні слова для кожного класу, що підтверджує можливість автоматичної класифікації на основі текстових ознак.

Забезпечено відсутність витоків інформації між навчальною та тестовою вибірками: відгуки одного продукту або користувача не потрапляли одночасно в різні підмножини. Виявлено, що близько 3% відгуків містять сарказм або іронію, а 2% – змішану тональність; у таких випадках клас визначався за переважаючим враженням.

Сформований набір даних забезпечує репрезентативність, збалансованість класів та розмаїття текстів, що створює надійну основу для оцінки ефективності запропонованого методу класифікації.

Dense Neural Network реалізовано як другий етап гібридної моделі класифікації. Прихований шар застосовує функцію активації ReLU, яка вводить нелінійність і дозволяє виявляти складні залежності між ознаками. Вихідний шар використовує Softmax для перетворення вихідних значень у ймовірнісний розподіл, де клас із найбільшою ймовірністю визначається як фінальний.

Навчання мережі здійснюється з використанням оптимізатора Adam зі швидкістю 0.001 для стабільної збіжності. Функція втрат стандартна для багатокласової класифікації. Динаміка процесу навчання відображена на рисунку 3.

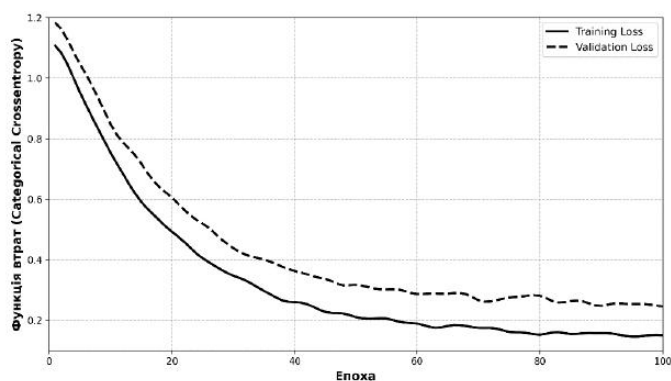


Рис. 3. Графік функції втрат при навчанні

На графіку (рис. 3) показано зміну значень функції втрат для навчальної та валідаційної вибірок протягом 100 епох навчання. На початку втрати були високими (~1,05) через випадкову ініціалізацію ваг. Протягом перших 20 епох спостерігалось різке зменшення втрат, що свідчить про швидке навчання моделі базовим закономірностям. Після 20-ї епохи темп зниження втрат уповільнюється, залишаючись стабільним до 60–70 епох; навчальна втрата на 100-й епісі досягає ~0,15, а валідаційна стабілізується близько 0,28. Після 70-ї епохи спостерігається початок перенавчання: навчальна втрата продовжує зменшуватися, тоді як валідаційна зростає.

На рисунку 4 представлено графік точності класифікації. Початкова точність становила ~65% для навчальної та ~63% для валідаційної вибірки, що перевищує випадковий рівень (~33%). Протягом перших 40 епох точність швидко зростає: навчальна до 88%, валідаційна – до 85%, що демонструє здатність моделі ефективно розпізнавати основні закономірності тональності відгуків.

Після 40-ї епохи темп зростання точності сповільнюється: на навчальній вибірці вона досягає 92% на 100-й епісі, тоді як на валідаційній – максимум ~88% на 70-й епісі, після чого стабілізується з незначними коливаннями. Стабілізація валідаційної точності свідчить про досягнення оптимального стану моделі на невідомих даних, а подальше навчання може призвести до перенавчання.

Аналіз кривих функції втрат та точності дозволив визначити оптимальну кількість епох навчання – 70, що забезпечує баланс між точністю класифікації та узагальнювальною здатністю. Поступова стабілізація валідаційної точності та невелика різниця між навчальною та валідаційною точністю до 60-ї епохи підтверджують ефективність підбору параметрів оптимізації та добру узагальнювальну здатність моделі.

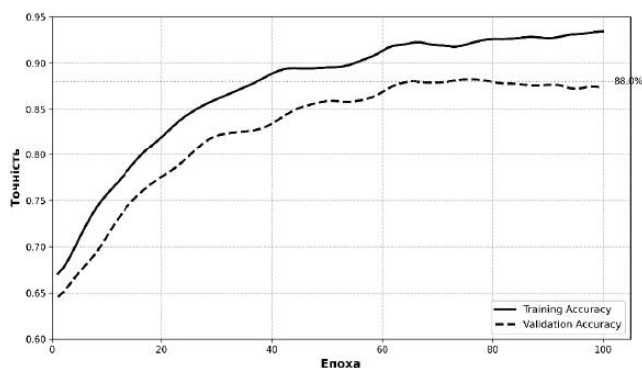


Рис. 4. Графік точності класифікації протягом процесу навчання

Графіки навчання (рис. 4) демонструють, що Dense Neural Network успішно коригує передбачення SVM-класифікатора, використовуючи інформацію про впевненість базової моделі. Одного прихованого шару виявилось достатньо для ефективної класифікації тональності, а більш глибокі архітектури не забезпечили суттєвого покращення результатів на цьому датасеті.

Після завершення навчання Dense Neural Network оцінювали на тестовій вибірці, де модель досягла точності 88,8%, що на 2,6% перевищує результати базового SVM-класифікатора. Це свідчить про здатність нейронної мережі виявляти додаткові закономірності та коригувати помилки базової моделі. Найсуттєвіше покращення точності Dense Neural Network спостерігається для нейтрального класу, що підтверджує її ефективність у вирішенні найбільш складного аспекту задачі класифікації.

Середнє значення F1-Score для Dense Neural Network становить 0.887, що на 0.032 (3.7%) перевищує результат SVM (0.855), забезпечуючи статистично значуще та практично вагоме покращення для задач аналізу тональності.

Для детальнішої оцінки було проведено порівняльний аналіз результатів класифікації обох моделей; його підсумки подано на рисунку 5. У 85% випадків обидві моделі формують правильні передбачення, що характерно для відгуків із чітко вираженою тональністю. У 7.2% випадків лише SVM забезпечує коректну класифікацію – переважно для коротких відгуків із простою структурою, де достатньо базових текстових ознак. Натомість у 3% випадків правильні результати отримує лише Dense Neural Network, що свідчить про її здатність розпізнавати складні контекстуальні зв'язки у довшіх і структурно насичених текстах, які SVM не виявляє.

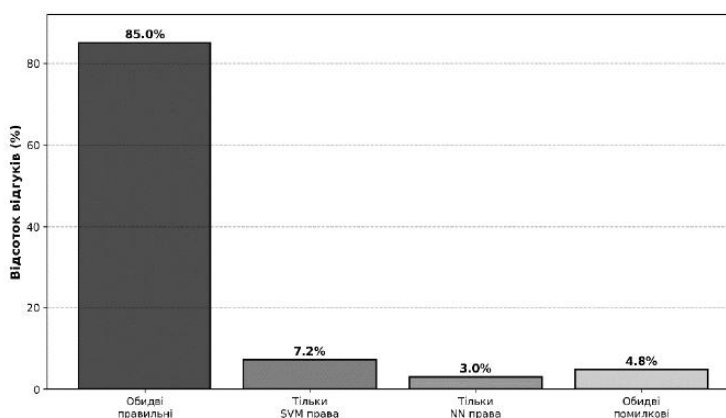


Рис.5. Порівняння помилок

Лише у 4.8% випадків обидві моделі формують некоректні передбачення, що характерно для відгуків із неоднозначною, змішаною або іронічною тональністю, які є складними для інтерпретації навіть експертами. Загалом у 10.2% випадків спостерігається розбіжність у помилках моделей, що підтверджує доцільність гібридного підходу: SVM ефективніший для одних типів текстів, тоді як Dense Neural Network краще опрацьовує інші. Така комплементарність моделей створює умови для підвищення загальної точності системи за рахунок комбінування їх передбачень.

Порівняння метрик для обох моделей за всіма класами тональності наведено на рисунку 6. Графік демонструє F1-Score для позитивного, негативного та нейтрального класів окремо, а також середнє значення, що дозволяє комплексно оцінити переваги Dense Neural Network над SVM.

