

КВАЛІФІКАЦІЙНА РОБОТА

Інформаційна система визначення емоційного ставлення аудиторії до бренду на основі аналізу даних соцмереж
Назва теми

Рівень вищої освіти перший (бакалаврський)

Галузь знань 12 «Інформаційні технології»
Шифр, назва

Спеціальність 126 «Інформаційні системи та технології»
Шифр, назва

Освітня програма «Інформаційні системи та технології»
Назва

Шифр КвРІСТ 2401217.22.01.02 ПЗ

Виконав здобувач IV курсу, група ІСТ-22-1



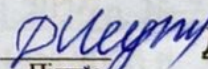
Підпис

Владислав САМАРУК

Ініціали, прізвище

Керівник

канд. техн. наук, доцент
Науковий ступінь, учене звання



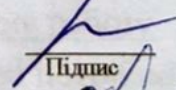
Підпис

Дмитро МЕДЗАТИЙ

Ініціали, прізвище

Нормоконтролер

канд. фіз.-мат. наук, доцент
Науковий ступінь, учене звання

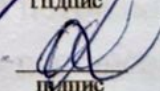


Підпис

Тетяна КИСІЛЬ

Ініціали, прізвище

До захисту допускаю:
завідувач кафедри КІС
«01» червня 2026 р.



Підпис

Ольга ПАВЛОВА

Ініціали, прізвище

дата

Хмельницький 2026

ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

Факультет ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Кафедра КОМП'ЮТЕРНОЇ ІНЖЕНЕРІЇ ТА ІНФОРМАЦІЙНИХ СИСТЕМ

Рівень вищої освіти ПЕРШИЙ (БАКАЛАВРСЬКИЙ)

Галузь знань 12 ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

Спеціальність 126 ІНФОРМАЦІЙНІ СИСТЕМИ ТА ТЕХНОЛОГІЇ

Освітня програма «ІНФОРМАЦІЙНІ СИСТЕМИ ТА ТЕХНОЛОГІЇ»

ЗАТВЕРДЖУЮ

Завідувачка кафедри КІС

 Ольга ПАВЛОВА

“ 10 ” 01 2026 р.

ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ

Самаруку Владиславу Володимировичу

Прізвище, ім'я, по батькові студента

1. Тема проекту (роботи) Інформаційна система визначення емоційного ставлення аудиторії до бренду на основі аналізу даних соцмереж

Керівник проекту (роботи) Медзатий Дмитро Миколайович, к.т.н., доцент.

Прізвище, ім'я, по батькові, науковий ступінь, вчене звання

Затверджена наказом ректора університету від 20.01.2026 р. № 7

2. Термін подання здобувачем роботи на кафедру 01.06.2026 р.

3. Вихідні дані до роботи Завдання на кваліфікаційну роботу

4. Зміст пояснювальної записки (перелік питань, які потрібно

розробити) Дослідження предметної області та постановка задачі розроблення інформаційної системи аналізу емоційного ставлення до бренду.

Проектування інформаційної системи аналізу емоційного ставлення аудиторії до бренду.

Програмна реалізація інформаційної системи.

5. Перелік графічного матеріалу (із зазначенням обов'язкових креслень)

Архітектура ПЗ проекту

Архітектура проекту

Алгоритм роботи ІС

6. Консультанти розділів кваліфікаційної роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв

7. Дата видачі завдання « 10 » 01 2026 р.

КАЛЕНДАРНИЙ ПЛАН

№з/п	Назва етапів (розділів) дипломного проекту (роботи)	Термін виконання етапів проекту (роботи)	Примітка
1	Вибір напрямку дослідження та узгодження тематики кваліфікаційної роботи з керівником	10.01.2026	виконано
2	Ознайомлення з предметною областю; формулювання мети та задач дослідження; визначення об'єкта та предмета дослідження	01.02.2026	виконано
3	Робота над розділом 1 - дослідження предметної області визначення емоційного ставлення аудиторії до бренду та постановка задачі розроблення інформаційної системи	01.03.2026	виконано
4	Робота над розділом 2 - проєктування інформаційної системи визначення емоційного ставлення аудиторії до бренду на основі аналізу даних соціальних мереж	01.04.2026	виконано
5	Робота над розділом 3 – програмна реалізація інформаційної системи	29.04.2026	виконано
6	Оформлення пояснювальної записки згідно вимог	25.05.2026	не виконано
7	Попередній захист ВКР	26.05.2025	не виконано
8	Захист ВКР на засіданні ЕК	Червень 2026 року	

Здобувач

Підпис

Ім'я, ПРІЗВИЩЕ

Владислав САМАРУК

Керівник кваліфікаційної роботи

Підпис

Дмитро МЕДЗАТИЙ

Ім'я, ПРІЗВИЩЕ

АНОТАЦІЯ

Тема кваліфікаційної роботи: «Інформаційна система визначення емоційного ставлення аудиторії до бренду на основі аналізу даних соцмереж».

Автор роботи: Владислав САМАРУК.

Керівник роботи: Дмитро МЕДЗАТИЙ.

Пояснювальна записка: 79 с., 13 рис., 5 табл., 1 дод., 50 джерел.

Графічна частина: 3 креслення.

АНАЛІЗ ДАНИХ, БАЗА ДАНИХ, БРЕНД, ЕМОЦІЙНЕ СТАВЛЕННЯ, ІНФОРМАЦІЙНА СИСТЕМА, МАШИННЕ НАВЧАННЯ, СОЦІАЛЬНІ МЕРЕЖІ.

Кваліфікаційна робота бакалавра присвячена розробці та дослідженню інформаційної системи визначення емоційного ставлення аудиторії до бренду на основі аналізу даних соціальних мереж. Актуальність теми зумовлена активним розвитком цифрових каналів комунікації, зростанням ролі соціальних мереж у формуванні репутації компаній та необхідністю оперативного виявлення позитивних, негативних і нейтральних реакцій користувачів.

У роботі розглянуто особливості збору, попередньої обробки, збереження та аналізу текстових даних із соціальних мереж. Досліджено методи визначення тональності тексту, підходи до класифікації емоційного забарвлення повідомлень, а також можливості використання алгоритмів машинного навчання та засобів обробки природної мови для автоматизованого аналізу користувацького контенту.

Метою роботи є проектування інформаційної системи, яка забезпечує збір повідомлень із соціальних мереж, їх очищення, аналіз тональності, збереження результатів у базі даних та подання аналітичної інформації у зручному для користувача вигляді. Практичне значення роботи полягає у можливості застосування розробленої системи для моніторингу репутації бренду, оцінювання ефективності маркетингових кампаній, виявлення негативних інформаційних хвиль та підтримки управлінських рішень у сфері цифрового маркетингу.



Підпис здобувача

30.05.2026

Дата

ЗМІСТ

Вступ.....	4
1 Інформаційна система визначення емоційного ставлення аудиторії до бренду та постановка задачі щодо її розроблення	6
1.1 Аналіз предметної області і виявлення наявних проблем і завдань.....	6
1.2 Порівняльний аналіз переваг та недоліків існуючих рішень.....	8
1.3 Підходи до вирішення задачі за темою дослідження	12
1.4 Постановка задачі.....	15
1.5 Висновки до першого розділу	16
2 Проєктування інформаційної системи визначення емоційного ставлення аудиторії до бренду.....	17
2.1 Визначення програмних підсистем інформаційної системи.....	17
2.2 Проєктування архітектури інформаційної системи	20
2.3 Проєктування бази даних інформаційної системи	26
2.4 Визначення функціональних та нефункціональних вимог до інформаційної системи	29
2.5 Моделювання потоків даних інформаційної системи.....	32
2.6 Проєктування користувацького інтерфейсу системи.....	34
2.7 Проєктування системи безпеки та захисту даних	35
2.8 Висновки до другого розділу.....	37
3 Програмна реалізація інформаційної системи визначення емоційного ставлення аудиторії до бренду	39
3.1 Опис реалізації модулів програмного забезпечення інформаційної системи	39
3.2 Опис середовища розроблення та структури програмного проєкту	40
3.3 Опис процесу створення баз даних	44

					КвРІСТ 2401217.22.01.02 ПЗ			
Зм.	Арк.	№ док.ум.	Підпис	Дата	Інформаційна система визначення емоційного ставлення аудиторії до бренду на основі аналізу даних соцмереж. Пояснювальна записка	Літера	Арк.вщ.	Арк.вщів
Виконав	Владислав Самарук					у	2	74
Перевір.	Дмитро Медзатий							
Н.контр.	Тетяна КИСІЛЬ					ХНУ ІСТ-22-1		
Затвер.	Ольга ПАВЛОВА							

3.4	Опис реалізації алгоритму визначення емоційного ставлення аудиторії до бренду	47
3.5	Опис реалізації користувацького інтерфейсу	52
3.6	Тестування роботи інформаційної системи	55
3.7	Висновки до третього розділу	59
	Висновки.....	63
	Перелік джерел посилань	66
	Додаток А «Архітектура ПЗ проєкту»	72
	Додаток Б «Архітектура проєкта».....	73
	Додаток В «Алгоритм роботи ІС».....	74

ВСТУП

Актуальність дослідження. У сучасних умовах цифрової трансформації соціальні мережі стали одним із головних каналів комунікації між брендами та їхньою аудиторією. Користувачі щоденно залишають велику кількість коментарів, відгуків, згадок, реакцій та повідомлень, які відображають їхнє ставлення до компаній, товарів, послуг, рекламних кампаній і публічних подій. Такі дані мають значну цінність для бізнесу, оскільки дозволяють оцінити репутацію бренду, визначити рівень задоволеності споживачів, виявити проблемні аспекти діяльності компанії та своєчасно реагувати на негативні інформаційні прояви.

Традиційні методи аналізу громадської думки, такі як опитування, анкетування або фокус-групи, потребують значних часових і фінансових витрат. Крім того, вони не завжди дозволяють оперативно отримати актуальну інформацію щодо настроїв аудиторії. Натомість аналіз даних соціальних мереж дає змогу виконувати моніторинг емоційного ставлення користувачів у режимі, близькому до реального часу. Саме тому виникає потреба у створенні інформаційних систем, які автоматизують процес збору, обробки, класифікації та візуалізації даних із соціальних мереж.

Емоційне ставлення аудиторії до бренду може бути позитивним, негативним або нейтральним. Позитивні згадки свідчать про прихильність користувачів, задоволеність якістю товарів чи послуг, підтримку маркетингових активностей. Негативні згадки можуть вказувати на проблеми з якістю обслуговування, невдалі рекламні кампанії, репутаційні ризики або поширення критичних інформаційних повідомлень. Нейтральні повідомлення, у свою чергу, часто містять фактичну інформацію без вираженого емоційного забарвлення. Особливу складність становить автоматизоване опрацювання текстів природною мовою. Повідомлення в соціальних мережах часто містять скорочення, помилки, жаргон, емодзі, хештеги, посилання та неформальні мовні конструкції. Це

					КвРІСТ 2401217.22.01.02 ПЗ	Арк.
						4
Зм.	Арк.	№ докум.	Підпис	Дата		

потребує використання спеціальних методів попередньої обробки тексту, алгоритмів машинного навчання та моделей аналізу тональності.

Метою дипломної роботи є проєктування та реалізація інформаційної системи визначення емоційного ставлення аудиторії до бренду на основі аналізу даних соціальних мереж. Для досягнення поставленої мети необхідно виконати такі завдання:

- 1) провести аналіз предметної області визначення емоційного ставлення аудиторії до бренду;
- 2) дослідити сучасні підходи до збору та аналізу даних соціальних мереж;
- 3) визначити функціональні та нефункціональні вимоги до інформаційної системи;
- 4) спроектувати архітектуру інформаційної системи;
- 5) розробити структуру бази даних для збереження повідомлень, брендів, результатів аналізу та статистичних показників;
- 6) описати реалізацію основних програмних модулів системи;
- 7) виконати тестування роботи інформаційної системи.

Об'єктом дослідження є процес аналізу текстових даних соціальних мереж для визначення емоційного ставлення аудиторії до бренду.

Предметом дослідження є методи, моделі та програмні засоби автоматизованого визначення тональності повідомлень користувачів соціальних мереж.

Практичне значення роботи полягає у створенні інформаційної системи, яка може бути використана маркетинговими підрозділами, PR-фахівцями, аналітиками та керівниками компаній для оперативного моніторингу репутації бренду та підтримки прийняття управлінських рішень.

					КвРІСТ 2401217.22.01.02 ПЗ	Арк.
						5
Зм.	Арк.	№ докум.	Підпис	Дата		

1 ІНФОРМАЦІЙНА СИСТЕМА ВИЗНАЧЕННЯ ЕМОЦІЙНОГО СТАВЛЕННЯ АУДИТОРІЇ ДО БРЕНДУ ТА ПОСТАНОВКА ЗАДАЧІ ЩОДО ЇЇ РОЗРОБЛЕННЯ

1.1 Аналіз предметної області і виявлення наявних проблем і завдань

У сучасному цифровому середовищі соціальні мережі виконують роль не лише комунікаційної платформи, а й потужного середовища формування репутації бренду. Користувачі щодня створюють величезні обсяги контенту: публікують відгуки, коментарі, огляди, рекомендації, скарги, а також мультимедійні матеріали - фотографії, відео, сторіс. На відміну від традиційних медіа, інформація в соціальних мережах поширюється миттєво та має високий рівень довіри серед аудиторії, оскільки часто базується на особистому досвіді.

Емоційне ставлення споживачів до бренду є динамічним показником. Воно може змінюватися під впливом рекламних кампаній, запуску нових продуктів, сервісних збоїв, публічних заяв компанії або зовнішніх інформаційних подій. Негативний інформаційний привід здатен за короткий час спричинити репутаційну кризу, тоді як позитивна кампанія – суттєво підвищити рівень лояльності. У таких умовах своєчасне виявлення змін у настроях аудиторії стає стратегічно важливим завданням.

Предметна область дослідження охоплює процеси збору, обробки та аналізу даних із відкритих онлайн-джерел з метою визначення емоційного ставлення до бренду. Основу цієї області складають технології соціального моніторингу (Social Listening) та аналізу тональності (Sentiment Analysis)[1], що використовують методи обробки природної мови (Natural Language Processing, NLP), машинного навчання та штучного інтелекту.

Незважаючи на активний розвиток таких технологій, у предметній області існує низка проблем.

					КвРІСТ 2401217.22.01.02 ПЗ	Арк.
						6
Зм.	Арк.	№ докум.	Підпис	Дата		

По-перше, це проблема великих обсягів даних. Щодня генеруються тисячі й мільйони повідомлень, що потребують автоматизованої обробки. Ручний аналіз є практично неможливим через масштаб і швидкість оновлення інформації.

По-друге, складність інтерпретації емоцій. Людська мова містить сарказм, іронію, сленг, багатозначність та культурні особливості. Алгоритмам складно коректно визначати справжній емоційний зміст повідомлення, особливо у випадках непрямой критики або жартівливих висловлювань.

По-третє, існує проблема мультимедійності контенту. Значна частина сучасних згадувань брендів відбувається у вигляді зображень або відео без текстового супроводу. Це потребує застосування методів комп'ютерного зору та аналізу візуальних ознак.[2]

Окремою проблемою є мовна локалізація. Багато міжнародних систем демонструють нижчу точність при роботі з українською мовою через обмеженість навчальних корпусів і специфіку мовних конструкцій.

Крім того, слід враховувати обмеження доступу до даних, що виникають через політику соціальних платформ, а також наявність спаму та бот-активності, які спотворюють реальну картину настроїв аудиторії.[14]

1) У межах предметної області можна сформулювати такі ключові завдання:

2) автоматизований збір даних із різних джерел (соціальні мережі, форуми, медіа);

3) фільтрація нерелевантного контенту, спаму та ботів;

4) визначення тональності повідомлень (позитивна, негативна, нейтральна);

5) глибший аналіз емоцій та аспектів продукту;

6) виявлення трендів і потенційних кризових ситуацій;

7) візуалізація результатів для підтримки управлінських рішень;

8) адаптація алгоритмів до мовних і галузевих особливостей.