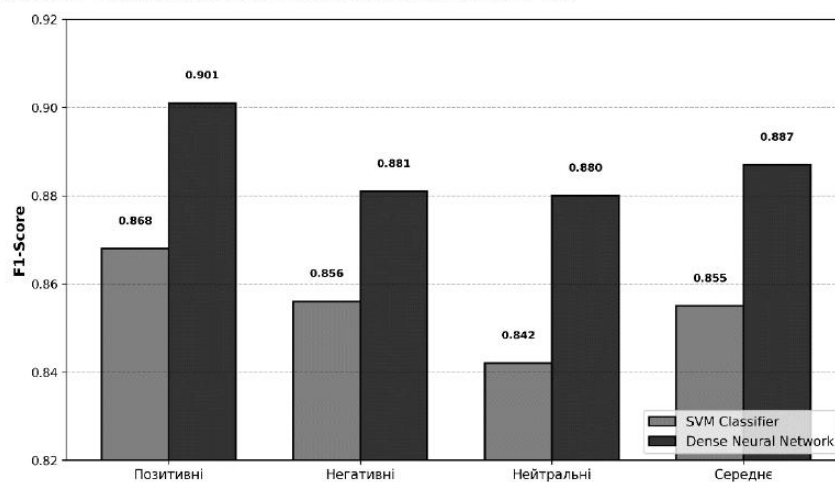


Рис. 6. Порівняння метрик Dense Neural Network та SVM-класифікатор

Для позитивного класу F1-Score Dense Neural Network становить 0.901, що на 0.033 перевищує результат SVM 0.868. Для негативного класу покращення становить 0.025 (0.881 проти 0.856), а для нейтрального – найбільше, 0.038 (0.880 проти 0.842). Середнє підвищення F1-Score дорівнює 0.032. Графік на рисунку 4.6 демонструє стабільну перевагу Dense Neural Network над SVM у всіх класах тональності, що особливо помітно для нейтрального класу, який є найскладнішим для класифікації.

Аналіз помилкових класифікацій виявив низку обмежень моделі. Dense Neural Network і надалі має труднощі з відгуками, що містять сарказм або іронію, наприклад: «чудовий продукт, якщо вам подобається викидати гроші», де позитивні лексеми вводять модель в оману. Складними залишаються також умовні конструкції – «було б добре, якби ціна була нижчою», які потребують тонкого контекстного аналізу. У випадках, коли текст містить багато фактів без явної оцінки, модель інколи визначає тональність як нейтральну, навіть якщо загальний контекст натякає на іншу оцінку.

Попри зазначені обмеження, Dense Neural Network демонструє суттєву перевагу над базовою SVM-моделлю: точність 88.8% та середній F1-Score 0.887 підтверджують ефективність нейронної мережі у коригуванні передбачень традиційних методів машинного навчання та її придатність для задачі тримірної класифікації тональності у розподілено-паралельному середовищі.

Фінальна гібридна модель поєднує переваги SVM-класифікатора та Dense Neural Network у послідовній архітектурі. На першому етапі SVM формує базове оцінювання тональності, тоді як Dense Neural Network виконує функцію коректора, аналізуючи як вихід SVM, так і вихідні текстові ознаки, що забезпечує більш точне фінальне рішення. Узагальнені результати порівняння представлені на рисунку 7.

Логістична регресія, використана як найпростіший базовий метод, забезпечила точність 81.2%, що підтверджує наявність лінійної компоненти у задачі класифікації тональності. Застосування SVM з RBF-ядром

підвищило точність до 86.2%, тобто на 5%, завдяки здатності моделі враховувати нелінійні залежності між ознаками. Додавання біграм до SVM дозволило досягти 87.5%, що на 1.3% перевищує результат базового SVM, та підтвердило важливість урахування локального контексту слів для підвищення точності класифікації.

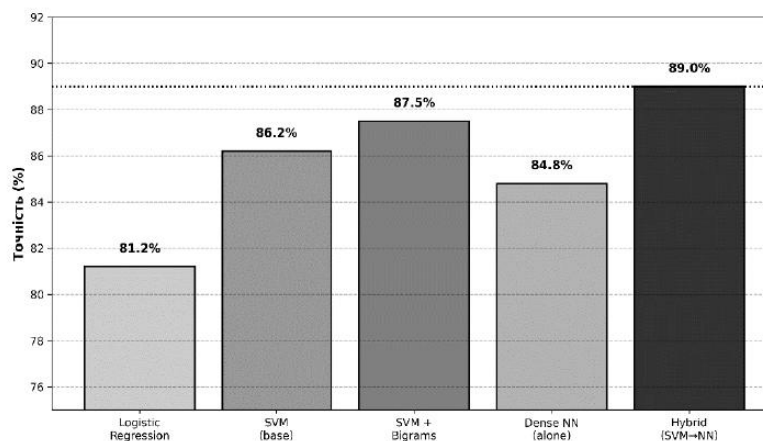


Рис. 7. Порівняння точності різних підходів до класифікації тональності

Dense Neural Network окремо, без використання передбачень SVM, забезпечила точність 84.8%. Це нижче за результат SVM з біграмами, але вище за базову SVM, що свідчить про здатність мережі виявляти складні закономірності навіть на відносно невеликих вибірках.

Фінальна гібридна модель, яка поєднує SVM з біграмами та Dense Neural Network у послідовній архітектурі, досягла найвищої точності – 89.0%. Покращення становить 1.5% порівняно з SVM з біграмами (87.5%) та 4.2% порівняно з окремою нейронною мережею (84.8%), що підтверджує ефективність запропонованого підходу.

Гібридний підхід демонструє стійкі переваги завдяки поєднанню сильних сторін обох методів: SVM забезпечує надійну базову класифікацію, біграми враховують локальний контекст, а Dense Neural Network коригує помилки, пов'язані з нелінійними залежностями у тексті.

Порівняння метрик якості між моделями показує середні значення F1-Score: 0.810 для логістичної регресії, 0.855 для базової SVM, 0.872 для SVM з біграмами, 0.847 для окремої Dense NN та 0.890 для гібридної моделі. Покращення спостерігається для всіх трьох класів: позитивного (з 0.871 до 0.901), негативного (з 0.870 до 0.881) та нейтрального (з 0.875 до 0.880). Незначне зростання для нейтрального класу пояснюється тим, що SVM з біграмами вже продемонструвала високу точність у цьому випадку.

Важливо, що гібридна модель не є середнім значенням окремих результатів: її точність 89.0% перевищує обидві складові моделі, що свідчить про те, що Dense Neural Network ефективно використовує інформацію про впевненість SVM для прийняття рішень. Коли передбачення SVM надійне, мережа його підтримує; за умов низької впевненості вона більше покладається на власні ознаки.

Аналіз помилок показав, що гібридна модель виправляє 35% хибних класифікацій SVM, що демонструє суттєве зниження кількості помилок. Проте модель все ще помиляється зокрема на відгуках з сарказмом, змішаною тональністю або культурно специфічними контекстами.

Попри ці обмеження, точність 89.0% є конкурентоспроможною та відповідає верхній межі результатів сучасних досліджень (85-92%). Простота та модульність запропонованої архітектури роблять її придатною для практичного застосування в розподілено-паралельних системах аналізу відгуків клієнтів. Результати експериментів підтверджують ефективність комбінування SVM та глибоких мереж для покращення класифікації тональності.

Узагальнення результатів експериментальних досліджень демонструє стабільне підвищення ефективності класифікації тональності під час переходу від простих лінійних моделей до комбінованих підходів. Логістична регресія формує базовий рівень якості з точністю 81.2%. Використання SVM з RBF-ядром забезпечує суттєве покращення (86.2%), що підтверджує переваги нелінійних методів у задачах текстової класифікації. Розширення ознакового простору за рахунок біграм підвищує точність до 87.5%, що підкреслює важливість врахування локального контексту слів.

Окрема Dense Neural Network демонструє 84.8% точності, дещо поступаючись SVM через обмежений обсяг навчальних даних. Найвищі результати забезпечує гібридна модель, що поєднує SVM та нейронну мережу в послідовній архітектурі, досягаючи точності 89.0% та збалансованих метрик якості для всіх класів. Покращення простежується для позитивних, негативних і нейтральних відгуків, що свідчить про здатність моделі коригувати помилки базового класифікатора.

Хоча гібридна модель потребує більше часу на навчання, продуктивність у класифікації нових текстів залишається високою та достатньою для використання в розподілено-паралельних системах оброблення даних. У порівнянні з іншими методами, що застосовуються в задачах аналізу тональності на подібних за масштабом датасетах, отримана точність є конкурентоспроможною та досягається за умов обмежених ресурсів.

Аналіз помилок свідчить, що найскладнішими для класифікації залишаються відгуки з сарказмом, іронією, культурними чи контекстними відсиланнями, а також змішаною тональністю. Водночас результати підтверджують ефективність послідовної гібридної архітектури, яка поєднує можливості традиційних методів машинного навчання та нейронних мереж і забезпечує підвищення точності порівняно з використанням кожного підходу окремо.

#### **Висновки**

У роботі проведено комплексне дослідження застосування технологій машинного навчання для автоматизованого аналізу тональності клієнтських відгуків у контексті розподілено-паралельних обчислень. Сформовано збалансований датасет, на основі якого виконано експериментальне порівняння традиційних та нейромережевих підходів. Оптимізація параметрів SVM-класифікатора з RBF-ядром засвідчила його високу точність 86.2%, що підтверджує доцільність використання методів опорних векторів для задач текстової класифікації.

На основі отриманих результатів розроблено гібридну модель послідовної архітектури, у якій нейронна мережа використовує оцінки впевненості SVM як додаткові ознаки. Запропонований метод продемонстрував найвищу точність – 89.0% перевищивши всі окремі моделі. Це підтвердило, що послідовне комбінування традиційних алгоритмів із глибокими нейронними мережами забезпечує суттєве зростання якості класифікації.

Розроблена модульна програмна система забезпечує гнучкість інтеграції, масштабованість та можливість ефективного розподіленого використання. Аналіз помилок показав, що складними залишаються тексти із сарказмом, іронією або культурно-специфічними відсиланнями, що визначає перспективи подальших досліджень, зокрема застосування попередньо навчених мовних моделей та контекстних ембеддингів.

Отримані результати підтверджують практичну цінність запропонованого гібридного підходу та його конкурентоспроможність у порівнянні з сучасними методами аналізу тональності. Метод забезпечує високу точність за умов обмежених навчальних даних і може бути ефективно впроваджений у системи автоматичного моніторингу відгуків у реальних IT-проектах та розподілено-паралельних обчислювальних середовищах.

#### **Література**

1. Kumar S., Roy P. P., Dogra D. P., Kim B.-G. A Comprehensive Review on Sentiment Analysis: Tasks, Approaches and Applications. arXiv, 2023. URL: <https://doi.org/10.48550/arXiv.2311.11250>.
2. Kirimlioglu E., Kung H., Orlando D. Were You Helpful -- Predicting Helpful Votes from Amazon Reviews. arXiv, 2024. URL: <https://doi.org/10.48550/arXiv.2412.02884>.

3. Jain P. K., Pamula R. A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews. arXiv, 2020. URL: <https://doi.org/10.48550/arXiv.2008.10282>.
4. Wang C., Zhu X., Yan L. Sentiment Analysis for E-Commerce Reviews Based on Deep Learning Hybrid Model / *2022 5th International Conference on Signal Processing and Machine Learning*, Dalian China, ACM, August 04, 2022. Pp. 38–46. URL: <https://doi.org/10.1145/3556384.3556391>.
5. Sunil N., Shirazi F. Customer Review Classification Using Machine Learning and Deep Learning Techniques / *Social Computing and Social Media: 15th International Conference, SCSM 2023, Held as Part of the 25th HCI International Conference, HCII 2023, Copenhagen, Denmark, July 23–28, 2023, Proceedings, Part I*, Berlin, Heidelberg, Springer-Verlag, 2023. Pp. 581–597. URL: [https://doi.org/10.1007/978-3-031-35915-6\\_42](https://doi.org/10.1007/978-3-031-35915-6_42).
6. Hong X., Guan S.-U., Wong P. W. H., Xue N., Man K. L., Liu D., Li Z. Lifelong Machine Learning-Based Quality Analysis for Product Review / *Proceedings of the 3rd International Conference on Advanced Information Science and System*, New York, NY, USA, Association for Computing Machinery, 2022. Pp. 1–5. URL: <https://doi.org/10.1145/3503047.3503154>.
7. Peng S., Chen J., Chen C., Dai M. Sentiment and User Churn Analysis for Online Platform Business using Yelp Studies / *Proceedings of the 2024 3rd International Conference on Algorithms, Data Mining, and Information Technology*, New York, NY, USA, Association for Computing Machinery, 2025. Pp. 18–25. URL: <https://doi.org/10.1145/3701100.3701105>.
8. Kumar S., Roy P. P., Dogra D. P., Kim B.-G. A Comprehensive Review on Sentiment Analysis: Tasks, Approaches and Applications. arXiv, 2023. URL: <https://doi.org/10.48550/arXiv.2311.11250>.
9. Hussain S., Dhanda N., Verma R. Sentiment Analysis of Amazon Product Reviews using VADER and RoBERTa Models / *2023 8th International Conference on Communication and Electronics Systems (ICCES)*, June 2023. Pp. 708–713. URL: <https://doi.org/10.1109/ICCES57224.2023.10192872>.
10. Das K. K. Kunal-Kumar-Das191049/Sentimental-Analysis-of-Amazon-Reviews : Jupyter Notebook URL: <https://github.com/Kunal-Kumar-Das191049/Sentimental-Analysis-of-Amazon-Reviews>.
11. Bharadwaj R., Shendurkar S., Kadam T., Patekar U., Waghule S. RoBERTa and VADER Sentiment Analysis: A Comparative Approach / *Big Data Analytics and Data Science*, Singapore, Springer Nature, 2024. Pp. 141–154. URL: [https://doi.org/10.1007/978-981-97-8666-4\\_12](https://doi.org/10.1007/978-981-97-8666-4_12).
12. Kong H. Investigating the Effect of Gender and Age in Sentiment Analysis Training and Testing: RoBERTa Vader, Naive Bayes, SVM and Logistic Regression / *Proceedings of the 2024 8th International Conference on Software and e-Business*, New York, NY, USA, Association for Computing Machinery, 2025. Pp. 8–18. URL: <https://doi.org/10.1145/3715885.3715899>.
13. Alantari H. J., Currim I. S., Deng Y., Singh S. An empirical comparison of machine learning methods for text-based sentiment analysis of online consumer reviews. *International Journal of Research in Marketing*. 2022. Vol. 39, No. 1. Pp. 1–19. URL: <https://doi.org/10.1016/j.ijresmar.2021.10.011>.
14. Hossain M. S., Rahman M. F. Customer Sentiment Analysis and Prediction of Insurance Products' Reviews Using Machine Learning Approaches. *FIIIB Business Review*. 2023. Vol. 12, No. 4. Pp. 386–402. URL: <https://doi.org/10.1177/23197145221115793>.
15. Manziuk E., Barmak O., Krak I., Petliak N., Jin Z., Radiuk P. Explainable Deep Learning for Interpretable Brain Tumor Diagnosis from MRI Images / *Lecture Notes in Data Engineering, Computational Intelligence, and Decision-Making, Volume 1*, Cham, Springer Nature Switzerland, 2024. Pp. 326–348. URL: [https://doi.org/10.1007/978-3-031-70959-3\\_17](https://doi.org/10.1007/978-3-031-70959-3_17).

16. Manziuk E., Krak I., Barmak O., Mazurets O., Kuznetsov V., Pylypiak O. Structural alignment method of conceptual categories of ontology and formalized domain. *CEUR-WS*. 2021. Vol. 3003. Pp. 11–22.
17. Pavlyshyn V., Ryzhanskyi O., Manziuk E., Radiuk P., Barmak O., Krak I. Establishing Patterns of the Urban Transport Flows on Clustering Analysis. *CEUR Workshop Proceedings*. 2025. Vol. 3974. Pp. 1–9. URL: <https://ceur-ws.org/Vol-3974/paper01.pdf>.
18. Barmak O., Krak I., Yakovlev S., Manziuk E., Radiuk P., Kuznetsov V. Toward explainable deep learning in healthcare through transition matrix and user-friendly features. *Frontiers in Artificial Intelligence*. 2024. Vol. 7. Pp. 1482141. URL: <https://doi.org/10.3389/frai.2024.1482141>.
19. Pavlyshyn V., Manziuk E., Barmak O., Radiuk P., Krak I. An Adaptive Machine Learning Approach to Sustainable Traffic Planning: High-Fidelity Pattern Recognition in Smart Transportation Systems. *Future Transportation*. 2025. Vol. 5, No. 4. URL: <https://doi.org/10.3390/futuretransp5040152>.
20. Ryzhanskyi O., Pavlyshyn V., Radiuk P., Manziuk E., Barmak O., Krak I. AI-Driven Traffic Signal Control System to Reduce CO2 Emissions. *CEUR Workshop Proceedings*. 2025. Vol. 3974. Pp. 18–27. URL: <https://ceur-ws.org/Vol-3974/paper02.pdf>.
21. An adaptive approach to detecting fake news based on generalized text features(Conference Paper) Shupta, A., Barmak, O., Wierzbicki, A., Skrypnyk, T. // 7th International Conference on Computational Linguistics and Intelligent Systems. Volume I: Machine Learning Workshop, CoLInS 2023; Kharkiv; Ukraine; 20 April 2023 до 21 April 2023; Код 188444 // *CEUR Workshop Proceedings Volume 3387*, 2023, Pages 300-310
22. E. Manziuk, O. Barmak, I. Krak, O. Mazurets, and T. Skrypnyk, “Formal Model of Trustworthy Artificial Intelligence Based on Standardization..” in *CEUR Workshop Proceedings*, Khmelnytskyi, Ukraine, Mar. 2021, vol. 2853, pp. 190–197. <http://ceur-ws.org/Vol-2853/>