					КвРІСТ 2401217.22.01.02 ПЗ	Арк. 7
Зм.	Арк.	№ докум.	Підпис	Дата		

Таким чином, предметна область характеризується високою динамічністю, значними обсягами даних та складністю інтерпретації емоційного контенту. Це обумовлює необхідність створення або адаптації ефективних інформаційних систем, здатних поєднати сучасні методи NLP, аналізу даних і гнучке налаштування під конкретні потреби бренду.

1.2 Порівняльний аналіз переваг та недоліків існуючих рішень

На сучасному етапі розвитку інформаційних технологій існує значна кількість програмних рішень для соціального моніторингу та аналізу тональності. Вони відрізняються рівнем функціональності, складністю впровадження, глибиною аналітики та вартістю використання. Загалом їх можна поділити на три основні групи: корпоративні платформи, інструменти для малого та середнього бізнесу, а також хмарні й open-source засоби для створення власних систем.[6]

Корпоративні платформи забезпечують найповніший набір функцій. Вони дозволяють здійснювати збір даних з великої кількості джерел, аналізувати тональність і емоції, визначати ключові теми обговорень, сегментувати аудиторію та формувати інтерактивні аналітичні панелі. Такі рішення характеризуються високою масштабованістю та підтримкою багатьох мов. Водночас їхніми недоліками є висока вартість ліцензії, складність впровадження та необхідність професійного налаштування для досягнення високої точності результатів.

Інструменти для малого та середнього бізнесу орієнтовані на швидкий запуск і простоту використання. Вони надають можливість моніторингу згадувань у реальному часі та базового визначення тональності повідомлень. Перевагою таких систем є доступна ціна та інтуїтивний інтерфейс. Однак глибина аналізу обмежена, а точність визначення емоцій у складних мовних конструкціях може бути недостатньою.

Хмарні NLP-сервіси пропонують інший підхід – вони надають інструменти для аналізу тексту у вигляді API, які можна інтегрувати у власну систему. Їхньою перевагою є гнучкість, масштабованість і можливість побудови індивідуальної

					КвРІСТ 2401217.22.01.02 ПЗ	Арк. 8
Зм.	Арк.	№ докум.	Підпис	Дата		

архітектури. Проте такі сервіси не здійснюють самостійний збір даних і потребують повноцінної технічної реалізації всієї системи, що збільшує складність проєкту.

Open-source рішення дозволяють створити власну модель аналізу тональності з максимальною кастомізацією. Використання сучасних трансформерних моделей забезпечує високу точність, особливо за умови донавчання на специфічних корпусах даних. Разом із тим такі рішення потребують значних обчислювальних ресурсів, кваліфікованих фахівців та розробки додаткових модулів збору й візуалізації даних.[17]

Порівняльний аналіз показує, що жодне з існуючих рішень не є універсальним. Корпоративні платформи забезпечують максимальну функціональність, але мають високу вартість. Доступні інструменти простіші, проте обмежені у глибині аналітики. Хмарні та open-source засоби дають найбільшу гнучкість, однак вимагають власної розробки та технічної експертизи.

Особливою проблемою залишається адаптація моделей до української мови, складність обробки сарказму, іронії та контекстно залежних висловлювань, а також вплив бот-активності на результати аналізу. Це підтверджує актуальність розробки або вдосконалення спеціалізованої системи, адаптованої до мовних і ринкових особливостей. Розберемо приклади таких існуючих платформ. [6]

Brandwatch є однією з найбільш функціонально насичених платформ корпоративного рівня. Система забезпечує масштабний збір даних із соціальних мереж, блогів, форумів та онлайн-медіа, підтримує багатомовний аналіз і надає розширені інструменти візуалізації результатів. Платформа дозволяє здійснювати кластеризацію тем, аналіз частки голосу (Share of Voice), відстеження змін настроїв у часовому розрізі та порівняння з конкурентами.

Перевагою Brandwatch є глибина аналітики та можливість обробки великих обсягів історичних даних. Водночас система потребує професійного налаштування запитів і фільтрів для досягнення високої точності. Основними

					КвРІСТ 2401217.22.01.02 ПЗ	Арк. 9
Зм.	Арк.	№ докум.	Підпис	Дата		

недоліками є висока вартість ліцензії та складність впровадження, що робить її доступною переважно для великого бізнесу.

Talkwalker відрізняється інтеграцією текстового та візуального аналізу. Платформа здатна розпізнавати логотипи й бренди на зображеннях і відео, що дозволяє враховувати згадування без текстового супроводу. Система також підтримує аналіз емоцій, виявлення трендів і автоматичне сповіщення про потенційні кризові ситуації.

Сильними сторонами Talkwalker є використання сучасних AI-технологій і глобальне покриття джерел. Однак висока вартість та складність конфігурації обмежують використання платформи невеликими компаніями. Крім того, для максимальної ефективності потрібна аналітична команда.

Meltwater поєднує соціальний моніторинг із медіа-аналітикою, що дозволяє формувати комплексне уявлення про інформаційне поле навколо бренду. Система зручна для PR-відділів, оскільки охоплює як соціальні мережі, так і новинні ресурси.

Перевагою є можливість інтегрованого аналізу та створення звітів для стратегічного планування. Недоліком є менш глибокий емоційний аналіз порівняно з AI-орієнтованими платформами, а також значні фінансові витрати на впровадження.

Sprinklr є омніканальною платформою управління клієнтським досвідом, яка включає модулі соціального моніторингу та аналізу настроїв. Вона дозволяє інтегрувати дані з соціальних мереж із CRM-системами та іншими корпоративними інструментами.

Основною перевагою є можливість кастомізації моделей машинного навчання під специфіку бренду або галузі. Недоліками виступають складність технічної реалізації, висока вартість і потреба в IT-підтримці.

Brand24 орієнтований на малий і середній бізнес. Система дозволяє швидко налаштувати моніторинг згадувань і отримувати базову оцінку тональності в реальному часі. Перевагами є доступність, простота використання та швидкий

					КвРІСТ 2401217.22.01.02 ПЗ	Арк. 10
Зм.	Арк.	№ докум.	Підпис	Дата		

запуск. Недоліками є обмежена глибина аналізу, нижча точність у складних мовних контекстах і відсутність розширених AI-функцій.

Awarío надає можливість створювати гнучкі пошукові запити для відстеження онлайн-репутації. Система забезпечує базовий аналіз тональності та формування аналітичних звітів. Перевагою є оперативність і відносно невисока вартість. Недоліком - обмежені можливості емоційної класифікації та відсутність глибокого аспектного аналізу.

Google Cloud Natural Language є хмарним сервісом, який надає API для аналізу тональності та визначення сутностей у тексті. Його можна використовувати як ядро власної інформаційної системи. Перевагою є масштабованість і проста інтеграція. Недоліком - відсутність механізму збору даних і потреба в розробці повної архітектури системи. [35]

AWS Comprehend забезпечує аналіз тональності, ключових фраз та сутностей із можливістю інтеграції в інфраструктуру AWS. Підходить для потокової обробки великих обсягів даних. Перевагою є гнучкість і масштабування, недоліком - технічна складність реалізації та платна модель використання.

Моделі на базі трансформерів, доступні через Hugging Face, забезпечують високу точність аналізу тональності та емоцій, особливо після донавчання на доменно-специфічних даних. Перевагою є максимальна гнучкість і можливість адаптації до української мови. Недоліками є потреба в обчислювальних ресурсах, спеціалістах з машинного навчання та створенні власної інфраструктури збору й обробки даних.

Узагальнюючи проведений аналіз, можна зробити висновок, що корпоративні платформи забезпечують найбільш комплексний функціонал, але мають високу вартість. Інструменти для малого бізнесу є доступнішими, проте обмежені в глибині аналітики. Хмарні та open-source рішення надають гнучкість і можливість адаптації, однак потребують технічної реалізації. Це підтверджує

					КвРІСТ 2401217.22.01.02 ПЗ	Арк. 11
Зм.	Арк.	№ докум.	Підпис	Дата		

доцільність розробки власної системи, адаптованої до конкретних вимог та мовних особливостей досліджуваного середовища.

1.3 Підходи до вирішення задачі за темою дослідження

Оскільки метою дослідження є створення інформаційної системи для визначення емоційного ставлення аудиторії до бренду в соціальних мережах, доцільно застосувати комплексний програмно-аналітичний підхід, що поєднує технології збору даних, обробки природної мови та машинного навчання.

Для реалізації системи доцільним є використання мови програмування Python, яка має розвинену екосистему бібліотек для аналізу текстових даних. Зокрема, для попередньої обробки тексту можуть застосовуватися бібліотеки NLTK, spaCy або Stanza, які забезпечують токенізацію, лематизацію та очищення тексту від шуму. Для реалізації модуля аналізу тональності доцільно використати сучасні трансформерні моделі (BERT, RoBERTa, XLM-R), доступні через бібліотеку Hugging Face Transformers.

Архітектурно система може складатися з кількох основних модулів: модуля збору даних (через API соціальних мереж або web-scraping), модуля попередньої обробки тексту, модуля аналізу тональності та емоцій, модуля збереження результатів у базі даних, а також аналітичної панелі для візуалізації результатів. Для зберігання даних може бути використана реляційна (PostgreSQL) або нереляційна (MongoDB) база даних. [33]

Особливу увагу доцільно приділити адаптації моделей до української мови шляхом донавчання на спеціалізованих корпусах текстів. Це дозволить підвищити точність визначення емоційного забарвлення повідомлень і мінімізувати помилки, пов'язані з контекстною неоднозначністю.

Таким чином, запропонований підхід передбачає створення гнучкої, масштабованої та адаптованої до конкретних потреб системи, що поєднує сучасні NLP-технології з можливістю подальшого розширення функціоналу. Важливою особливістю аналізу емоційного ставлення аудиторії до бренду є те, що

повідомлення користувачів у соціальних мережах не завжди мають чітко виражену тональність. Часто один і той самий текст може містити як позитивні, так і негативні оцінки. Наприклад, користувач може позитивно оцінювати якість товару, але водночас висловлювати невдоволення швидкістю доставки або роботою служби підтримки. У таких випадках простого визначення загальної тональності повідомлення може бути недостатньо, оскільки для повноцінного аналізу необхідно враховувати окремі аспекти взаємодії користувача з брендом.[2]

У межах предметної області доцільно розрізняти загальний аналіз тональності та аспектний аналіз. Загальний аналіз тональності дозволяє визначити підсумкове емоційне забарвлення повідомлення: позитивне, негативне або нейтральне. Аспектний аналіз, у свою чергу, передбачає визначення ставлення користувача до окремих характеристик бренду, товару або послуги. До таких аспектів можуть належати ціна, якість, доставка, обслуговування, дизайн, зручність користування, швидкість роботи сервісу та інші параметри. Саме аспектний аналіз є особливо корисним для бізнесу, оскільки дозволяє не лише виявити наявність негативу, а й зрозуміти його причину.

Ще одним важливим завданням є визначення джерела інформації та контексту появи повідомлення. Реакція користувачів може суттєво відрізнятися залежно від соціальної платформи, формату публікації та цільової аудиторії. Наприклад, коментарі в соціальних мережах можуть бути більш емоційними та короткими, тоді як відгуки на спеціалізованих платформах зазвичай містять детальніший опис досвіду користувача. Тому під час розроблення інформаційної системи необхідно враховувати не лише текст повідомлення, а й додаткові метадані: дату публікації, джерело, ключові слова, автора або умовний ідентифікатор користувача, а також зв'язок повідомлення з конкретним брендом.

Окрему роль у визначенні емоційного ставлення відіграє попередня обробка текстових даних. Повідомлення у соціальних мережах часто мають неформальний характер і можуть містити граматичні помилки, скорочення, повторення символів,

емодзі, хештеги, згадки інших користувачів та посилання. Без попереднього очищення такі елементи можуть знижувати точність аналізу. Наприклад, посилання або службові символи не несуть прямого емоційного змісту, однак можуть ускладнювати обробку тексту алгоритмами. Водночас емодзі можуть мати важливе емоційне значення, тому під час обробки їх доцільно або зберігати, або перетворювати на відповідні текстові ознаки.

Під час аналізу українськомовного контенту виникають додаткові труднощі, пов'язані з морфологічною складністю мови. Одне й те саме слово може мати багато форм залежно від відмінка, числа, роду або часу. Тому для підвищення точності аналізу доцільно застосовувати нормалізацію тексту, зокрема лематизацію, яка дозволяє привести слова до початкової форми. Це дає змогу зменшити кількість різних варіантів одного й того самого поняття та покращити роботу алгоритмів класифікації.

Також важливим є питання достовірності отриманих даних. Не всі повідомлення в соціальних мережах відображають реальне ставлення користувачів до бренду. Частина контенту може створюватися ботами, конкурентами або користувачами з метою штучного формування позитивного чи негативного інформаційного фону. Через це інформаційна система повинна передбачати можливість фільтрації підозрілих повідомлень, повторів, спаму та нерелевантних згадувань. Такий підхід дозволяє підвищити якість аналітичних результатів і зробити висновки більш обґрунтованими. [14]

Таким чином, аналіз емоційного ставлення аудиторії до бренду є комплексним завданням, яке поєднує збір даних, попередню обробку тексту, визначення тональності, фільтрацію нерелевантного контенту та подання результатів у зручній для користувача формі. Саме тому розроблення інформаційної системи для автоматизованого аналізу повідомлень соціальних мереж є доцільним і практично значущим напрямом дослідження. У наявному документі розділ 1 вже містить аналіз предметної області, порівняння існуючих

рішень, підходи до вирішення задачі та постановку задачі, тому цей фрагмент логічно доповнює його зміст.

1.4 Постановка задачі

Метою роботи є розробка інформаційної системи для автоматичного визначення емоційного ставлення аудиторії до бренду на основі аналізу текстових даних із соціальних мереж.

Для досягнення поставленої мети необхідно вирішити такі завдання:

- 1) дослідити теоретичні основи соціального моніторингу та аналізу тональності;
- 2) провести аналіз предметної області та визначити її особливості;
- 3) здійснити порівняльний аналіз існуючих програмних рішень;
- 4) виявити основні проблеми сучасних систем аналізу емоцій;
- 5) сформулювати функціональні та нефункціональні вимоги до розроблюваної системи;
- 6) розробити архітектурну модель інформаційної системи;
- 7) реалізувати модуль збору та попередньої обробки даних;
- 8) реалізувати модель аналізу тональності та емоцій;
- 9) провести тестування точності та ефективності роботи системи;
- 10) оцінити результати реалізації та сформулювати висновки.

Об'єктом дослідження є процес визначення емоційного ставлення аудиторії до бренду в цифровому середовищі.

Предметом дослідження є методи та алгоритми обробки природної мови, що використовуються для автоматичного аналізу тональності та емоцій у текстових повідомленнях. [3]

У результаті виконання роботи має бути створена працездатна система, здатна здійснювати автоматичний аналіз емоційного забарвлення повідомлень,

формувати аналітичні звіти та підтримувати прийняття управлінських рішень на основі отриманих даних. [8]

1.5 Висновки до першого розділу

У першому розділі було досліджено предметну область, пов'язану з визначенням емоційного ставлення аудиторії до бренду в соціальних мережах. Встановлено, що соціальні мережі є важливим джерелом формування репутації бренду та отримання зворотного зв'язку від споживачів, а великі обсяги користувацького контенту потребують автоматизованого збору, обробки й аналізу.

У ході дослідження визначено основні проблеми сучасних систем аналізу тональності, зокрема складність інтерпретації емоційного змісту повідомлень, наявність сарказму та іронії, мультимедійність контенту, вплив бот-активності, а також недостатню адаптацію багатьох рішень до української мови. Показано, що ефективне вирішення цих проблем потребує використання методів обробки природної мови, машинного навчання та штучного інтелекту.