## Додаток Б

### Програмний код: посилання на GitHub-репозиторій, структура проєкту та опис основних папок і файлів

#### Посилання на репозиторій на GitHub:

<https://github.com/bohdanzahurskiy1/project>

#### Вигляд сторінки репозиторію

File Name	Action	Time
base_classifier.py	Added project	6 minutes ago
bigram_generator.py	Added project	6 minutes ago
config.py	Added project	6 minutes ago
data_loader.py	Added project	6 minutes ago
hybrid_classifier.py	Added project	6 minutes ago
model_evaluator.py	Added project	6 minutes ago
nn_classifier.py	Added project	6 minutes ago
normalizer.py	Added project	6 minutes ago
predict.py	Added project	6 minutes ago
stop_words.py	Added project	6 minutes ago
svm_classifier.py	Added project	6 minutes ago
text_preprocessor.py	Added project	6 minutes ago
tfidf_vectorizer.py	Added project	6 minutes ago
tokenizer.py	Added project	6 minutes ago
train.py	Added project	6 minutes ago
vectorizer_manager.py	Added project	6 minutes ago

#### Конфігурація та дані

- `config.py` - централізована конфігурація: шляхи до даних, параметри моделей, константи
- `data_loader.py` - завантаження CSV з відгуками та розподіл на `train/val/test` вибірки

#### Попередня обробка тексту

- `tokenizer.py` - токенизація тексту на окремі слова, видалення пунктуації
- `stop_words.py` - видалення стоп-слів (службових слів)
- `normalizer.py` - нормалізація токенів: стеммінг (відсікання закінчень), приведення до нижнього регістру
- `text_preprocessor.py` - об'єднує всі етапи попередньої обробки в один пайплайн

## Векторизація

- `bigram_generator.py` - генерація та фільтрація біграм для кращого представлення контексту
- `tfidf_vectorizer.py` - TF-IDF векторизація для перетворення токенів у числові вектори
- `vectorizer_manager.py` - управління всією векторизацією: комбінує TF-IDF та біграми

## Моделі класифікації

- `base_classifier.py` - абстрактний базовий клас для всіх класифікаторів
- `svm_classifier.py` - SVM класифікатор з RBF ядром для класифікації тональності
- `nn_classifier.py` - Dense Neural Network з прихованим шаром та ReLU активацією
- `hybrid_classifier.py` - гібридна модель: SVM + Neural Network (SVM передає ймовірності в NN)

## Оцінка та використання

- `model_evaluator.py` - обчислення метрик: accuracy, precision, recall, F1-score, матриця помилок
- `train.py` - основний скрипт навчання гібридної моделі
- `predict.py` - завантаження навченої моделі та передбачення тональності для нових відгуків

## Додаток В

### Презентаційний матеріал

КВАЛІФІКАЦІЙНА РОБОТА МАГІСТРА

## МЕТОД АНАЛІЗУ ТОНАЛЬНОСТІ ВІДГУКІВ КЛІЄНТІВ З ВИКОРИСТАННЯМ МАШИННОГО НАВЧАННЯ



**Виконав:**

*студент 2 курсу, групи КНм-24-1*

*Богдан ЗАГУРСЬКИЙ*

**Керівник:**

*д.т.н., професор кафедри КН*

*Едуард МАНЗЮК*



2

## Актуальність

Практичне значення дослідження обумовлене зростаючою потребою компаній у швидкій та об'єктивній обробці великих масивів текстових відгуків від споживачів. Традиційний підхід до аналізу, який передбачає ручну обробку коментарів спеціалістами, характеризується низькою швидкістю виконання, високою вартістю, схильністю до суб'єктивних оцінок і неможливістю масштабування для великих підприємств.

Сучасні досягнення у сфері обробки природної мови та машинного навчання відкривають нові можливості для створення автоматизованих систем, здатних з високою точністю визначати настрої текстових повідомлень. Впровадження таких систем дозволяє організаціям оперативно виявляти проблемні аспекти своїх продуктів чи послуг, своєчасно реагувати на негативні тенденції, покращувати якість обслуговування клієнтів та підвищувати загальну конкурентоспроможність на ринку.

## Мета і задачі роботи

**Метою роботи** є підвищення точності автоматичної класифікації тональності клієнтських відгуків через розробку гібридного методу, що базується на інтеграції різних алгоритмів машинного навчання.

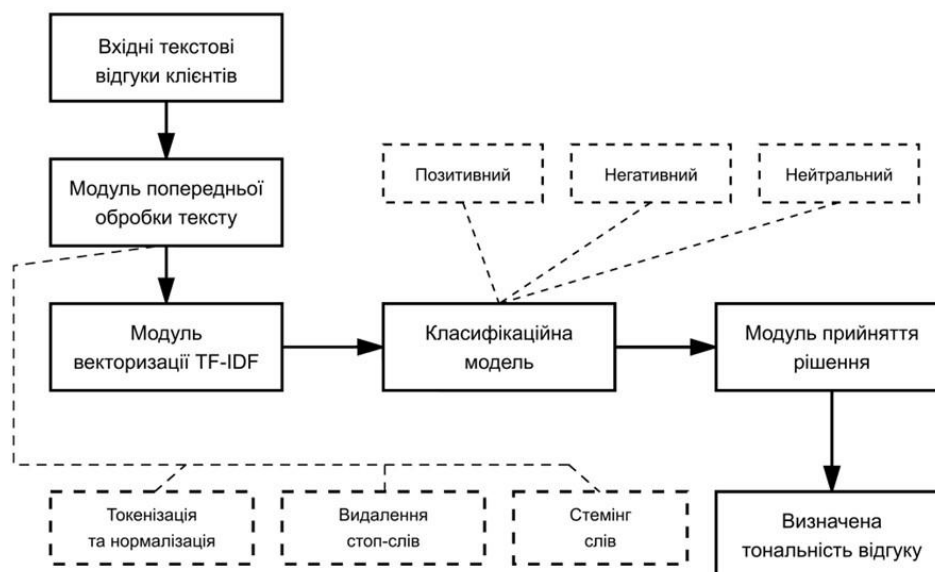
**Об'єкт дослідження** – процес автоматизованого визначення емоційного забарвлення у текстових відгуках споживачів.

**Предмет дослідження** – архітектури, алгоритми та програмні засоби для класифікації тональності текстів із використанням комбінованих методів машинного навчання.