Порівняльний аналіз існуючих програмних рішень показав, що корпоративні платформи мають високий рівень функціональності, але є дорогими та складними у впровадженні. Доступніші інструменти для малого й середнього бізнесу часто мають обмежені аналітичні можливості, тоді як хмарні та open-source рішення потребують додаткового налаштування й технічної підтримки.

На основі отриманих результатів обґрунтовано доцільність розробки власної інформаційної системи, адаптованої до мовних і функціональних особливостей поставленої задачі. Визначено доцільність використання Python, сучасних NLP-бібліотек і трансформерних моделей для аналізу тональності та емоційного забарвлення тексту. Отримані результати стали основою для подальшого проектування архітектури системи, реалізації програмних модулів і оцінювання ефективності розробленого рішення.

2 ПРОЄКТУВАННЯ ІНФОРМАЦІЙНОЇ СИСТЕМИ ВИЗНАЧЕННЯ ЕМОЦІЙНОГО СТАВЛЕННЯ АУДИТОРІЇ ДО БРЕНДУ

2.1 Визначення програмних підсистем інформаційної системи

Проектування інформаційної системи визначення емоційного ставлення аудиторії до бренду передбачає формування набору взаємопов'язаних програмних підсистем, кожна з яких відповідає за виконання окремих функцій. Поділ системи на підсистеми дозволяє спростити розроблення програмного забезпечення, підвищити модульність архітектури та забезпечити можливість подальшого масштабування системи. [13]

Функціонування інформаційної системи базується на принципі послідовної обробки даних. Повідомлення із соціальних мереж проходять декілька етапів обробки: від отримання даних до формування підсумкових аналітичних звітів. Для реалізації такого підходу систему доцільно поділити на окремі функціональні модулі.

Першою та однією з найважливіших є підсистема збору даних. Її основним призначенням є автоматизоване отримання повідомлень із соціальних мереж, форумів, блогів та інших відкритих інформаційних джерел. Робота підсистеми може базуватися на використанні API соціальних платформ або засобів web-scraping. Основними завданнями підсистеми є пошук повідомлень за ключовими словами, отримання текстового контенту, фільтрація дублікатів та передавання інформації до наступних модулів системи.

Отримані текстові повідомлення надходять до підсистеми попередньої обробки тексту. Її призначення полягає у підготовці тексту до подальшого аналізу тональності. Повідомлення у соціальних мережах часто містять помилки, скорочення, емодзі, спеціальні символи, гіперпосилання та інші елементи, що можуть негативно впливати на результати аналізу. Тому підсистема виконує очищення тексту, токенізацію, нормалізацію слів, видалення стоп-слів та інші операції обробки природної мови. [3, 8]

					КвРІСТ 2401217.22.01.02 ПЗ	Арк. 17
Зм.	Арк.	№ докум.	Підпис	Дата		

Наступним компонентом є підсистема аналізу тональності, яка виконує безпосереднє визначення емоційного забарвлення повідомлень. Для цього можуть використовуватися словникові методи, алгоритми машинного навчання або сучасні трансформерні моделі. Результатом роботи підсистеми є визначення класу тональності повідомлення та числового індексу емоційного забарвлення.

Важливим елементом системи є підсистема збереження даних. Вона відповідає за збереження повідомлень, результатів аналізу, інформації про бренди, користувачів та сформовані звіти. Використання бази даних дозволяє забезпечити централізоване керування інформацією та швидкий доступ до результатів аналізу.

Для представлення результатів користувачу використовується підсистема формування звітів. Вона забезпечує побудову таблиць, діаграм, графіків та інших елементів візуалізації. Завдяки цьому користувач отримує можливість швидко оцінити зміну настроїв аудиторії та виявити негативні інформаційні тенденції.

Окреме значення має підсистема користувацького інтерфейсу. Вона забезпечує взаємодію користувача із системою, надаючи можливість налаштування параметрів аналізу, запуску обробки даних та перегляду результатів.

Для забезпечення стабільної роботи системи також передбачено підсистему адміністрування, яка відповідає за керування користувачами, налаштування параметрів системи, моніторинг роботи модулів та забезпечення інформаційної безпеки.

Всі модулі повинні взаємодіяти між собою, для того, щоб підтримувати коректну роботу ІС. Схематичне зображення цих відносин наведено на рисунку 2.1, як саме відбувається взаємодія між ними і користувачем ІС. А також у таблиці 2.1 наведено основні підсистеми інформаційної системи та їхні функції, передбачені проєктом. Такий поділ дозволяє чітко визначити призначення кожного компонента, встановити його роль у загальному процесі обробки даних і спростити подальше проєктування програмної реалізації системи.

Таблиця 2.1 – Основні функції підсистем

Підсистема	Основне призначення
Збір даних	Отримання повідомлень із соціальних мереж
Попередня обробка	Очищення та нормалізація тексту
Аналіз тональності	Визначення емоційного забарвлення
Збереження даних	Робота з базою даних
Формування звітів	Побудова графіків і статистики
Користувацький інтерфейс	Взаємодія користувача із системою
Адміністрування	Керування системою та безпекою



Рисунок 2.1 – Схематичне зображення взаємодії підсистем

2.2 Проектування архітектури інформаційної системи

Архітектура інформаційної системи є одним із ключових етапів проектування програмного забезпечення, оскільки саме вона визначає структуру системи, принципи взаємодії між окремими компонентами та способи обробки інформації. Від правильно обраної архітектури залежить продуктивність системи, можливість її масштабування, зручність супроводу та подальшого розширення функціоналу.

Для інформаційної системи визначення емоційного ставлення аудиторії до бренду доцільно використовувати багаторівневу клієнт-серверну архітектуру. [10] Такий підхід дозволяє розділити систему на окремі логічні рівні, кожен з яких виконує власні функції та взаємодіє з іншими компонентами через визначені механізми обміну даними. Використання багаторівневої архітектури забезпечує модульність програмного забезпечення, спрощує модернізацію системи та дозволяє незалежно змінювати окремі програмні модулі без порушення роботи всієї системи.

У межах розроблюваної системи доцільно виділити три основні рівні архітектури: рівень представлення, рівень бізнес-логіки та рівень даних.

Рівень представлення забезпечує взаємодію користувача з інформаційною системою. Основним призначенням цього рівня є відображення інформації та отримання команд від користувача. Реалізація рівня представлення може здійснюватися у вигляді веб-інтерфейсу з використанням HTML, CSS та JavaScript[40, 41]. Через інтерфейс користувач отримує можливість виконувати налаштування параметрів аналізу, додавати бренди, запускати процес збору даних, переглядати результати аналізу та формувати звіти.

Важливою перевагою веб-інтерфейсу є його універсальність та доступність. Користувач може працювати із системою незалежно від операційної системи чи типу пристрою. Крім того, використання веб-технологій дозволяє спростити

оновлення програмного забезпечення та забезпечити централізоване адміністрування системи.

Другим рівнем архітектури є рівень бізнес-логіки. Саме цей рівень виконує основні операції обробки інформації та реалізує функціональні можливості системи. До складу рівня бізнес-логіки входять модулі збору даних, попередньої обробки тексту, аналізу тональності, формування статистики та генерації звітів.

Модуль збору даних відповідає за отримання повідомлень із соціальних мереж, форумів та інших онлайн-джерел. Для отримання інформації можуть використовуватися офіційні API платформ або механізми web-scraping. Підсистема повинна забезпечувати регулярне оновлення даних, перевірку коректності отриманих повідомлень та уникнення дублювання інформації.

Після отримання повідомлень інформація передається до модуля попередньої обробки тексту. На цьому етапі виконується очищення тексту від HTML-тегів, спеціальних символів, посилань, службових елементів та зайвих пробілів. Також можуть виконуватися токенізація, нормалізація слів, видалення стоп-слів і лематизація. Попередня обробка тексту дозволяє підвищити точність роботи алгоритмів аналізу тональності та зменшити кількість помилок класифікації.

Одним із центральних компонентів системи є модуль аналізу тональності. Його основним завданням є визначення емоційного забарвлення повідомлень користувачів. Для реалізації аналізу можуть використовуватися як словникові підходи, так і алгоритми машинного навчання. Найбільш перспективним рішенням є використання трансформерних моделей, таких як BERT, RoBERTa або XLM-R, які забезпечують високу точність класифікації тексту та здатні враховувати контекст повідомлення.

У процесі аналізу система визначає, чи є повідомлення позитивним, негативним або нейтральним. Додатково може обчислюватися числовий індекс тональності та рівень впевненості моделі. Це дозволяє формувати більш детальну аналітику та оцінювати якість класифікації.

					КвРІСТ 2401217.22.01.02 ПЗ	Арк. 21
Зм.	Арк.	№ докум.	Підпис	Дата		

Після завершення аналізу результати передаються до модуля формування статистики та звітності. Даний модуль виконує узагальнення результатів аналізу та побудову аналітичних матеріалів. Користувач отримує можливість переглядати кількість позитивних, негативних і нейтральних повідомлень, динаміку зміни емоційного ставлення аудиторії, рівень активності користувачів та інші статистичні показники.

Для підвищення наочності результати можуть подаватися у вигляді графіків, гістограм, кругових діаграм та часових графіків. Візуалізація даних суттєво спрощує сприйняття інформації та дозволяє швидко виявляти негативні тенденції або репутаційні ризики.

Третім рівнем архітектури є рівень даних. Він відповідає за збереження інформації у базі даних та забезпечення доступу до неї. Для реалізації системи доцільно використовувати реляційну систему керування базами даних PostgreSQL або MySQL[33, 49]. Використання реляційної бази даних дозволяє забезпечити цілісність інформації, підтримку складних SQL-запитів та ефективну роботу зі структурованими даними[12].

У базі даних зберігаються відомості про бренди, повідомлення користувачів, джерела даних, результати аналізу тональності, користувачів системи та сформовані звіти. Для підвищення продуктивності можуть використовуватися механізми індексування таблиць та оптимізації запитів.

Взаємодія між рівнями системи здійснюється за допомогою API-запитів. Користувацький інтерфейс передає запити до серверної частини, після чого модулі бізнес-логіки виконують необхідну обробку даних та повертають результати у форматі JSON[34, 39]. Такий підхід дозволяє забезпечити незалежність клієнтської та серверної частини системи, а також спрощує інтеграцію з іншими програмними продуктами.

Загальна схема взаємодії компонентів інформаційної системи наведена на рисунку 2.2.

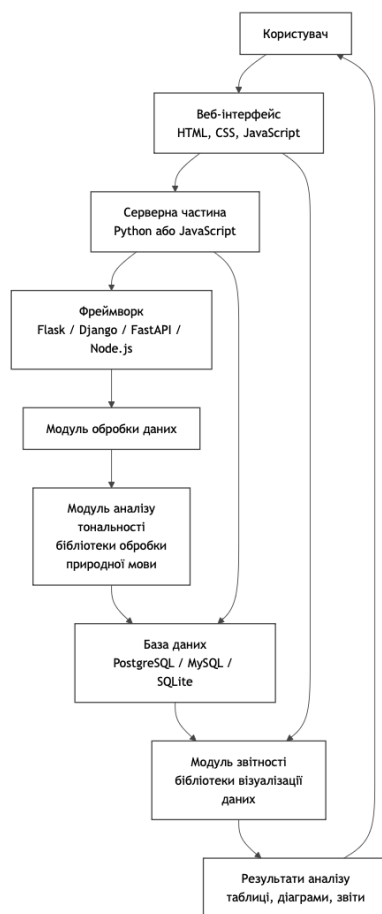


Рисунок 2.2 – Спрощена схема загальної архітектури системи

Для аналізу тональності тексту може бути використаний алгоритм, який враховує частоту позитивних і негативних слів у повідомленні. У простому випадку індекс тональності можна визначити за формулою:

$$S = P - N, \quad (2.1)$$

де S – індекс тональності повідомлення;

P – кількість позитивних ознак у повідомленні;

N – кількість негативних ознак у повідомленні.

Якщо значення S більше нуля, повідомлення можна віднести до позитивних.

Якщо значення S менше нуля, повідомлення має негативне забарвлення. Якщо

значення S дорівнює нулю або знаходиться у визначеному нейтральному діапазоні, повідомлення вважається нейтральним.

Для більш точного аналізу можна використовувати ймовірнісну модель класифікації, яка визначає ймовірність належності повідомлення до кожного класу:

$$C = \operatorname{argmax} P(c|t), \quad (2.2)$$

де C - визначений клас тональності;

c – один із можливих класів тональності;

t – текст повідомлення;

$P(c|t)$ – ймовірність належності тексту до класу c .

У такому випадку повідомлення відноситься до того класу, для якого ймовірність є найбільшою. Такий підхід дозволяє краще враховувати особливості тексту та використовувати навчальні вибірки.

Також під час проектування архітектури інформаційної системи важливо враховувати не лише функціональний поділ на рівні, а й особливості реального використання програмного продукту. Оскільки система призначена для аналізу повідомлень із соціальних мереж, вона повинна підтримувати роботу з даними, які можуть надходити нерівномірно. У певні періоди активність користувачів може бути низькою, однак під час запуску рекламної кампанії, появи резонансної новини або виникнення негативного інформаційного приводу кількість повідомлень може суттєво збільшуватися. Тому архітектура має бути орієнтована на поступове масштабування, щоб за потреби можна було збільшити кількість оброблюваних повідомлень без повної перебудови системи.

Для цього доцільно передбачити можливість незалежного функціонування окремих модулів. Наприклад, модуль збору даних може працювати за розкладом і періодично отримувати нові повідомлення, тоді як дааналізу тональності може запускатися окремо після накопичення певної кількості записів. Такий підхід

дозволяє уникнути перевантаження серверної частини та забезпечити більш стабільну роботу системи. Крім того, окреме виконання модулів спрощує їх тестування, оновлення та заміну у випадку переходу на інші технології.

Важливою особливістю архітектури є організація взаємодії між клієнтською та серверною частинами. Користувацький інтерфейс не повинен безпосередньо взаємодіяти з базою даних, оскільки це створює ризики порушення цілісності інформації та ускладнює контроль доступу. Усі запити користувача мають проходити через серверний рівень, де виконується перевірка прав доступу, валідація введених даних, обробка запиту та формування відповіді. Це дозволяє централізовано контролювати логіку роботи системи та забезпечувати безпечну взаємодію з даними.

Під час проєктування серверної частини доцільно використовувати REST-підхід, за якого кожна функція системи представлена окремим маршрутом або API-запитом. Наприклад, додавання нового бренду, отримання списку повідомлень, запуск аналізу тональності або формування звіту можуть реалізовуватися як окремі запити. Це робить структуру програмного забезпечення зрозумілою та спрощує подальше розширення функціоналу. У разі необхідності до системи можна буде підключити мобільний застосунок або зовнішню аналітичну панель, використовуючи вже наявний API.

Окрему увагу слід приділити обробці помилок. Інформаційна система повинна коректно реагувати на помилки підключення до джерел даних, некоректний формат повідомлень, відсутність відповіді від бази даних або помилки під час аналізу тональності. Для цього необхідно передбачити механізм логування подій. У журналі подій можуть зберігатися відомості про час виконання операції, тип помилки, модуль, у якому вона виникла, та короткий опис проблеми. Це дасть змогу адміністратору швидко виявляти причини збоїв і підтримувати стабільну роботу системи.

Також доцільно передбачити можливість подальшої інтеграції інформаційної системи з іншими сервісами. Наприклад, результати аналізу

					КвРІСТ 2401217.22.01.02 ПЗ	Арк. 25
Зм.	Арк.	№ докум.	Підпис	Дата		

можуть експортуватися у форматах CSV, XLSX або PDF, передаватися до CRM-системи чи використовуватися для побудови зовнішніх аналітичних панелей. Такий підхід підвищує практичну цінність системи, оскільки результати аналізу можуть застосовуватися не лише в межах самої програми, а й у ширшому процесі управління репутацією бренду.

Таким чином, проєктована архітектура повинна забезпечувати не лише виконання основних функцій, а й гнучкість, надійність, безпечність та можливість подальшого розвитку. Саме модульний підхід дозволяє поступово розширювати систему, додавати нові джерела даних, змінювати алгоритми аналізу тональності та вдосконалювати інструменти візуалізації без порушення роботи вже реалізованих компонентів.

2.3 Проєктування бази даних інформаційної системи

База даних є важливим компонентом інформаційної системи, оскільки забезпечує збереження вхідних повідомлень, результатів аналізу, інформації про бренди, джерела даних і користувачів. Вона повинна мати логічну структуру, яка дозволяє швидко отримувати необхідну інформацію та формувати аналітичні звіти. [11]

Для розроблюваної інформаційної системи доцільно передбачити такі основні таблиці:

- 1) Brands;
- 2) Sources;
- 3) Messages;
- 4) SentimentResults;
- 5) Users;
- 6) Reports.

Таблиця Brands призначена для збереження інформації про бренди, які аналізуються системою. Вона може містити такі поля: ідентифікатор бренду, назва бренду, опис, ключові слова, дата додавання.

Таблиця Sources містить інформацію про джерела даних. До її полів належать ідентифікатор джерела, назва соціальної мережі, тип джерела, адреса або параметри доступу, статус активності.

Таблиця Messages використовується для збереження повідомлень користувачів. Вона містить ідентифікатор повідомлення, ідентифікатор бренду, ідентифікатор джерела, текст повідомлення, дату публікації, автора або умовний ідентифікатор користувача, дату завантаження до системи.

Таблиця SentimentResults призначена для збереження результатів аналізу тональності. Вона містить ідентифікатор результату, ідентифікатор повідомлення, визначений клас тональності, числовий індекс тональності, рівень упевненості моделі та дату аналізу.

Таблиця Users містить дані про користувачів інформаційної системи. Вона може включати ідентифікатор користувача, ім'я, логін, пароль у захищеному вигляді, роль користувача та дату створення облікового запису.

Таблиця Reports використовується для збереження сформованих звітів. Вона містить ідентифікатор звіту, ідентифікатор бренду, період аналізу, кількість позитивних, негативних і нейтральних повідомлень, дату створення звіту та посилання на файл звіту.

Між зазначеними таблицями повинні бути встановлені логічні зв'язки, які забезпечують цілісність даних і коректну роботу інформаційної системи. Таблиця Messages пов'язується з таблицею Brands за допомогою поля brand_id, оскільки кожне повідомлення належить до певного бренду. Також таблиця Messages пов'язується з таблицею Sources через поле source_id, що дозволяє визначити джерело отримання кожного повідомлення. Таблиця SentimentResults пов'язується з таблицею Messages через поле message_id, тому результат аналізу закріплюється за конкретним повідомленням. Таблиця Reports пов'язується з

таблицею Brands, оскільки звіти формуються для окремого бренду за вибраний період часу. Така структура зв'язків дозволяє швидко отримувати повідомлення, результати аналізу та статистичні дані для формування звітів.

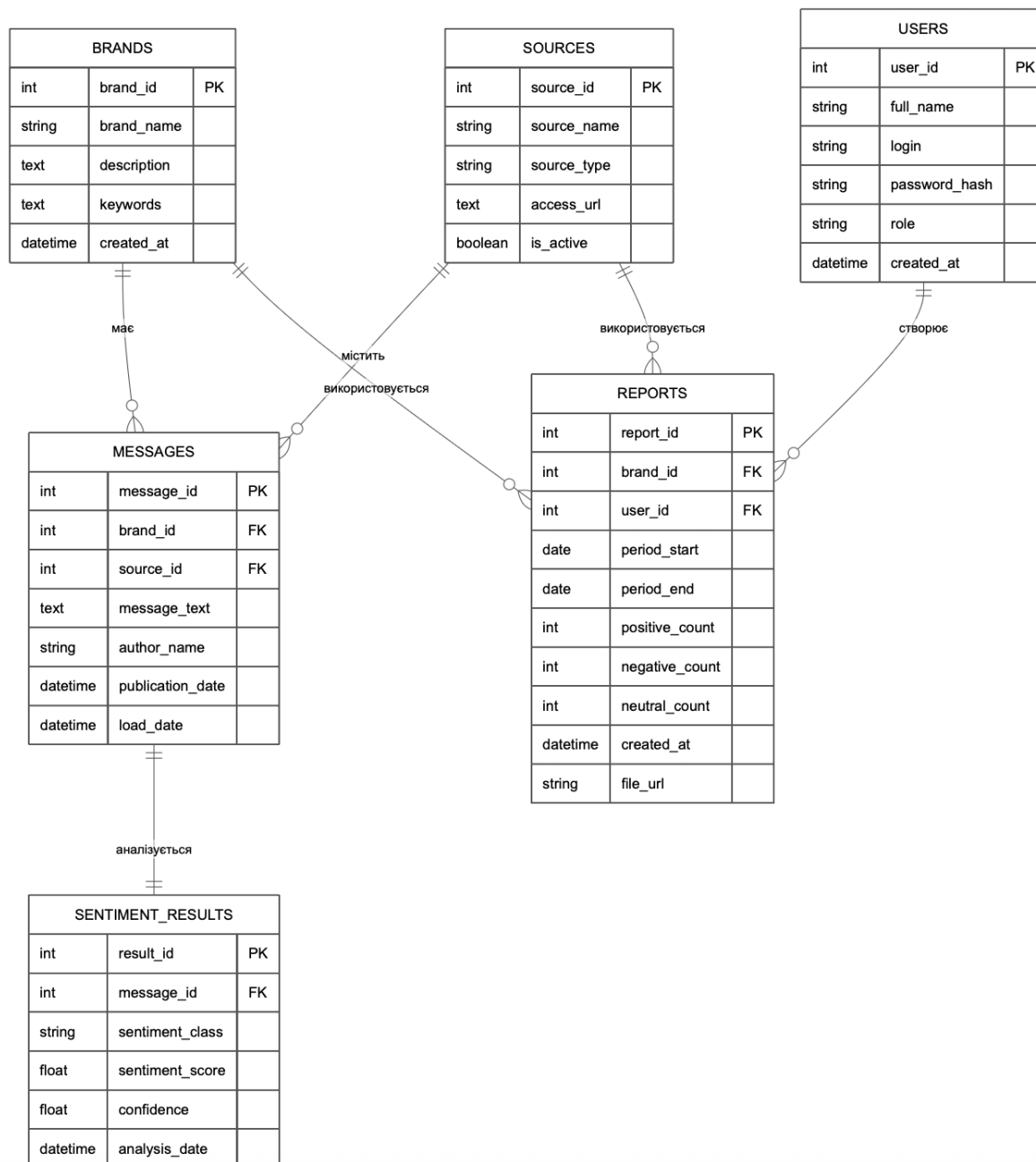


Рисунок 2.3 – Схематичне зображення ER-моделі бази даних

Під час проєктування архітектури інформаційної системи необхідно враховувати особливості роботи з текстовими даними соціальних мереж. Такі дані

можуть надходити з різних джерел, мати різний формат і потребувати попередньої обробки перед аналізом. Тому архітектура системи повинна забезпечувати узгоджену роботу модулів збору, обробки, аналізу, збереження та відображення інформації.

Доцільним є використання модульного підходу, за якого кожен компонент відповідає за окрему функцію. Це спрощує розроблення, тестування та подальше вдосконалення системи. У майбутньому такий підхід дозволить додати нові джерела даних, змінити алгоритм аналізу тональності або розширити засоби формування звітів без повної перебудови програмного забезпечення.

Архітектура системи повинна передбачати поділ на користувацький інтерфейс, серверну частину та базу даних. Користувацький інтерфейс забезпечує введення параметрів аналізу й перегляд результатів. Серверна частина виконує основну логіку роботи системи, а база даних зберігає повідомлення, результати аналізу, інформацію про бренди, джерела та користувачів.

Таким чином, обрана архітектура повинна забезпечувати модульність, надійність і можливість подальшого розвитку інформаційної системи.

2.4 Визначення функціональних та нефункціональних вимог до інформаційної системи

Важливим етапом проектування інформаційної системи є формування функціональних та нефункціональних вимог, які визначають основні можливості системи, особливості її роботи та критерії ефективності[13]. Саме вимоги дозволяють сформулювати цілісне бачення майбутньої системи та забезпечити відповідність програмного продукту поставленим завданням.

Функціональні вимоги визначають перелік операцій, які повинна виконувати система для забезпечення процесу аналізу емоційного ставлення аудиторії до бренду. Оскільки основним джерелом інформації є соціальні мережі, система повинна підтримувати автоматизований збір повідомлень із зовнішніх джерел. Для цього користувач має можливість задавати параметри пошуку: назву

бренду, ключові слова, часові інтервали аналізу та джерела отримання даних. Після отримання повідомлень система повинна виконувати їх попередню обробку, яка включає очищення тексту від зайвих символів, видалення посилань, хештегів, службових конструкцій та нормалізацію текстових даних.

Однією з ключових функцій системи є автоматичне визначення тональності повідомлень. У процесі аналізу система повинна класифікувати повідомлення на позитивні, негативні та нейтральні, а також визначати числові показники емоційного забарвлення. Результати аналізу повинні зберігатися у базі даних та бути доступними для подальшого опрацювання.

Крім аналізу окремих повідомлень, система має забезпечувати формування узагальнених статистичних показників. До таких показників належать кількість позитивних, негативних та нейтральних повідомлень, динаміка зміни настроїв аудиторії у часі, активність користувачів та інтенсивність згадувань бренду. Для зручності сприйняття інформації результати повинні відображатися у вигляді таблиць, графіків та діаграм.

Окрему увагу необхідно приділити підсистемі адміністрування. Адміністратор системи повинен мати можливість керувати джерелами даних, редагувати параметри аналізу, здійснювати контроль користувачів та переглядати історію виконаних операцій. Наявність такої функціональності забезпечує контрольованість та безпечність роботи системи.

Нефункціональні вимоги визначають якісні характеристики програмного продукту. Однією з найважливіших вимог є продуктивність системи. Оскільки соціальні мережі генерують великі обсяги інформації, система повинна забезпечувати швидку обробку текстових повідомлень та своєчасне формування результатів аналізу. Затримки у роботі можуть знизити актуальність отриманих даних, особливо у випадку кризових ситуацій або різких змін громадської думки.

Важливим аспектом є також масштабованість системи. Архітектура програмного забезпечення повинна дозволяти розширення функціоналу, підключення нових джерел даних та збільшення кількості оброблюваних

повідомлень без суттєвого зниження продуктивності. Це особливо актуально у випадку збільшення кількості брендів або переходу до аналізу міжнародних інформаційних потоків [10].

Не менш важливою є надійність системи. Інформаційна система повинна забезпечувати коректне збереження даних навіть у випадку помилок або тимчасової недоступності зовнішніх сервісів. Для цього необхідно передбачити механізми резервного збереження інформації, повторного виконання запитів та обробки помилок.

Окреме значення має безпека інформації. Система повинна забезпечувати захист облікових записів користувачів, шифрування паролів та контроль доступу до адміністративних функцій. Також необхідно враховувати вимоги щодо захисту персональних даних і дотримання політики використання даних соціальних платформ.

Окреме значення має безпека інформації. Система повинна забезпечувати захист облікових записів користувачів, шифрування паролів та контроль доступу до адміністративних функцій. Також необхідно враховувати вимоги щодо захисту персональних даних і дотримання політики використання даних соціальних платформ.

Крім зазначених вимог, під час проєктування інформаційної системи необхідно враховувати зручність подальшого супроводу та можливість розвитку програмного забезпечення. Система не повинна бути жорстко прив'язана до одного джерела даних або одного способу аналізу повідомлень. У майбутньому може виникнути потреба підключити нові соціальні мережі, розширити перелік ключових слів, змінити алгоритм аналізу тональності або додати нові типи звітів. Тому вимоги до системи мають передбачати модульність і можливість поступового розширення функціоналу.

Основні функціональні та нефункціональні вимоги до інформаційної системи наведено в таблиці 2.2.

Таблиця 2.2 – Основні вимоги до інформаційної системи

Тип вимоги	Характеристика
Функціональна	Збір повідомлень із соціальних мереж
Функціональна	Попередня обробка текстових даних
Функціональна	Аналіз тональності повідомлень
Функціональна	Формування статистики та звітів
Функціональна	Візуалізація результатів аналізу
Функціональна	Керування користувачами та брендами
Нефункціональна	Висока швидкість обробки даних
Нефункціональна	Масштабованість архітектури
Нефункціональна	Надійність збереження інформації
Нефункціональна	Захист даних користувачів
Нефункціональна	Зручність користувацького інтерфейсу

Таким чином, сформовані вимоги визначають основні напрями реалізації інформаційної системи та створюють основу для подальшого програмного розроблення окремих модулів.

2.5 Моделювання потоків даних інформаційної системи

Одним із важливих етапів проєктування інформаційної системи є моделювання потоків даних між окремими компонентами системи [13]. Аналіз потоків даних дозволяє визначити послідовність передачі інформації, встановити

взаємозв'язки між програмними модулями та оцінити ефективність організації процесів обробки інформації.

У розроблюваній інформаційній системі основним об'єктом обробки є текстові повідомлення користувачів соціальних мереж. Потік даних формується починаючи з моменту отримання повідомлень із зовнішніх джерел та завершується формуванням аналітичних звітів для користувача [14].

На початковому етапі користувач через веб-інтерфейс задає параметри аналізу. До таких параметрів належать назва бренду, ключові слова, джерела даних та часові межі дослідження. Введені параметри передаються до серверної частини системи, де формується запит до модуля збору даних.

Модуль збору даних взаємодіє із зовнішніми джерелами інформації через API соціальних платформ або засоби web-scraping. Отримані повідомлення проходять первинну перевірку та передаються до модуля попередньої обробки тексту. На цьому етапі виконується очищення повідомлень від службових символів, посилань, HTML-тегів та інших елементів, які можуть негативно впливати на результати аналізу.

Після завершення попередньої обробки текст передається до модуля аналізу тональності. Тут виконується класифікація повідомлень за емоційним забарвленням. Результати аналізу разом із текстом повідомлення та службовою інформацією зберігаються у базі даних.

Наступним етапом є формування статистики та звітності. На основі накопичених даних система виконує узагальнення результатів аналізу та формує статистичні показники. Після цього результати передаються до модуля візуалізації, де відображаються у вигляді таблиць, графіків і діаграм.

Модель потоків даних дозволяє визначити основні інформаційні зв'язки між підсистемами та забезпечити узгоджену роботу всіх компонентів системи.

2.6 Проектування користувацького інтерфейсу системи

Користувацький інтерфейс є важливим компонентом інформаційної системи, оскільки саме через нього користувач взаємодіє з програмним забезпеченням. Від зручності та зрозумілості інтерфейсу залежить ефективність роботи користувача та швидкість отримання необхідної аналітичної інформації.

Під час проектування інтерфейсу необхідно враховувати принципи простоти, логічності та доступності. Інтерфейс системи повинен забезпечувати швидкий доступ до основних функцій, мінімізувати кількість дій користувача та забезпечувати наочне відображення результатів аналізу.

Основним елементом інтерфейсу є головна сторінка системи, на якій користувач може задати параметри аналізу. Для цього передбачаються поля введення назви бренду, ключових слів, часових інтервалів та вибору джерел даних. Після запуску аналізу користувач отримує доступ до результатів обробки повідомлень.

Для відображення результатів доцільно використовувати таблиці, графіки та діаграми. Графічне подання статистики дозволяє швидко оцінити загальний рівень емоційного ставлення аудиторії та виявити негативні тенденції. Наприклад, кругові діаграми можуть використовуватися для відображення співвідношення позитивних, негативних і нейтральних повідомлень, а часові графіки - для аналізу динаміки зміни настроїв аудиторії.

Важливим елементом інтерфейсу є система фільтрації даних. Користувач повинен мати можливість переглядати результати за окремими соціальними мережами, часовими періодами або категоріями повідомлень. Це значно підвищує зручність аналізу великих обсягів інформації.