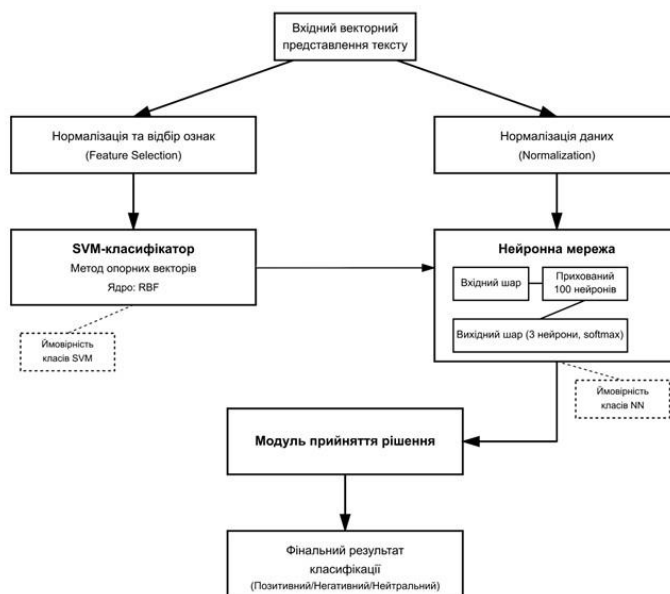
### Задачі дослідження:

- здійснити огляд наявних методологій та технологічних рішень у сфері автоматичного визначення тональності текстових даних, включаючи традиційні статистичні підходи та сучасні нейромережеві архітектури;
- розробити метод класифікації емоційного забарвлення відгуків, який використовує послідовну інтеграцію для досягнення підвищеної точності розпізнавання;
- розробити модифікацію базового методу шляхом включення аналізу біграм на етапі формування текстових ознак, що дозволить враховувати локальний контекст слів та покращити розпізнавання складних лінгвістичних конструкцій;
- створити програмну реалізацію запропонованого методу з використанням модульної архітектури, що забезпечить можливість незалежного розвитку окремих компонентів системи;
- провести експериментальне дослідження розробленого методу на репрезентативному наборі реальних клієнтських відгуків з оцінкою ключових метрик якості класифікації.

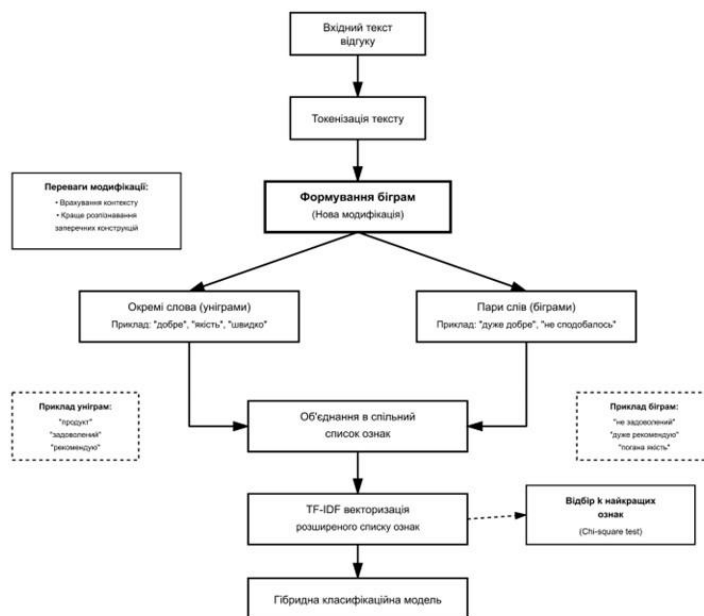
Загальна схема методу аналізу тональності



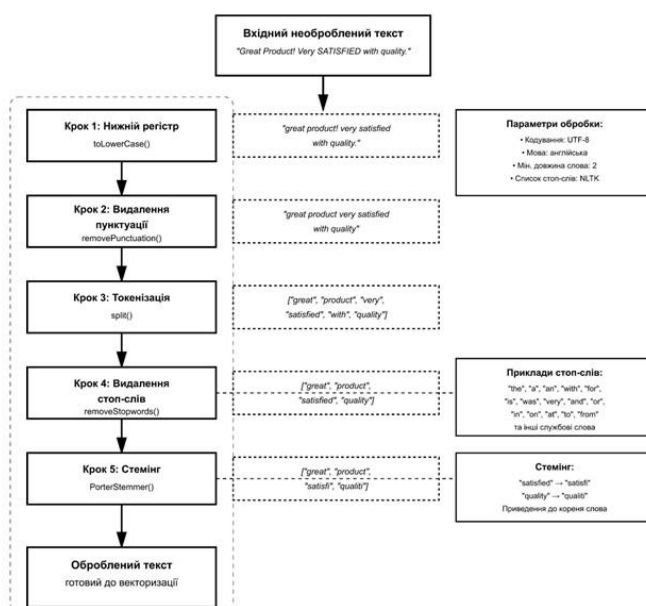
## Архітектура гібридної моделі класифікації



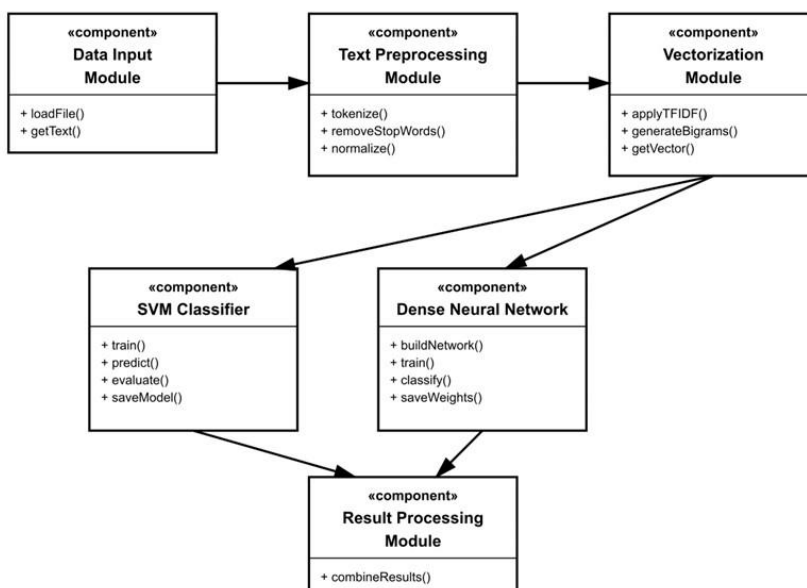
## Модифікована архітектура з аналізом біграм



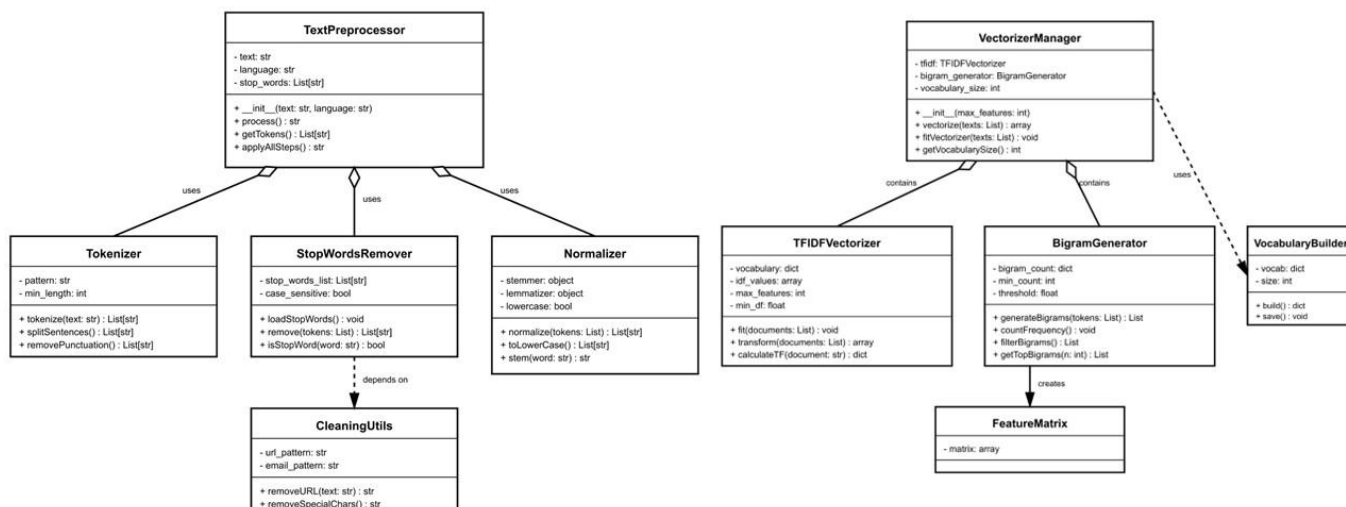
## Схема попередньої обробки текстових даних