Інтерфейс адміністратора повинен надавати додаткові можливості керування системою. До таких функцій належать додавання нових джерел даних, керування користувачами, перегляд журналу подій та налаштування параметрів аналізу.

					КвРІСТ 2401217.22.01.02 ПЗ	Арк. 34
Зм.	Арк.	№ докум.	Підпис	Дата		

Під час реалізації користувацького інтерфейсу можуть використовуватися сучасні веб-технології HTML, CSS та JavaScript [40]. Для створення інтерактивних елементів доцільно застосовувати бібліотеки Bootstrap або React, які дозволяють забезпечити адаптивність інтерфейсу та коректне відображення системи на різних пристроях [40].

Таким чином, правильно спроектований користувацький інтерфейс забезпечує ефективну взаємодію користувача із системою та спрощує процес аналізу емоційного ставлення аудиторії до бренду.

2.7 Проектування системи безпеки та захисту даних

Проектування системи безпеки є важливою складовою розроблення інформаційної системи визначення емоційного ставлення аудиторії до бренду. Оскільки система працює з повідомленнями користувачів, результатами аналітичної обробки, обліковими записами та службовими параметрами, необхідно забезпечити захист даних на всіх етапах їх обробки. Безпека системи повинна охоплювати не лише збереження інформації в базі даних, а й контроль доступу, перевірку вхідних даних, захист передавання інформації та відновлення даних у разі збоїв.

Одним із основних елементів захисту є автентифікація користувачів. Доступ до системи має надаватися лише після введення коректного логіна та пароля. Це дозволяє ідентифікувати користувача та визначити, які функції системи йому доступні. Після автентифікації виконується авторизація, тобто перевірка прав доступу до окремих модулів. У системі доцільно передбачити щонайменше дві ролі: звичайний користувач і адміністратор. Звичайний користувач може переглядати результати аналізу, запускати обробку повідомлень і формувати звіти, тоді як адміністратор має додаткові права на керування користувачами, джерелами даних і параметрами системи.

Особливу увагу необхідно приділити захисту паролів. Паролі користувачів не повинні зберігатися у відкритому вигляді. Для цього застосовуються алгоритми хешування, які перетворюють пароль у спеціальний набір символів. У разі несанкціонованого доступу до бази даних це ускладнює отримання справжніх паролів користувачів. Додатково може використовуватися додавання випадкового значення до пароля перед хешуванням, що підвищує рівень захисту облікових записів.

Важливим напрямом захисту є безпечне передавання даних між клієнтською та серверною частинами системи. Для цього доцільно використовувати протокол HTTPS, який забезпечує шифрування інформації під час передавання мережею. Це дозволяє знизити ризик перехоплення логінів, паролів, параметрів аналізу та інших службових даних [34, 39].

Окреме значення має перевірка вхідних даних. Усі дані, які користувач вводить через інтерфейс, повинні проходити валідацію. Це стосується назв брендів, ключових слів, дат, параметрів фільтрації та завантажених файлів. Перевірка даних дозволяє запобігти некоректній роботі системи, помилкам у базі даних та потенційним атакам, пов'язаним із передаванням шкідливих або неправильно сформованих запитів.

Для підвищення контрольованості роботи системи доцільно передбачити журналювання основних подій. У журналі можуть фіксуватися входи користувачів до системи, запуск аналізу, створення звітів, зміна налаштувань, помилки обробки даних та інші важливі дії. Наявність журналу подій дає змогу адміністратору виявляти проблеми в роботі системи, аналізувати причини помилок і контролювати дії користувачів.

Не менш важливим є резервне копіювання бази даних. Оскільки в системі зберігаються повідомлення, результати аналізу та сформовані звіти, втрата цих даних може призвести до порушення роботи інформаційної системи. Регулярне створення резервних копій дозволяє відновити інформацію у випадку технічного збою, помилки користувача або пошкодження бази даних.

Таким чином, система безпеки повинна включати автентифікацію та авторизацію користувачів, захист паролів, шифрування передавання даних, перевірку вхідної інформації, журналювання подій і резервне копіювання. Реалізація цих заходів забезпечує стабільну роботу інформаційної системи, збереження даних і захист користувацької інформації від несанкціонованого доступу.

2.8 Висновки до другого розділу

У другому розділі було виконано проектування інформаційної системи визначення емоційного ставлення аудиторії до бренду на основі аналізу даних соціальних мереж. У процесі проектування визначено основні програмні підсистеми, їх функціональне призначення та принципи взаємодії між окремими компонентами системи. Встановлено, що для ефективної реалізації поставлених завдань система повинна містити підсистеми збору даних, попередньої обробки тексту, аналізу тональності, збереження інформації, формування звітів, користувацького інтерфейсу та адміністрування.

У межах розділу було спроектовано архітектуру інформаційної системи на основі багаторівневого клієнт-серверного підходу. Використання такого підходу дозволяє забезпечити модульність програмного забезпечення, розмежування функцій між окремими рівнями системи та можливість подальшого масштабування. Виділено рівень представлення, рівень бізнес-логіки та рівень даних, кожен із яких виконує окремі функції в процесі обробки інформації. [10]

Під час проектування архітектури було обґрунтовано використання сучасних веб-технологій та мови програмування Python, що забезпечує гнучкість розроблення, підтримку засобів обробки природної мови та інтеграцію з бібліотеками машинного навчання. Також розглянуто особливості використання REST API для взаємодії між клієнтською та серверною частинами системи.

У роботі було спроектовано структуру бази даних інформаційної системи. Визначено основні таблиці для збереження інформації про бренди, повідомлення,

джерела даних, результати аналізу тональності, користувачів та сформовані звіти. Описано логічні зв'язки між таблицями та принципи організації збереження інформації. Запропонована структура бази даних забезпечує цілісність інформації, зручність доступу до даних та можливість формування статистичних звітів за різними параметрами.

Крім того, у розділі визначено функціональні та нефункціональні вимоги до системи. Серед основних функціональних можливостей визначено автоматизований збір повідомлень із соціальних мереж, попередню обробку тексту, аналіз тональності, формування статистики та візуалізацію результатів. Нефункціональні вимоги охоплюють питання продуктивності, масштабованості, безпеки та надійності програмного забезпечення.

Також було розглянуто моделювання потоків даних між окремими компонентами системи. Визначено послідовність передачі інформації від моменту отримання повідомлень із зовнішніх джерел до формування аналітичних звітів для користувача. Це дозволило сформувавши цілісне уявлення про процес функціонування інформаційної системи та забезпечити узгоджену взаємодію між її підсистемами.

Окрему увагу приділено проектуванню користувацького інтерфейсу та системи захисту даних. Визначено основні принципи побудови інтерфейсу, спрямовані на забезпечення зручності роботи користувача та ефективного відображення результатів аналізу. Розглянуто механізми автентифікації, авторизації, шифрування паролів та резервного копіювання інформації, що дозволяє забезпечити безпечну роботу системи.

Таким чином, у другому розділі було сформовано повну концептуальну модель інформаційної системи визначення емоційного ставлення аудиторії до бренду. Отримані результати створюють основу для подальшої програмної реалізації системи, розроблення окремих модулів програмного забезпечення та проведення тестування ефективності роботи створеного програмного продукту.

3 ПРОГРАМНА РЕАЛІЗАЦІЯ ІНФОРМАЦІЙНОЇ СИСТЕМИ ВИЗНАЧЕННЯ ЕМОЦІЙНОГО СТАВЛЕННЯ АУДИТОРІЇ ДО БРЕНДУ

3.1 Опис реалізації модулів програмного забезпечення інформаційної системи

Програмна реалізація інформаційної системи визначення емоційного ставлення аудиторії до бренду передбачає створення комплексу взаємопов'язаних модулів, які забезпечують повний цикл роботи з даними соціальних мереж. Система повинна виконувати збір повідомлень, їх попередню обробку, визначення тональності, збереження результатів у базі даних, формування статистики та подання результатів користувачеві у зручному вигляді.

З урахуванням вимог, визначених у попередніх розділах, програмне забезпечення інформаційної системи доцільно реалізувати на основі багаторівневої архітектури. Основними рівнями системи є рівень користувацького інтерфейсу, серверний рівень, рівень обробки даних, рівень аналізу тональності та рівень збереження даних.

Рівень користувацького інтерфейсу забезпечує взаємодію користувача із системою. Через веб-інтерфейс користувач може додавати бренд для моніторингу, завантажувати повідомлення, запускати аналіз, переглядати результати, застосовувати фільтри та формувати звіти. Для створення інтерфейсу можуть бути використані HTML, CSS та JavaScript.

Серверний рівень відповідає за обробку запитів користувача, виконання логіки роботи системи та взаємодію з базою даних. Для реалізації серверної частини можуть використовуватися мова програмування Python або JavaScript, а також фреймворки Flask, Django, FastAPI або Node.js [10]. Серверна частина приймає запити від веб-інтерфейсу, передає дані до відповідних модулів і повертає користувачеві результати обробки.

Рівень обробки даних забезпечує очищення та підготовку текстових повідомлень до аналізу [3, 8]. На цьому етапі виконується видалення посилань,

зайвих символів, службової інформації, повторів, непотрібних пробілів та інших шумових елементів. Також текст може приводитися до нижнього регістру, розбиватися на окремі слова, очищатися від стоп-слів і нормалізуватися.

Рівень аналізу тональності призначений для визначення емоційного забарвлення повідомлень. У межах системи кожне повідомлення класифікується як позитивне, негативне або нейтральне. Для цього може використовуватися словниковий підхід, алгоритми машинного навчання або комбінована модель, яка враховує наявність позитивних і негативних ознак у тексті.

Рівень збереження даних забезпечує роботу з базою даних. У базі даних зберігаються відомості про бренди, джерела повідомлень, отримані повідомлення, результати аналізу тональності, користувачів системи та сформовані звіти. Для реалізації цього рівня можуть використовуватися системи керування базами даних PostgreSQL, MySQL або SQLite.

Як було наведено вище, основними програмними модулями інформаційної системи є:

- 1) модуль користувацького інтерфейсу;
- 2) модуль адміністрування;
- 3) модуль збору даних;
- 4) модуль попередньої обробки тексту;
- 5) модуль аналізу тональності;
- 6) модуль збереження даних;
- 7) модуль формування звітів.

3.2 Опис середовища розроблення та структури програмного проєкту

Для реалізації інформаційної системи визначення емоційного ставлення аудиторії до бренду було обрано середовище розроблення, яке дозволяє створити веб-застосунок із можливістю роботи з базою даних, обробки текстових повідомлень та відображення результатів аналізу користувачу. Програмна

					КвРІСТ 2401217.22.01.02 ПЗ	Арк. 40
Зм.	Арк.	№ докум.	Підпис	Дата		

частина системи може бути реалізована мовою Python, оскільки вона є зручною для роботи з текстовими даними, реалізації алгоритмів аналізу тональності та взаємодії з базами даних [9].

Розроблення програмного забезпечення може виконуватися у середовищі Visual Studio Code або PyCharm. Ці середовища забезпечують зручне редагування програмного коду, підсвічування синтаксису, роботу з файлами проєкту та запуск окремих частин програми. Для перевірки роботи бази даних можуть використовуватися вбудовані інструменти SQLite або зовнішні засоби адміністрування баз даних, наприклад DB Browser for SQLite, pgAdmin або phpMyAdmin залежно від обраної системи керування базами даних.

На етапі навчальної реалізації доцільно використовувати локальний варіант запуску системи. У такому випадку серверна частина, база даних і користувацький інтерфейс працюють на одному комп'ютері. Це спрощує процес розроблення, тестування та налагодження програми. У разі подальшого практичного впровадження система може бути перенесена на сервер, що дозволить забезпечити доступ до неї з різних пристроїв через веб-браузер.

Структура програмного проєкту повинна бути зрозумілою та розділеною на окремі частини відповідно до їх призначення. Це дозволяє швидко знаходити потрібні файли, змінювати окремі компоненти системи та уникати змішування програмної логіки, шаблонів інтерфейсу й допоміжних файлів. Орієнтовна структура програмного проєкту може мати такий вигляд:

- 1) app.py - головний файл запуску інформаційної системи;
- 2) database.py - файл для підключення до бази даних і виконання SQL-запитів;
- 3) preprocessing.py - файл із функціями очищення та підготовки тексту;
- 4) sentiment.py - файл із реалізацією алгоритму визначення тональності;
- 5) reports.py - файл для формування статистики та звітів;
- 6) config.py - файл із налаштуваннями підключення до бази даних та параметрами системи;

- 7) templates - каталог HTML-сторінок користувацького інтерфейсу;
- 8) static - каталог CSS-стилів, JavaScript-файлів і зображень;
- 9) data - каталог для тестових наборів повідомлень;
- 10) logs - каталог для збереження журналів подій і помилок.

Файл `app.py` є центральним елементом програмного проєкту. Він відповідає за запуск веб-застосунку, приймання запитів від користувача та виклик відповідних функцій інших файлів. Наприклад, якщо користувач додає новий бренд, запит надходить до `app.py`, після чого викликається функція запису даних у базу. Якщо користувач запускає аналіз повідомлень, `app.py` передає текстові дані до функцій попередньої обробки та аналізу тональності.

Файл `database.py` призначений для роботи з базою даних. У ньому доцільно розмістити функції підключення до бази, створення таблиць, додавання нових записів, вибірки повідомлень, збереження результатів аналізу та отримання статистичних даних. Винесення SQL-запитів в окремий файл робить структуру програми зрозумілішою та спрощує подальше редагування бази даних.

Файл `preprocessing.py` містить функції попередньої обробки тексту. До таких функцій належать видалення посилань, зайвих пробілів, спеціальних символів, хештегів, згадок користувачів, а також приведення тексту до нижнього регістру. Саме цей файл відповідає за підготовку повідомлень до подальшого аналізу. Завдяки цьому основний алгоритм аналізу тональності отримує вже очищений і стандартизований текст.

Файл `sentiment.py` містить реалізацію алгоритму визначення емоційного ставлення аудиторії до бренду. У ньому розміщуються словники позитивних і негативних ознак, функція підрахунку кількості позитивних та негативних слів, обчислення індексу тональності та визначення підсумкового класу повідомлення. Результатом роботи цього файлу є значення тональності: позитивне, негативне або нейтральне.

Файл `reports.py` використовується для формування підсумкової статистики. Він може містити функції підрахунку кількості повідомлень кожного класу,

					КВРІСТ 2401217.22.01.02 ПЗ	Арк. 42
Зм.	Арк.	№ докум.	Підпис	Дата		

формування зведених таблиць за брендом і часовим періодом, а також підготовки даних для побудови діаграм. Завдяки цьому логіка формування звітів відокремлюється від логіки аналізу повідомлень.

Каталог `templates` містить HTML-шаблони сторінок інформаційної системи. До нього можуть входити сторінка авторизації, головна сторінка, сторінка керування брендами, сторінка завантаження повідомлень, сторінка перегляду результатів і сторінка формування звітів. Каталог `static` використовується для збереження CSS-файлів, JavaScript-файлів і графічних елементів, які забезпечують оформлення та інтерактивність інтерфейсу.

Каталог `data` може використовуватися для збереження тестових повідомлень, які застосовуються під час перевірки роботи алгоритму. Каталог `logs` призначений для журналювання подій, наприклад помилок підключення до бази даних, некоректного формату вхідних даних або помилок під час аналізу повідомлень. Наявність журналу подій спрощує пошук і виправлення помилок у процесі тестування.

Загальна послідовність роботи програмного проєкту є такою: користувач через веб-інтерфейс виконує певну дію, запит передається до головного файлу `app.py`, після чого викликаються відповідні функції з інших файлів. Якщо потрібно отримати або зберегти дані, використовується `database.py`. Якщо потрібно підготувати текст до аналізу, використовується `preprocessing.py`. Якщо потрібно визначити тональність повідомлення, використовується `sentiment.py`. Якщо потрібно сформуванати статистику або звіт, використовується `reports.py`.