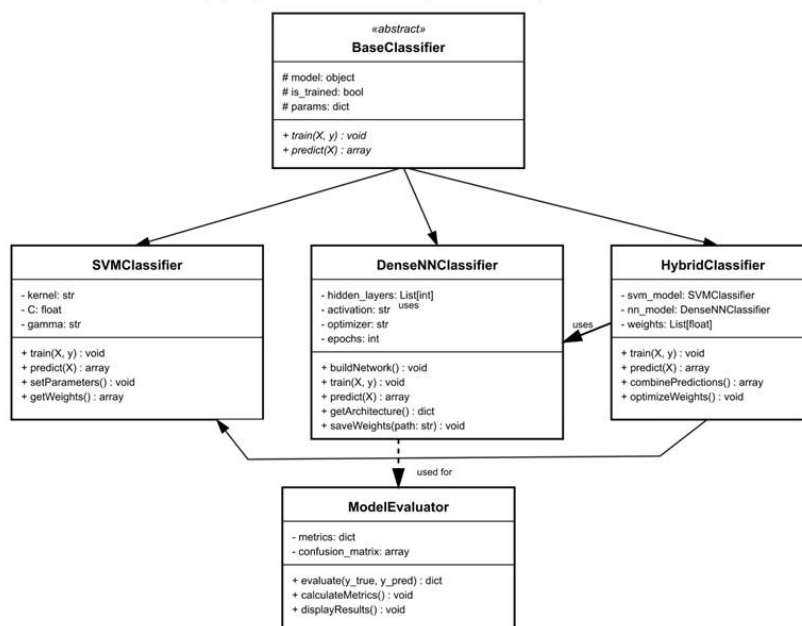
## Діаграма компонентів методу аналізу тональності відгуків



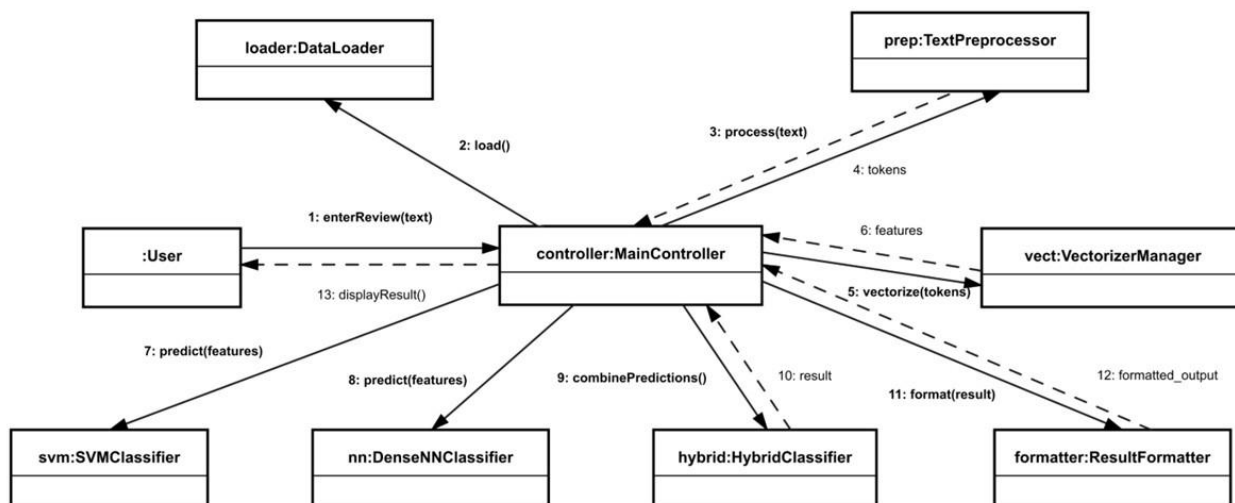
Діаграма класів модуля векторизації



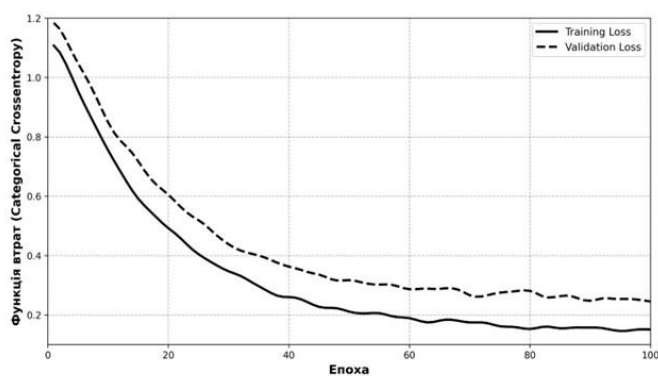
Діаграма класів модуля класифікації



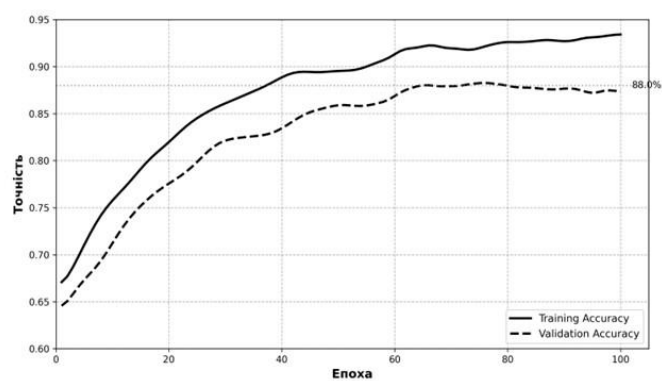
Діаграма взаємодії елементів методу



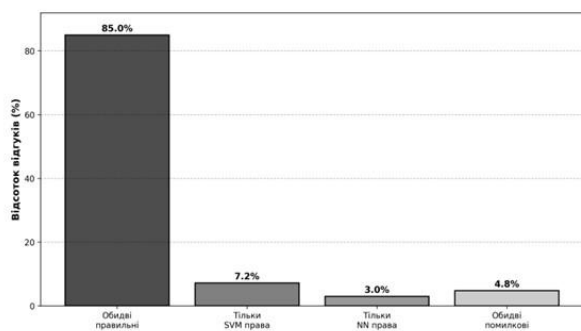
Графік функції втрат при навчанні



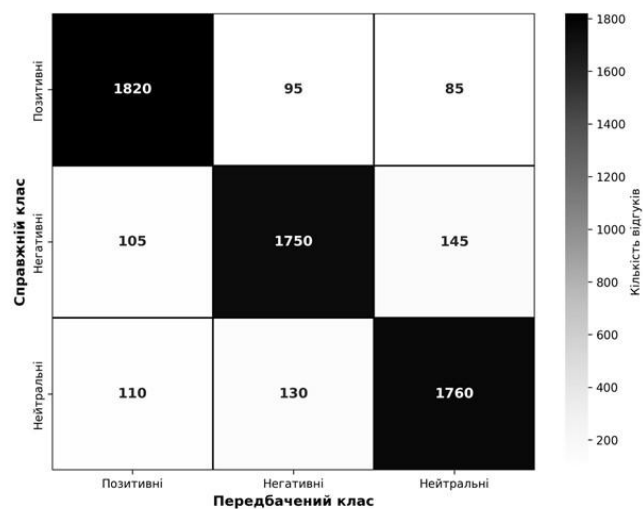
Графік точності класифікації протягом процесу навчання



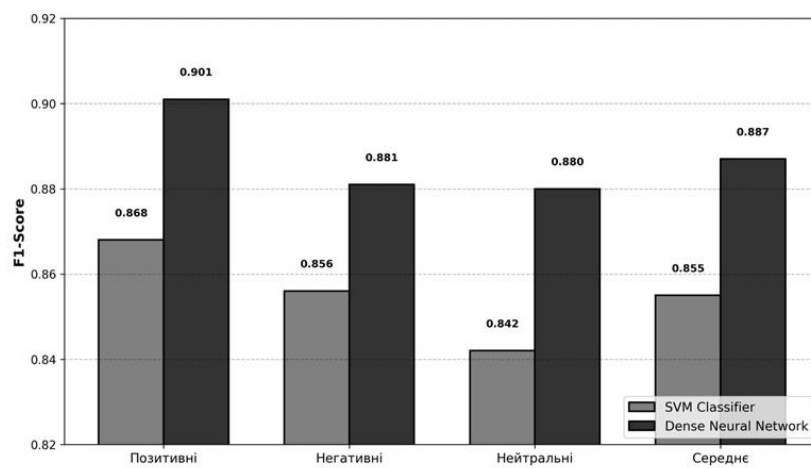
## Порівняння помилок



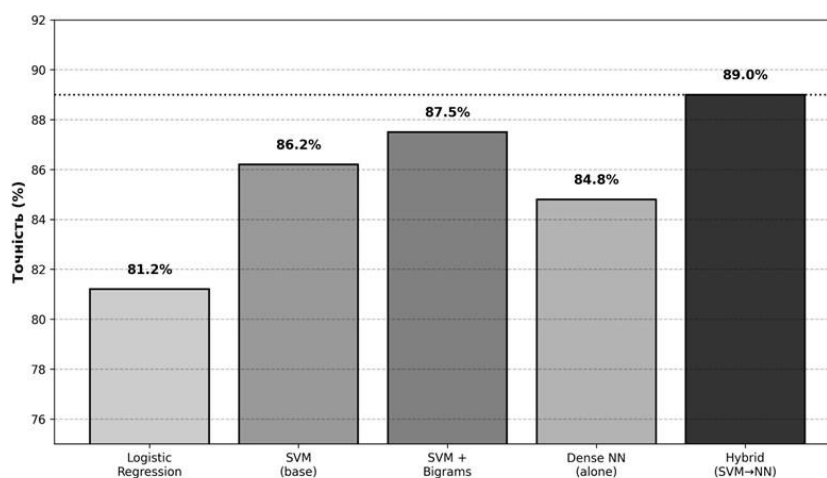
## Матриця помилок для Dense Neural Network



## Порівняння метрик Dense Neural Network та SVM-класифікатор



### Порівняння точності різних підходів до класифікації тональності



## Висновки

У кваліфікаційній роботі магістра вирішено актуальну задачу підвищення точності автоматизованого визначення емоційного забарвлення клієнтських відгуків через розробку та впровадження гібридного методу класифікації на основі інтеграції традиційних алгоритмів машинного навчання з нейромережевими компонентами.

Задачі дослідження:

- зроблено огляд наявних методологій та технологічних рішень у сфері автоматичного визначення тональності текстових даних, включаючи традиційні статистичні підходи та сучасні нейромережеві архітектури;
- розроблено метод класифікації емоційного забарвлення відгуків, який використовує послідовну інтеграцію для досягнення підвищеної точності розпізнавання;
- розроблено модифікацію базового методу шляхом включення аналізу біграм на етапі формування текстових ознак, що дозволить враховувати локальний контекст слів та покращити розпізнавання складних лінгвістичних конструкцій;
- створено програмну реалізацію запропонованого методу з використанням модульної архітектури, що забезпечить можливість незалежного розвитку окремих компонентів системи;
- проведено експериментальне дослідження розробленого методу на репрезентативному наборі реальних клієнтських відгуків з оцінкою ключових метрик якості класифікації.

ДЯКУЮ ЗА УВАГУ!

---

## Протокол аналізу звіту подібності науковим керівником

Заявляю, що я ознайомився (-лась) з Повним звітом подібності, який був згенерований Системою виявлення і запобігання плагіату щодо роботи:

**Автор:** Богдан ЗАГУРСЬКИЙ

**Співавтор:**

**Назва:** КВАЛІФІКАЦІЙНА РОБОТА на тему Метод аналізу тональності відгуків клієнтів з використанням машинного навчання

**Науковий керівник:** Едуард МАНЗІОК, д.т.н., професор

**Підрозділ:** Кафедра комп'ютерних наук

**Коефіцієнт подібності 1:** 2.9%

**Коефіцієнт подібності 2:** 0.7%

**Мікропробіли:** 0

**Заміна букв:** 1

**Інтервали:** 0

**Білі знаки:** 0

**Дата створення звіту:** 2025-12-13 21:04:57.0

Після аналізу Звіту подібності констатую наступне:

Запозичення, виявлені в роботі є законними і не є плагіатом. Рівень подібності не перевищує допустимої межі. Таким чином робота незалежна і приймається.

Запозичення не є плагіатом, але перевищено граничне значення рівня подібностей. Таким чином робота повертається на доопрацювання.

Виявлено запозичення і плагіат або навмисні текстові спотворення (маніпуляції), як передбачувані спроби укриття плагіату, які роблять роботу невідповідною вимогам законодавства (Ст. 32. ЗУ Про вищу освіту, пункт 3.1, Ст. 42. ЗУ Про освіту) та вимог НАЗЯВО (Критерій 5), а також кодексу етики і процедурам. Таким чином робота не приймається.

Обґрунтування:

2025-12-13

Дата

експерт

*Петровський С. Р. etn*

## Anti-Plagiarism (UA) v-15.281 Educational

The maximum coincidence with one document 1.0%

Dictionaries check: en\_US, ru\_RU, ua\_UA. Errors in the documents: 9%

ID: 252745 Title: КВАЛІФІКАЦІЙНА РОБОТА на тему Метод аналізу тональності відгуків клієнтів з використанням машинного навчання Added in a DB: 2025-12-13 Authors: Богдан ЗАГУРСЬКИЙ Heads: Едуард МАНЗЮК Consultants: Opponents:	Document		Sum coincidence on the DB	
	Symbols	Lexemes	Symbols	Lexemes
	110752	1706	2584 (2%)	45 (3%)

### Plagiarism sources

ID	Description	Plagiarism presence in the document	
		Symbols	Lexemes

**РІШЕННЯ ЕКСПЕРТНОЇ КОМІСІЇ КАФЕДРИ КОМП'ЮТЕРНИХ НАУК**

**ПРО ДОПУСК КВАЛІФІКАЦІЙНОЇ РОБОТИ ДО ЗАХИСТУ**

Назва кваліфікаційної роботи Метод аналізу тональності відгуків клієнтів з використанням машинного навчання

Автор студент групи КНм-24-1 Богдан ЗАГУРСЬКИЙ

Освітня програма Комп'ютерні науки

Рівень вищої освіти другий (магістерський)

Спеціальність 122 – Комп'ютерні науки

Науковий керівник: д.т.н., проф. каф. комп'ютерних наук Едуард МАНЗЮК

На основі аналізу кваліфікаційної роботи на дотримання вимог академічної доброчесності (у т.ч. відсутності ознак академічного плагіату) з урахуванням результатів перевірки роботи спеціалізованим програмними засобами комісія зробила такий висновок:

№	Висновок	Позначка про відповідність
1	Ознаки академічного плагіату	
1.1	Запозичення, виявлені в роботі, є законними і не є академічним плагіатом (далі – зазначаються підстави віднесення запозичень до правомірних, якщо потрібно). Робота приймається до захисту.	<i>відповідає</i>
1.2	Виявлені запозичення не є академічним плагіатом, розміщені в розділах, які не описують безпосередньо авторське дослідження, але кількість цитат перевищує обсяг, виправданий поставленою метою роботи (далі – зазначаються детальні та аргументовані підстави віднесення запозичень до правомірних). Робота приймається до захисту, але має бути відкоригована.	
1.3	Виявлені запозичення не є академічним плагіатом, але частково розміщені в розділах, які описують безпосередньо авторське дослідження, а кількість цитат перевищує обсяг, виправданий поставленою метою роботи. Робота може бути допущена до захисту після того як буде відкоригована та доопрацьована і успішно пройде повторну перевірку на академічний плагіат.	
1.4	Робота містить навмисні текстові спотворення, передбачувані спроби укриття текстових запозичень або інші прояви академічного плагіату. Робота містить фабрикацію або фальсифікацію даних. Робота не допускається до захисту.	
2	Інші види порушень академічної доброчесності	<i>відсутні</i>

Підтвердження:

*Запозичення, виявлені в роботі Богдана ЗАГУРСЬКОГО, не є плагіатом, оскільки: запозичення розміщені в розділі огляду існуючих підходів, не описують безпосередньо авторську роботу і не стосуються її результатів; усі запозичення фрагментарні; до запозичень входять фрагменти, які не мають авторства і містять поширені конструкції та загальновідомі терміни, скорочення. Рівень подібності не перевищує допустимої межі. Таким чином, робота є законною та приймається до захисту.*

*Обсяг запозичень, визначений системами виявлення збігів/ідентичності/схожості:*

*- за системою Anti-Plagiarism: 1%;*

*- за системою StrikePlagiarism КПІ: 2,93%*

13.12.2025

Завідувач кафедри



Олександр БАРМАК

Гарант освітньої програми



Руслан БАГРІЙ

Керівник кваліфікаційної роботи



Едуард МАНЗЮК



ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ  
МОН УКРАЇНИ



Кафедра комп'ютерних наук

## ВІДГУК ОПОНЕНТА

на кваліфікаційну роботу магістра

студента гр. КНм-24-1 Богдана ЗАГУРСЬКОГО

за темою Метод аналізу тональності відгуків клієнтів з використанням машинного навчання

### 1. Актуальність обраної теми

Актуальність обраної теми зумовлена критичною необхідністю автоматизації процесів аналізу великих обсягів клієнтських відгуків у сучасних системах електронної комерції та обслуговування споживачів. Застосування методів машинного навчання для визначення емоційного забарвлення текстів дозволяє значно підвищити ефективність роботи компаній, швидко виявляти проблемні аспекти продуктів чи послуг, оперативно реагувати на негативні відгуки та покращувати загальну якість обслуговування клієнтів.