Таким чином, описана структура програмного проєкту забезпечує логічний поділ системи на окремі файли та каталоги відповідно до їхнього функціонального призначення. Це спрощує розроблення, тестування і супровід інформаційної системи, оскільки кожен компонент відповідає за окрему частину роботи програмного забезпечення. Крім того, така організація проєкту дозволяє надалі розширювати функціональність системи, додавати нові модулі або змінювати окремі частини програмного коду без повної перебудови всієї системи.

3.3 Опис процесу створення баз даних

Спочатку необхідно побудувати логічно-фізичну модель бази даних інформаційної системи. Для цього визначаються основні сутності, які будуть використовуватися під час роботи системи, їхні атрибути та зв'язки між ними. У межах розроблюваної інформаційної системи основними сутностями є бренд, джерело даних, повідомлення, результат аналізу тональності, користувач і звіт.

ER модель бази даних інформаційної системи наведена вище на рисунку 2.3.

Для створення схеми бази даних передбачено такі таблиці:

- 1) Brands;
- 2) Sources;
- 3) Messages;
- 4) SentimentResults;
- 5) Users;
- 6) Reports.

Приклад створення таблиці Brands наведено на рисунку 3.1.

```
CREATE TABLE Brands (  
    brand_id INTEGER PRIMARY KEY,  
    brand_name VARCHAR(100) NOT NULL,  
    description TEXT,  
    keywords TEXT,  
    created_at DATETIME NOT NULL  
);
```

Рисунок 3.1 – Приклад створення таблиці Brands

Приклад створення таблиці Sources наведено на рисунку 3.2.

```

CREATE TABLE Sources (
    source_id INTEGER PRIMARY KEY,
    source_name VARCHAR(100) NOT NULL,
    source_type VARCHAR(50),
    access_url TEXT,
    is_active BOOLEAN DEFAULT TRUE
);

```

Рисунок 3.2 – Приклад створення таблиці Sources

Приклад створення таблиці Messages наведено на рисунку 3.3.

```

CREATE TABLE Messages (
    message_id INTEGER PRIMARY KEY,
    brand_id INTEGER NOT NULL,
    source_id INTEGER NOT NULL,
    message_text TEXT NOT NULL,
    author_name VARCHAR(100),
    publication_date DATETIME,
    load_date DATETIME NOT NULL,
    FOREIGN KEY (brand_id) REFERENCES Brands(brand_id),
    FOREIGN KEY (source_id) REFERENCES Sources(source_id)
);

```

Рисунок 3.3 – Приклад створення таблиці Messages

Приклад створення таблиці SentimentResults наведено на рисунку 3.4.

```

CREATE TABLE SentimentResults (
    result_id INTEGER PRIMARY KEY,
    message_id INTEGER NOT NULL,
    sentiment_class VARCHAR(20) NOT NULL,
    sentiment_score FLOAT,
    confidence FLOAT,
    analysis_date DATETIME NOT NULL,
    FOREIGN KEY (message_id) REFERENCES Messages(message_id)
);

```

Рисунок 3.4 - Приклад створення таблиці SentimentResults

Приклад створення таблиці Users наведено на рисунку 3.5.

```
CREATE TABLE Users (  
    user_id INTEGER PRIMARY KEY,  
    full_name VARCHAR(100) NOT NULL,  
    login VARCHAR(50) NOT NULL UNIQUE,  
    password_hash VARCHAR(255) NOT NULL,  
    role VARCHAR(30) NOT NULL,  
    created_at DATETIME NOT NULL  
);
```

Рисунок 3.5 – Приклад створення таблиці Users

Приклад створення таблиці Reports наведено на рисунку 3.6.

```
CREATE TABLE Reports (  
    report_id INTEGER PRIMARY KEY,  
    brand_id INTEGER NOT NULL,  
    user_id INTEGER NOT NULL,  
    period_start DATE NOT NULL,  
    period_end DATE NOT NULL,  
    positive_count INTEGER DEFAULT 0,  
    negative_count INTEGER DEFAULT 0,  
    neutral_count INTEGER DEFAULT 0,  
    created_at DATETIME NOT NULL,  
    file_url VARCHAR(255),  
    FOREIGN KEY (brand_id) REFERENCES Brands(brand_id),  
    FOREIGN KEY (user_id) REFERENCES Users(user_id)  
);
```

Рисунок 3.6 – Приклад створення таблиці Reports

Зв'язки між таблицями бази даних мають такий вигляд:

- 1) один бренд може мати багато повідомлень;
- 2) одне джерело може містити багато повідомлень;

- 3) одне повідомлення може мати один або декілька результатів аналізу;
- 4) один бренд може мати багато звітів;
- 5) один користувач може створювати багато звітів.

Після створення таблиць виконується перевірка правильності структури бази даних. Для цього додаються тестові записи про бренд, джерело, повідомлення та результат аналізу. Потім виконується вибірка повідомлень за брендом і перевіряється правильність зв'язків між таблицями.

Для підвищення швидкості пошуку інформації доцільно створити індекси для полів, які найчастіше використовуються у запитах. До таких полів належать `brand_id`, `source_id`, `publication_date` і `sentiment_class`.

Приклад створення індексів:

```
CREATE INDEX idx_messages_brand_id ON Messages(brand_id);
CREATE INDEX idx_messages_source_id ON Messages(source_id);
CREATE INDEX idx_messages_publication_date ON
Messages(publication_date);
CREATE INDEX idx_sentiment_class ON SentimentResults(sentiment_class);
```

Створена база даних забезпечує збереження всієї необхідної інформації для функціонування системи. Вона дозволяє зберігати первинні повідомлення, результати їх аналізу, інформацію про бренди та джерела, а також формувати підсумкові звіти за обраний період.

3.4 Опис реалізації алгоритму визначення емоційного ставлення аудиторії до бренду

Алгоритм визначення емоційного ставлення аудиторії до бренду реалізується як послідовність операцій над текстовими повідомленнями користувачів соціальних мереж [6]. Основним завданням алгоритму є визначення тональності кожного повідомлення та присвоєння йому одного з трьох класів: позитивне, негативне або нейтральне.

На вхід алгоритму надходить текст повідомлення з таблиці Messages. Спочатку система виконує очищення тексту. Після цього текст розбивається на окремі слова, які порівнюються зі словниками позитивних і негативних ознак. На основі знайдених ознак обчислюється індекс тональності.

Загальна послідовність роботи алгоритму представлена на Рисунку 3.7:

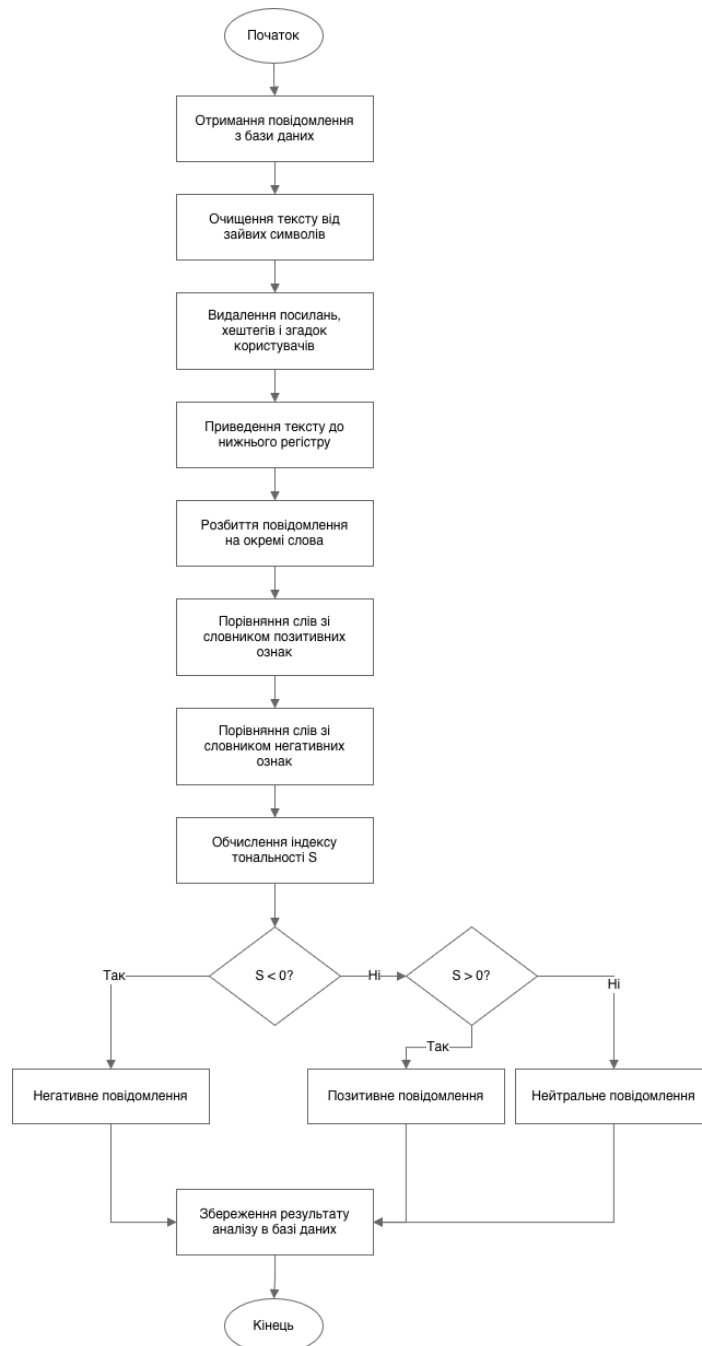


Рисунок 3.7 – Послідовність роботи алгоритму

Для базової реалізації аналізу тональності використовується формула 2.1.

Приклад словників, які можуть використовуватися у системі:

```
positive_words = [  
    "добре", "чудово", "якісний", "зручно", "швидко",  
    "подобається", "рекомендую", "задоволений", "супер",  
    "приємно", "дякую", "класний"  
]  
  
negative_words = [  
    "погано", "жахливо", "незручно", "дорого", "затримка",  
    "проблема", "розчарований", "неякісний", "скарга",  
    "повільно", "гірше", "не рекомендую"  
]
```

Рисунок 3.8 – Приклад словника

Приклад програмної реалізації функції очищення тексту:

```
import re  
  
def clean_text(text):  
    text = text.lower()  
    text = re.sub(r"http\S+|www\S+", " ", text)  
    text = re.sub(r"@w+", " ", text)  
    text = text.replace("#", "")  
    text = re.sub(r"^[а-щьюягґііа-z0-9\s]", " ", text)  
    text = re.sub(r"\s+", " ", text).strip()  
    return text
```

Рисунок 3.9 – Приклад функції очищення тексту

У системі також можна враховувати прості заперечення. Наприклад, фраза «не подобається» не повинна визначатися як позитивна лише через наявність слова «подобається». Для цього перевіряється попереднє слово перед емоційною

ознакою. Якщо перед позитивним словом стоїть «не», така ознака враховується як негативна.

Результати аналізу кожного повідомлення після проходження алгоритму зберігаються у таблиці `SentimentResults`. Для цього до бази даних записуються ідентифікатор повідомлення, клас тональності, числовий індекс, рівень упевненості та дата аналізу.

```
INSERT INTO SentimentResults (  
    message_id,  
    sentiment_class,  
    sentiment_score,  
    confidence_level,  
    analysis_date  
)  
VALUES  
    (1, 'Позитивне', 2, 0.92, CURRENT_TIMESTAMP),  
    (2, 'Негативне', -2, 0.89, CURRENT_TIMESTAMP),  
    (3, 'Нейтральне', 0, 0.76, CURRENT_TIMESTAMP),  
    (4, 'Негативне', -1, 0.84, CURRENT_TIMESTAMP),  
    (5, 'Позитивне', 3, 0.95, CURRENT_TIMESTAMP);
```

Рисунок 3.10 – Приклад SQL запиту вставки результатів в `SentimentResults`

Для перевірки роботи алгоритму доцільно використати набір тестових повідомлень із різним емоційним забарвленням. До такого набору варто включити позитивні, негативні та нейтральні повідомлення, а також приклади із запереченнями. Це дозволяє оцінити, чи правильно система визначає тональність тексту в типових ситуаціях. [18]

У процесі аналізу для кожного повідомлення виконується попереднє очищення тексту, після чого система підраховує кількість позитивних і негативних ознак. На основі цих значень обчислюється індекс тональності. Якщо індекс має додатне значення, повідомлення вважається позитивним. Якщо індекс

має від'ємне значення, повідомлення класифікується як негативне. Якщо індекс дорівнює нулю, повідомлення визначається як нейтральне.

Такий підхід дозволяє наочно показати принцип роботи алгоритму та перевірити правильність класифікації повідомлень. У таблиці наведено приклади текстів, значення індексу тональності та результат, який отримує система після виконання аналізу. Приклад результатів роботи алгоритму наведено в таблиці 3.1.

Таблиця 3.1 – Результати визначення тональності повідомлень

№	Текст повідомлення	Індекс тональності	Результат аналізу
1	Дуже швидка доставка, задоволений сервісом	2	Позитивне
2	Погане обслуговування, замовлення затримали	-2	Негативне
3	Компанія відкрила новий магазин	0	Нейтральне
4	Не подобається якість товару	-1	Негативне
5	Зручно, швидко, рекомендую	3	Позитивне

Таким чином, реалізований алгоритм дозволяє виконувати базову класифікацію повідомлень за емоційним забарвленням. Його перевагою є простота реалізації та зрозумілість принципу роботи. Недоліком є залежність точності від повноти словників і складність обробки сарказму, жаргону та складних мовних конструкцій.

3.5 Опис реалізації користувацького інтерфейсу

Користувацький інтерфейс інформаційної системи забезпечує доступ до основних функцій програмного забезпечення. Його структура розроблена таким чином, щоб користувач міг виконувати аналіз даних соціальних мереж без прямої роботи з базою даних або програмним кодом.

Основними сторінками інтерфейсу є:

- 1) сторінка авторизації;
- 2) головна сторінка;
- 3) сторінка керування брендами;
- 4) сторінка завантаження повідомлень;
- 5) сторінка перегляду повідомлень;
- 6) сторінка результатів аналізу;
- 7) сторінка формування звітів;
- 8) сторінка адміністрування.

Сторінка авторизації призначена для входу користувача до системи. Користувач вводить логін і пароль, після чого система перевіряє наявність відповідного облікового запису в таблиці Users. Якщо введені дані правильні, користувач отримує доступ до функцій системи відповідно до своєї ролі.

Головна сторінка містить коротке зведення про результати аналізу. На ній відображається загальна кількість повідомлень, кількість позитивних, негативних і нейтральних згадок, а також загальний індекс ставлення аудиторії до бренду.

Сторінка керування брендами дозволяє додавати нові бренди та редагувати наявні записи. Для кожного бренду користувач вказує назву, опис і ключові слова. Ці дані використовуються під час пошуку та класифікації повідомлень.

Сторінка завантаження повідомлень використовується для імпорту даних у систему. Користувач обирає файл із повідомленнями, після чого система перевіряє його структуру та завантажує коректні записи до бази даних. Після

завершення імпорту користувач отримує інформацію про кількість доданих і пропущених повідомлень.

Сторінка перегляду повідомлень відображає список усіх завантажених повідомлень. Для кожного запису виводяться текст повідомлення, бренд, джерело, дата публікації, автор і результат аналізу. Користувач може виконувати фільтрацію за брендом, датою, джерелом і тональністю.

Сторінка результатів аналізу містить таблиці та діаграми. Кругова діаграма може використовуватися для відображення співвідношення позитивних, негативних і нейтральних повідомлень. Стовпчикова діаграма дозволяє порівнювати кількість повідомлень за джерелами, а лінійний графік - переглядати зміну тональності в часі.

Сторінка формування звітів дозволяє створити аналітичний звіт за обраний період. Користувач задає бренд, початкову дату та кінцеву дату. Після цього система підраховує кількість повідомлень кожного типу та формує підсумкову таблицю. За потреби звіт може бути експортований у файл.