### 2. Відповідність роботи предметній області спеціальності 122 Комп'ютерні науки та загальним вимогам до наукових робіт

Магістерська робота повністю відповідає предметній області спеціальності 122 Комп'ютерні науки, оскільки ґрунтується на застосуванні методів машинного навчання, обробки природної мови та нейронних мереж для вирішення задачі класифікації тональності текстових відгуків. Робота також відповідає загальним вимогам до наукових робіт, маючи чітку структуру.

### 3. Повнота розкриття мети та завдань дослідження

Мета та завдання дослідження розкриті повністю. Автор чітко формулює мету роботи – підвищення точності автоматичної класифікації тональності клієнтських відгуків через розробку гібридного методу. Для досягнення мети послідовно вирішуються поставлені завдання, що включають аналіз існуючих підходів, розробку гібридної архітектури з використанням SVM-класифікатора та Dense Neural Network, реалізацію попередньої обробки текстових даних, модифікацію базового методу шляхом включення біграм.

### 4. Наявність наукової новизни

Наукова новизна роботи полягає в удосконаленні методу визначення тональності клієнтських відгуків, який відрізняється від існуючих застосувань послідовної

гібридної архітектури, де Dense Neural Network використовує як вхідні дані не лише векторизовані текстові ознаки з біграмами.

#### **5. Зміст кожного розділу роботи**

Робота містить чотири розділи. В першому розділі представлено аналіз існуючих методів аналізу тональності текстових документів та систем автоматизованої обробки відгуків клієнтів. Другий розділ містить розробку методу визначення тональності з використанням гібридної архітектури. Третій розділ присвячено програмній реалізації розробленого методу та його модифікації з включенням біграм до векторного представлення. Розділ чотири містить експериментальне дослідження ефективності методу.

#### **6. Ступінь розкриття теми роботи**

Тема роботи розкрита повністю. Автор аналізує проблематику аналізу тональності відгуків, розглядає існуючі методи векторизації тексту та машинного навчання, обґрунтовує необхідність розробки удосконаленого гібридного методу. Детально описано запропонований підхід на основі послідовної інтеграції SVM-класифікатора з RBF-ядром та Dense Neural Network.

#### **7. Якість оформлення кваліфікаційної роботи**

Якість оформлення кваліфікаційної роботи відповідає встановленим академічним стандартам, демонструючи чіткість, послідовність та професійність у структурі, форматуванні та презентації матеріалу. Робота містить необхідні ілюстрації, таблиці, діаграми та графіки, що допомагають у розумінні запропонованого гібридного методу та результатів експериментальних досліджень.


#### **8. Недоліки кваліфікаційної роботи**

Робота могла б включати більш глибокий аналіз помилок класифікації для випадків з саркастичним або іронічним контекстом, де навіть удосконалена модель може мати труднощі.

#### **9. Загальний висновок (допускається чи не допускається до захисту), якої оцінки заслуговує кваліфікаційна робота**

Враховуючи рівень виконання та забезпечення усіх необхідних вимог, якість проведених експериментальних досліджень та отримані результати, робота може бути допущена до захисту. Рекомендована оцінка – відмінно.

Опонент С.Т.И., проректор АКИТКУ

  
Мероніон Б. Б.



## ВІДГУК НАУКОВОГО КЕРІВНИКА

### на кваліфікаційну роботу магістра

студента КНм-24-1 Богдана ЗАГУРСЬКОГО

за темою Метод аналізу тональності відгуків клієнтів з використанням машинного навчання

#### 1. Актуальність теми

Актуальність теми зумовлена критичною необхідністю автоматизації процесів аналізу клієнтських відгуків та обробки великих обсягів текстових даних у сучасних системах електронної комерції та обслуговування клієнтів. Можливість автоматично визначати емоційне забарвлення відгуків за допомогою методів машинного навчання відіграє ключову роль у підвищенні ефективності роботи компаній, оптимізації якості продуктів та послуг і зменшенні витрат на ручну обробку відгуків.

#### 2. Відповідність роботи предметній області Стандарту спеціальності 122 Комп'ютерні науки

Робота відповідає предметній області спеціальності 122 "Комп'ютерні науки", оскільки вона ґрунтується на застосуванні методів машинного навчання, зокрема методу опорних векторів та нейронних мереж прямого поширення. Дослідження передбачає використання алгоритмів векторизації тексту TF-IDF, структур даних, аналізу та попередньої обробки текстових даних.

#### 3. Професійні та особистісні якості

Під час роботи над магістерським дослідженням Богдан ЗАГУРСЬКИЙ продемонстрував високий рівень професійної компетентності в галузі комп'ютерних наук, відповідально та ефективно вирішуючи завдання з розробки методу класифікації тональності відгуків клієнтів з використанням гібридного підходу на основі інтеграції SVM-класифікатора та нейронної мережі. Студент проявив наполегливість у проведенні експериментальних досліджень та глибоке розуміння принципів машинного навчання.

#### 4. Ступінь самостійності під час виконання кваліфікаційної роботи

При виконанні магістерської роботи студент виявив високий рівень самостійності, запропонувавши удосконалення базового методу класифікації шляхом включення біграм до векторного представлення тексту. Студент самостійно провів аналіз літератури, розробив гібридну архітектуру моделі, реалізував програмне забезпечення та провів експериментальні дослідження.

#### 5. Наукова новизна та оригінальність запропонованих підходів

Удосконалено метод визначення тональності клієнтських відгуків, який відрізняється від існуючих застосуванням послідовної гібридної архітектури, де Dense Neural Network використовує як вхідні дані не лише векторизовані текстові ознаки з біграмами, а й ймовірнісні оцінки від попередньо навченого SVM-класифікатора, що забезпечило підвищення точності визначення емоційного забарвлення текстів.

#### **6. Ступінь оволодіння методами дослідження**

Під час роботи над магістерським дослідженням студент продемонстрував глибоке розуміння та вміле застосування методів комп'ютерних наук, зокрема методів обробки природної мови, векторизації тексту TF-IDF, методу опорних векторів, нейронних мереж прямого поширення та статистичного аналізу результатів, для вирішення задачі класифікації тональності текстових документів, що свідчить про його високий рівень оволодіння сучасними методами дослідження в галузі машинного навчання.

#### **7. Повнота та якість розкриття теми роботи**

У магістерській роботі тема розкрита повністю та ґрунтовно. Робота відзначається логічною структурою, глибиною аналізу існуючих підходів до аналізу тональності текстів, детальним описом розробленого гібридного методу, його програмною реалізацією та експериментальним дослідженням, що свідчить про високий рівень розуміння предметної області та здатність до самостійного наукового пошуку.

#### **8. Логічність, послідовність, аргументованість, літературна грамотність викладення матеріалу**

Магістерська робота характеризується чіткою логічною структурою, послідовним викладенням матеріалу від аналізу проблематики та існуючих підходів до розробки та експериментальної перевірки власного гібридного методу, аргументованістю висновків. Автор демонструє високий рівень літературної грамотності, дотримуючись наукового стилю викладення та забезпечуючи легкість сприйняття тексту.

#### **9. Можливість практичного застосування кваліфікаційної роботи, окремих її частин**

Розроблений у магістерській роботі метод аналізу тональності відгуків клієнтів з використанням машинного навчання має широкі можливості практичного застосування в галузях електронної комерції, управління репутацією брендів, аналізу соціальних мереж та моніторингу якості обслуговування клієнтів.

#### **10. Висновок про можливість допуску кваліфікаційної роботи до захисту, на яку оцінку заслуговує робота**

Враховуючи належний рівень виконання та забезпечення усіх необхідних вимог, робота може бути допущена до захисту. Рекомендована оцінка «відмінно».

Керівник



д.т.н., професор каф. КН Едуард МАНЗЮК