Сторінка адміністрування доступна лише користувачам із відповідними правами. На цій сторінці адміністратор може керувати користувачами, змінювати їх ролі, переглядати джерела даних і налаштовувати параметри системи.

Під час реалізації користувацького інтерфейсу важливо забезпечити послідовність виконання основних дій користувача. Після входу до системи користувач повинен мати можливість швидко перейти до потрібного розділу та виконати необхідну операцію без додаткового пошуку функцій. Для цього всі основні можливості системи доцільно згрупувати в головному меню.

Головне меню повинно відображати основні етапи роботи з інформаційною системою: перегляд загальної інформації, керування брендами, роботу з повідомленнями, запуск аналізу, перегляд результатів і формування звітів. Така структура дозволяє зробити інтерфейс зрозумілим навіть для користувача, який не має спеціальних технічних знань.

Окрему увагу потрібно приділити зручності перегляду результатів аналізу. Користувач повинен мати можливість не лише побачити загальну статистику, а й перейти до окремих повідомлень, переглянути їхню тональність, джерело та дату публікації. Це дозволяє детальніше аналізувати реакцію аудиторії та швидко знаходити повідомлення з негативним або позитивним емоційним забарвленням.

Приклад структури головного меню інформаційної системи наведено в таблиці 3.2.

Таблиця 3.2 - Структура головного меню інформаційної системи

№	Розділ	Опис
1	Головна	Перегляд загальних показників системи
2	Бренди	Додавання та редагування брендів
3	Повідомлення	Перегляд і фільтрація отриманих повідомлень
4	Аналіз	Запуск аналізу та перегляд результатів
5	Звіти	Формування й експорт аналітичних звітів
6	Адміністрування	Керування користувачами та налаштуваннями системи

Для зручності роботи інтерфейс повинен містити повідомлення про успішне виконання операцій і повідомлення про помилки. Наприклад, після завантаження файлу система повинна повідомити, скільки повідомлень було додано, а скільки пропущено через помилки або дублювання.

Таким чином, користувацький інтерфейс забезпечує повний доступ до функцій інформаційної системи та дозволяє користувачу виконувати моніторинг емоційного ставлення аудиторії до бренду в зручній формі.

3.6 Тестування роботи інформаційної системи

Тестування інформаційної системи визначення емоційного ставлення аудиторії до бренду є важливим етапом програмної реалізації, оскільки воно дозволяє перевірити коректність роботи основних модулів, відповідність системи визначеним вимогам і правильність обробки вхідних даних. Тестування повинно охоплювати всі ключові функції системи: додавання брендів, завантаження повідомлень, попередню обробку тексту, визначення тональності, збереження результатів у базі даних, формування статистики та відображення результатів у користувацькому інтерфейсі.

Основною метою тестування є виявлення можливих помилок у роботі програмного забезпечення до його практичного використання. Для цього необхідно перевірити, чи правильно взаємодіють між собою окремі модулі системи, чи коректно зберігаються дані в базі, чи правильно працює алгоритм аналізу тональності та чи зручно користувачеві виконувати основні операції через інтерфейс.

Тестування доцільно проводити поетапно. На першому етапі перевіряється робота модуля керування брендами. Для цього необхідно виконати додавання нового бренду, вказати його назву, опис і ключові слова. Після збереження система повинна створити відповідний запис у таблиці Brands і відобразити цей бренд у списку доступних для аналізу. Також потрібно перевірити ситуації, коли користувач залишає обов'язкові поля порожніми або вводить некоректні дані.

На другому етапі перевіряється процес завантаження повідомлень. Система повинна приймати набір текстових повідомлень, перевіряти їхню структуру та зберігати коректні записи у таблиці Messages. Під час такого тестування необхідно врахувати різні типи вхідних даних: повідомлення з позитивним, негативним і нейтральним змістом, повідомлення з посиланнями, хештегами, зайвими символами, різними регістрами літер, а також можливі дублікати. Якщо вхідні дані мають неправильний формат або не містять повідомлень, система повинна

вивести повідомлення про помилку та не створювати некоректні записи в базі даних.

На третьому етапі необхідно перевірити роботу модуля попередньої обробки тексту. Цей модуль повинен очищати повідомлення від зайвих символів, посилань, службової інформації, хештегів, згадок користувачів і зайвих пробілів. Також текст повинен приводитися до нижнього регістру для подальшого аналізу. Така перевірка дозволяє встановити, чи однаково система обробляє повідомлення різного формату та чи правильно готує текст до визначення тональності.

На четвертому етапі перевіряється алгоритм визначення емоційного ставлення аудиторії до бренду. Для цього використовується набір повідомлень із задалегідь визначеним очікуваним результатом. Система повинна знайти позитивні та негативні ознаки в тексті, обчислити індекс тональності й визначити клас повідомлення: позитивне, негативне або нейтральне.

Для перевірки роботи алгоритму може бути використано набір тестових повідомлень, наведений у таблиці 3.3.

Таблиця 3.3 - Тестування алгоритму визначення тональності повідомлень

№	Тестове повідомлення	Очікуваний результат	Отриманий результат
1	Дуже якісний сервіс, швидка доставка	Позитивне	Позитивне
2	Погане обслуговування, товар не сподобався	Негативне	Негативне
3	Компанія представила новий продукт	Нейтральне	Нейтральне
4	Не рекомендую цей бренд	Негативне	Негативне
5	Зручно, швидко і приємно користуватися	Позитивне	Позитивне
6	Сьогодні магазин працює до 20:00	Нейтральне	Нейтральне

За результатами тестування встановлено, що алгоритм коректно класифікує прості повідомлення з явно вираженою тональністю. Позитивні повідомлення були віднесені до позитивного класу, негативні - до негативного, а інформаційні повідомлення без емоційного забарвлення - до нейтрального. Це підтверджує працездатність базового алгоритму аналізу тональності.

Окремо було перевірено обробку заперечень. У повідомленнях типу «не подобається якість товару» система повинна враховувати, що наявність слова «подобається» не завжди означає позитивне ставлення. Якщо перед позитивною ознакою стоїть заперечення «не», така конструкція повинна враховуватися як негативна. Тестування показало, що врахування простих заперечень дозволяє підвищити точність класифікації повідомлень і зменшити кількість помилкових позитивних результатів.

На п'ятому етапі тестувалася робота збереження результатів аналізу у базі даних. Після виконання аналізу для кожного повідомлення система повинна створити запис у таблиці SentimentResults. У цьому записі зберігається ідентифікатор повідомлення, визначений клас тональності, числовий індекс тональності, рівень упевненості та дата виконання аналізу. Перевірка показала, що результати аналізу коректно зберігаються у базі даних і можуть бути використані для формування статистичних звітів.

На шостому етапі було перевірено формування статистики та звітів. Система повинна підраховувати кількість позитивних, негативних і нейтральних повідомлень за вибраним брендом та періодом часу. Також перевірялося відображення результатів у вигляді таблиць і діаграм. У результаті тестування встановлено, що система коректно формує підсумкові показники та дозволяє користувачеві швидко оцінити загальний рівень емоційного ставлення аудиторії до бренду.

Основні результати функціонального тестування наведено в таблиці 3.4.

					КвРІСТ 2401217.22.01.02 ПЗ	Арк. 57
Зм.	Арк.	№ докум.	Підпис	Дата		

Таблиця 3.4 - Результати функціонального тестування системи

№	Функція системи	Очікуваний результат	Результат тестування
1	Додавання бренду	Бренд збережено в базі даних	Виконано
2	Завантаження повідомлень	Повідомлення додано до таблиці Messages	Виконано
3	Очищення тексту	Видалено зайві символи, посилання та службові елементи	Виконано
4	Аналіз тональності	Повідомлення класифіковано як позитивне, негативне або нейтральне	Виконано
5	Збереження результату	Результат записано до таблиці SentimentResults	Виконано
6	Формування статистики	Підраховано кількість повідомлень кожного класу	Виконано
7	Відображення результатів	Дані показано у вигляді таблиць і діаграм	Виконано
8	Фільтрація повідомлень	Дані відфільтровано за брендом, датою або тональністю	Виконано

Під час тестування також було перевірено коректність роботи системи у випадку введення некоректних або неповних даних. Наприклад, якщо користувач не вказує назву бренду або намагається завантажити порожній набір повідомлень, система повинна повідомити про помилку та не виконувати подальшу обробку. Така перевірка дозволяє уникнути появи некоректних записів у базі даних та підвищує надійність роботи програмного забезпечення.

Важливим критерієм тестування була зручність користувацького інтерфейсу. Під час перевірки оцінювалася логічність розміщення основних елементів, доступність функцій, зрозумілість повідомлень про помилки та простота перегляду результатів аналізу. Встановлено, що структура інтерфейсу дозволяє користувачу послідовно виконувати основні дії: додати бренд,

завантажити повідомлення, запустити аналіз, переглянути результати та сформувавши звіт.

За результатами тестування можна зробити висновок, що розроблена інформаційна система виконує основні функції, визначені під час проєктування. Система забезпечує збереження повідомлень, їх попередню обробку, визначення емоційного забарвлення, запис результатів до бази даних і формування аналітичних показників. Разом із тим базовий словниковий алгоритм має певні обмеження. Він добре працює з простими повідомленнями, однак може помилятися у випадках сарказму, іронії, складних мовних конструкцій або використання неформальної лексики. Тому подальше вдосконалення системи може передбачати використання моделей машинного навчання, розширення словників тональності та адаптацію алгоритму до особливостей українськомовного контенту [17, 31].

3.7 Висновки до третього розділу

У межах третього розділу було описано програмну реалізацію інформаційної системи визначення емоційного ставлення аудиторії до бренду на основі аналізу даних соціальних мереж. Основну увагу приділено практичній реалізації програмних модулів, створенню бази даних, опису алгоритму аналізу тональності, побудові користувацького інтерфейсу та перевірці працездатності системи.

Розглянуто реалізацію основних модулів програмного забезпечення інформаційної системи. До складу системи входять модуль користувацького інтерфейсу, модуль адміністрування, модуль збору даних, модуль попередньої обробки тексту, модуль аналізу тональності, модуль збереження даних та модуль формування звітів. Кожен із цих модулів виконує окрему функцію в загальному процесі роботи системи, а їх взаємодія забезпечує повний цикл обробки

повідомлень: від отримання текстових даних до формування підсумкових аналітичних результатів.

Описано процес створення бази даних інформаційної системи. Визначено основні таблиці бази даних, зокрема Brands, Sources, Messages, SentimentResults, Users та Reports. Для кожної таблиці визначено її призначення, основні поля та зв'язки з іншими таблицями. Така структура дозволяє зберігати інформацію про бренди, джерела даних, повідомлення користувачів, результати аналізу тональності, облікові записи користувачів і сформовані звіти. Також наведено приклади SQL-запитів для створення таблиць та індексів, що забезпечують ефективніший пошук, вибірку й обробку даних.

Окрему увагу приділено реалізації алгоритму визначення емоційного ставлення аудиторії до бренду. Було розглянуто послідовність обробки текстового повідомлення, яка включає отримання повідомлення з бази даних, очищення тексту від зайвих символів, видалення посилань і службових елементів, приведення тексту до нижнього регістру, розбиття повідомлення на окремі слова, порівняння зі словниками позитивних і негативних ознак та обчислення індексу тональності. На основі отриманого значення індексу система визначає, до якого класу належить повідомлення: позитивного, негативного або нейтрального.

У процесі опису алгоритму також було враховано можливість обробки простих заперечень, що є важливим для підвищення точності аналізу. Наприклад, фраза «не подобається» не повинна класифікуватися як позитивна лише через наявність слова з позитивним емоційним забарвленням. Тому в алгоритмі передбачено перевірку наявності заперечення перед емоційною ознакою. Це дозволяє зменшити кількість помилкових результатів і зробити класифікацію повідомлень більш коректною.

Також було описано реалізацію користувацького інтерфейсу інформаційної системи. Визначено основні сторінки системи, серед яких сторінка авторизації, головна сторінка, сторінка керування брендами, сторінка завантаження повідомлень, сторінка перегляду повідомлень, сторінка результатів аналізу,

сторінка формування звітів і сторінка адміністрування. Така структура інтерфейсу забезпечує зручну взаємодію користувача із системою та дозволяє виконувати основні операції без безпосередньої роботи з програмним кодом або базою даних.

У підрозділі тестування було перевірено працездатність основних функцій інформаційної системи. Зокрема, було протестовано додавання брендів, завантаження повідомлень, попередню обробку тексту, визначення тональності повідомлень, збереження результатів у базі даних, формування статистики та відображення результатів у користувацькому інтерфейсі. Результати тестування показали, що система коректно виконує основні операції та забезпечує послідовну обробку даних відповідно до поставлених функціональних вимог.

Під час тестування алгоритму було використано набір повідомлень із позитивним, негативним і нейтральним змістом. Отримані результати підтвердили, що базовий алгоритм правильно класифікує прості повідомлення з явно вираженою тональністю. Позитивні повідомлення були віднесені до позитивного класу, негативні - до негативного, а інформаційні повідомлення без вираженого емоційного забарвлення - до нейтрального. Також підтверджено коректність збереження результатів аналізу у таблиці `SentimentResults` та можливість подальшого використання цих даних для побудови статистики й формування звітів.

За результатами виконаної програмної реалізації та тестування можна зробити висновок, що розроблена інформаційна система забезпечує повний цикл роботи з текстовими повідомленнями соціальних мереж. Вона дозволяє зберігати повідомлення, виконувати їх попередню обробку, визначати емоційне забарвлення, записувати результати аналізу до бази даних і подавати підсумкову інформацію користувачеві у зручному вигляді.

Разом із тим встановлено, що використання базового словникового підходу має певні обмеження. Такий підхід є простим у реалізації та зрозумілим для аналізу, однак його точність залежить від повноти словників позитивних і негативних ознак. Крім того, алгоритм може мати труднощі з обробкою сарказму,

					КвРІСТ 2401217.22.01.02 ПЗ	Арк. 61
Зм.	Арк.	№ докум.	Підпис	Дата		

іронії, жаргону, складних мовних конструкцій та контекстно залежних висловлювань. Тому подальше вдосконалення системи може передбачати використання складніших моделей аналізу тексту, зокрема методів машинного навчання або трансформерних моделей, розширення словників тональності, покращення обробки українськомовних повідомлень та інтеграцію з API соціальних мереж.

Таким чином, у третьому розділі було підтверджено практичну можливість реалізації інформаційної системи визначення емоційного ставлення аудиторії до бренду. Описані програмні модулі, база даних, алгоритм аналізу тональності, користувацький інтерфейс і результати тестування свідчать про працездатність запропонованого рішення та його відповідність основним завданням кваліфікаційної роботи.

					КвРІСТ 2401217.22.01.02 ПЗ	Арк. 62
Зм.	Арк.	№ докум.	Підпис	Дата		

ВИСНОВКИ

У роботі за результатами виконаних теоретичних та практичних досліджень було розроблено підхід до створення інформаційної системи визначення емоційного ставлення аудиторії до бренду на основі аналізу даних соціальних мереж. У ході виконання роботи було розглянуто особливості збору, обробки, збереження та аналізу текстових повідомлень користувачів, а також запропоновано структуру програмного забезпечення, бази даних і алгоритм визначення тональності повідомлень.

У першому розділі проведено аналіз предметної області визначення емоційного ставлення аудиторії до бренду. Встановлено, що соціальні мережі є важливим джерелом інформації про ставлення користувачів до бренду, оскільки вони містять коментарі, відгуки, згадки, реакції та інші повідомлення, які можуть мати позитивне, негативне або нейтральне забарвлення. Також було розглянуто основні проблеми аналізу текстових даних соціальних мереж, зокрема великі обсяги інформації, наявність неформальної лексики, сарказму, іронії, помилок, хештегів і скорочень.

У межах першого розділу також було здійснено порівняльний аналіз існуючих рішень для соціального моніторингу та аналізу тональності. Визначено, що корпоративні платформи мають широкий функціонал, але є дорогими та складними у впровадженні. Доступніші інструменти для малого й середнього бізнесу простіші у використанні, однак мають обмежені можливості аналітики. Хмарні та open-source рішення забезпечують гнучкість, проте потребують додаткової технічної реалізації. Це підтвердило доцільність розроблення власної інформаційної системи, адаптованої до поставлених вимог.

У другому розділі проведено проєктування інформаційної системи визначення емоційного ставлення аудиторії до бренду. Було визначено основні підсистеми системи, зокрема підсистему збору даних, підсистему попередньої обробки тексту, підсистему аналізу тональності, підсистему збереження даних,

					КвРІСТ 2401217.22.01.02 ПЗ	Арк. 63
Зм.	Арк.	№ докум.	Підпис	Дата		

підсистему формування звітів, підсистему користувацького інтерфейсу та підсистему адміністрування. Також розроблено архітектуру інформаційної системи, описано взаємодію користувача з веб-інтерфейсом і серверною частиною, визначено програмні засоби для реалізації системи та побудовано логічну схему бази даних.

У процесі проєктування було визначено функціональні та нефункціональні вимоги до інформаційної системи. До основних функціональних вимог віднесено збір повідомлень, попередню обробку тексту, визначення тональності, збереження результатів аналізу, формування статистики та звітів, а також візуалізацію результатів. Нефункціональні вимоги охоплюють продуктивність, масштабованість, надійність, безпеку даних і зручність користувацького інтерфейсу.

У третьому розділі описано програмну реалізацію інформаційної системи. Розглянуто реалізацію основних модулів програмного забезпечення, середовище розроблення та структуру програмного проєкту. Описано логічне розміщення основних файлів і каталогів, зокрема файлів для запуску системи, роботи з базою даних, попередньої обробки тексту, аналізу тональності, формування звітів, а також каталогів для HTML-шаблонів, стилів, тестових даних і журналів подій.

Також у третьому розділі описано процес створення бази даних інформаційної системи. Було визначено основні таблиці бази даних: Brands, Sources, Messages, SentimentResults, Users та Reports. Наведено приклади SQL-запитів для створення таблиць та індексів. Запропонована структура бази даних дозволяє зберігати повідомлення, результати їх аналізу, інформацію про бренди, джерела даних, користувачів системи та сформовані звіти.

Описано алгоритм визначення емоційного ставлення аудиторії до бренду. Алгоритм передбачає отримання повідомлення з бази даних, очищення тексту, приведення його до нижнього регістру, розбиття на окремі слова, порівняння зі словниками позитивних і негативних ознак, обчислення індексу тональності та класифікацію повідомлення як позитивного, негативного або нейтрального.

Додатково розглянуто можливість урахування простих заперечень, що дозволяє підвищити точність класифікації.

У роботі також описано структуру користувацького інтерфейсу системи. Визначено основні сторінки, які забезпечують авторизацію користувача, керування брендами, завантаження повідомлень, перегляд повідомлень, запуск аналізу, перегляд результатів, формування звітів і адміністрування. Такий інтерфейс дозволяє користувачеві працювати із системою без прямого доступу до бази даних або програмного коду.

Виконано тестування роботи інформаційної системи. У процесі тестування перевірено додавання брендів, завантаження повідомлень, попередню обробку тексту, роботу алгоритму визначення тональності, збереження результатів у базі даних, формування статистики, відображення результатів і фільтрацію повідомлень. Результати тестування підтвердили, що система коректно виконує основні функції та забезпечує послідовну обробку даних.

Отже, поставлену мету роботи було досягнуто. У результаті виконання роботи спроектовано та описано програмну реалізацію інформаційної системи, яка забезпечує збір повідомлень із соціальних мереж, попередню обробку текстових даних, визначення тональності повідомлень, збереження результатів у базі даних і формування аналітичних звітів. Запропонована система може бути використана для моніторингу репутації бренду, оцінювання реакції аудиторії на маркетингові кампанії, виявлення негативних інформаційних проявів і підтримки прийняття управлінських рішень.

Подальше вдосконалення інформаційної системи може передбачати інтеграцію з API соціальних мереж, використання моделей машинного навчання, розширення словників тональності, покращення обробки українськомовних повідомлень, додавання засобів аналізу сарказму та іронії, а також реалізацію модуля прогнозування змін емоційного ставлення аудиторії до бренду.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ

1. Liu B. Sentiment Analysis and Opinion Mining. Morgan & Claypool Publishers, 2012. 168 p. URL: <https://www.cs.uic.edu/~liub/FBS/SentimentAnalysis-and-OpinionMining.pdf> (дата зверення 21.04.2026)
2. Pang B., Lee L. Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*. 2008. Vol. 2, № 1-2. P. 1-135. URL: <https://www.cs.cornell.edu/home/llee/omsa/omsa.pdf> (дата зверення 22.04.2026)
3. Jurafsky D., Martin J. H. Speech and Language Processing. 3rd ed. Stanford University, 2023. URL: <https://web.stanford.edu/~jurafsky/slp3/> (дата зверення 04.05.2026)
4. Aggarwal C. C. Machine Learning for Text. Springer, 2018. 493 p. URL: <https://charuaggarwal.net/Text-Learning.pdf> (дата зверення 04.05.2026)
5. Feldman R. Techniques and Applications for Sentiment Analysis. *Communications of the ACM*. 2013. URL: <https://dl.acm.org/doi/abs/10.1145/2436256.2436274> (дата зверення 23.04.2026)
6. Medhat W., Hassan A., Korashy H. Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*. 2014. URL: <https://www.sciencedirect.com/science/article/pii/S2090447914000550> (дата зверення 23.04.2026)
7. Cambria E., Schuller B., Xia Y., Havasi C. New Avenues in Opinion Mining and Sentiment Analysis. *IEEE Intelligent Systems*. 2013. URL: <https://ieeexplore.ieee.org/abstract/document/6468032/> (дата зверення 04.05.2026)
8. Bird S., Klein E., Loper E. Natural Language Processing with Python. O'Reilly Media, 2009. 504p. URL: https://books.google.de/books?hl=uk&lr=&id=KGIbfiiP1i4C&oi=fnd&pg=PR5&dq=1.%09Bird+S.,+Klein+E.,+Loper+E.+Natural+Language+Processing+with+Python.+O'Reilly+Media,+2009.+504+p.&ots=Y5FnC7FBJ0&sig=616CFtAcLOkqMrW-J-97p_3svro&redir_esc=y#v=onepage&q&f=false (дата зверення 19.04.2026)

					КвРІСТ 2401217.22.01.02 ПЗ	Арк. 66
Зм.	Арк.	№ докум.	Підпис	Дата		

9. Raschka S., Mirjalili V. Python Machine Learning. 3rd ed. Packt Publishing, 2019. URL: https://books.google.de/books?hl=uk&lr=&id=sKXIDwAAQBAJ&oi=fnd&pg=PP1&dq=9.%09Raschka+S.,+Mirjalili+V.+Python+Machine+Learning.+3rd+ed.+Packt+Publishing,+2019.+&ots=VaJpsVQCLq&sig=EX8po6Xz0PCQwTLuuzwEHPZRNDw&redir_esc=y#v=onepage&q&f=false (дата зверення 04.05.2026)

10. Grinberg M. Flask Web Development: Developing Web Applications with Python. 2nd ed. O'Reilly Media, 2018. URL: https://books.google.de/books?hl=uk&lr=&id=cVIPDwAAQBAJ&oi=fnd&pg=PT255&dq=10.%09Grinberg+M.+Flask+Web+Development:+Developing+Web+Applications+with+Python.+2nd+ed.+O'Reilly+Media,+2018.+&ots=xPI_hpYI9W&sig=PANmthvRzUq_iOtFS0EGGoHLW-U&redir_esc=y#v=onepage&q&f=false (дата зверення 25.04.2026)

11. Elmasri R., Navathe S. B. Fundamentals of Database Systems. 7th ed. Boston : Pearson, 2016. 1280 p. URL: <http://diglib.globalcollege.edu.et:8080/xmlui/bitstream/handle/123456789/2640/Fundamental%20of%20Database%20Systems.pdf?sequence=1&isAllowed=y> (дата зверення 07.05.2026)

12. Silberschatz A., Korth H. F., Sudarshan S. Database System Concepts. 7th ed. New York : McGraw-Hill Education, 2019. 1376 p. URL: <http://eng1.mu.edu.tr/~pembeci/tmp/Korth/DBMS%20Korth-%20Solution.pdf> (дата зверення 29.04.2026)

13. Pressman R. S., Maxim B. R. Software Engineering: A Practitioner's Approach. 8th ed. New York : McGraw-Hill Education, 2014. 976 p. URL: <https://studylib.net/doc/26338727/2.-software-engineering-a-practitioner's-approach?p=2> (дата зверення 04.05.2026)

14. Russell M. A., Klassen M. Mining the Social Web: Data Mining Facebook, Twitter, LinkedIn, Instagram, GitHub, and More. 3rd ed. Sebastopol : O'Reilly Media, 2019. 426 p. URL: https://books.google.de/books?hl=uk&lr=&id=_VkrAQAAQBAJ (дата зверення 07.05.2026)

15. Han J., Kamber M., Pei J. Data Mining: Concepts and Techniques. 3rd ed. Waltham : Morgan Kaufmann, 2012. 744 p. URL: <https://homes.di.unimi.it/ceselli/IM/2012-13/slides/02-KnowYourData.pdf> (дата зверення 08.05.2026)

16. Duo W. L., Zhou M. C., Abusorrah A. A survey of cyber attacks on cyber physical systems: Recent advances and challenges. *IEEE/CAA Journal of Automatica Sinica*. 2022. Vol. 9, № 5. P. 784–800. URL: <https://ieeexplore.ieee.org/abstract/document/9763485> (дата зверення 08.05.2026)

17. Devlin J., Chang M. W., Lee K., Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of NAACL-HLT*. 2019. P. 4171–4186. URL: <https://arxiv.org/abs/1810.04805> (дата зверення 04.05.2026)

18. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser Ł., Polosukhin I. Attention Is All You Need. *Advances in Neural Information Processing Systems*. 2017. Vol. 30. URL: <https://arxiv.org/abs/1706.03762> (дата зверення 07.05.2026)

19. Brownlee J. Deep Learning for Natural Language Processing. Vermont : Machine Learning Mastery, 2017. 414 p. URL: https://books.google.de/books?hl=uk&lr=&id=_pmoDwAAQBAJ (дата зверення 22.05.2026)

20. Mikolov T., Chen K., Corrado G., Dean J. Efficient Estimation of Word Representations in Vector Space. arXiv preprint arXiv:1301.3781. 2013. URL: <https://arxiv.org/abs/1301.3781> (дата зверення 10.05.2026)

21. Goldberg Y. Neural Network Methods for Natural Language Processing. San Rafael : Morgan & Claypool Publishers, 2017. URL: <https://cris.bgu.ac.il/en/publications/neural-network-methods-for-natural-language-processing-2/> (дата зверення 11.05.2026)

22. Manning C. D., Raghavan P., Schütze H. Introduction to Information Retrieval. Cambridge : Cambridge University Press, 2008. URL: <https://nlp.stanford.edu/IR-book/> (дата зверення 04.05.2026)

					КвРІСТ 2401217.22.01.02 ПЗ	Арк. 68
Зм.	Арк.	№ докум.	Підпис	Дата		

23. Goodfellow I., Bengio Y., Courville A. Deep Learning. Cambridge : MIT Press, 2016. 800 p. URL: <https://www.deeplearningbook.org/> (дата зверення 05.05.2026)

24. Chollet F. Deep Learning with Python. 2nd ed. Shelter Island : Manning Publications, 2021. URL: <https://books.google.de/books?hl=uk&lr=&id=XHpKEAAAQBAJ> (дата зверення 07.05.2026)

25. Lane H., Howard C., Hapke H. Natural Language Processing in Action. Shelter Island : Manning Publications, 2019. 544 p. URL: <https://books.google.de/books?hl=uk&lr=&id=WkIAEQAAQBAJ> (дата зверення 03.05.2026)

26. Baeza-Yates R., Ribeiro-Neto B. Modern Information Retrieval: The Concepts and Technology Behind Search. 2nd ed. Harlow : Pearson, 2011. URL: <https://people.ischool.berkeley.edu/~hearst/irbook/print/chap10.ps.gz> (дата зверення 02.05.2026)

27. Alpaydin E. Introduction to Machine Learning. 4th ed. Cambridge : MIT Press, 2020. 712 p. URL: <https://books.google.de/books?hl=uk&lr=&id=uZnSDwAAQBAJ> (дата зверення 02.05.2026)

28. Géron A. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. 3rd ed. Sebastopol : O'Reilly Media, 2022. 864 p. URL: <https://books.google.de/books?hl=uk&lr=&id=X5ySEAAAQBAJ> (дата зверення 01.05.2026)

29. McKinney W. Python for Data Analysis. 3rd ed. Sebastopol : O'Reilly Media, 2022. URL: https://books.google.de/books?hl=uk&lr=&id=v3n4_AK8vu0C (дата зверення 14.05.2026)

30. Reimers N., Gurevych I. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*. Hong Kong : Association for Computational Linguistics, 2019. P. 3982–3992. URL: <https://arxiv.org/abs/1908.10084> (дата зверення 18.05.2026)

					КвРІСТ 2401217.22.01.02 ПЗ	Арк. 69
Зм.	Арк.	№ докум.	Підпис	Дата		

31. Wolf T. et al. Transformers: State-of-the-Art Natural Language Processing. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations. 2020. P. 38–45. URL: <https://aclanthology.org/2020.emnlp-demos.6/> (дата звернення: 24.05.2026).
32. Apache Kafka Documentation. URL: <https://kafka.apache.org/documentation/> (дата звернення: 23.05.2026).
33. PostgreSQL Global Development Group. PostgreSQL Documentation. URL: <https://www.postgresql.org/docs/> (дата звернення: 26.05.2026).
34. Mozilla Developer Network. REST API Introduction. URL: <https://developer.mozilla.org/en-US/docs/Glossary/REST> (дата звернення: 23.05.2026).
35. TensorFlow Documentation. URL: <https://www.tensorflow.org/> (дата звернення: 29.05.2026).
36. Hugging Face Documentation. URL: <https://huggingface.co/docs> (дата звернення: 16.05.2026).
37. spaCy Industrial-Strength NLP Documentation. URL: <https://spacy.io/> (дата звернення: 15.05.2026).
38. NLTK Documentation. URL: <https://www.nltk.org/> (дата звернення: 22.05.2026).
39. FastAPI Documentation. URL: <https://fastapi.tiangolo.com/> (дата звернення: 21.05.2026).
40. React Documentation. URL: <https://react.dev/> (дата звернення: 28.05.2026).
41. Bootstrap Documentation. URL: <https://getbootstrap.com/> (дата звернення: 18.05.2026).
42. Kelleher J. D., Tierney B. Data Science. Cambridge : MIT Press, 2018. (дата звернення 07.05.2026)
43. Raschka S. Machine Learning Q and AI. Leanpub, 2023. URL: <https://books.google.de/books?hl=uk&lr=&id=dT7hEAAAQBAJ> (дата звернення 07.05.2026)

					КвРІСТ 2401217.22.01.02 ПЗ	Арк. 70
Зм.	Арк.	№ докум.	Підпис	Дата		

44. Leskovec J., Rajaraman A., Ullman J. D. Mining of Massive Datasets. Cambridge : Cambridge University Press, 2020. URL: <http://www.mmnds.org/> (дата зверення 14.05.2026)

45. Witten I. H., Frank E., Hall M. A., Pal C. J. Data Mining: Practical Machine Learning Tools and Techniques. 4th ed. Amsterdam : Morgan Kaufmann, 2016. URL: [http://alvarestech.com/temp/Industria4.0/2019/\(Morgan%20Kaufmann%20Series%20in%20Data%20Management%20Systems\)%20Ian%20H.%20Witten,%20Eibe%20Frank,%20Mark%20A.%20Hall,%20Christopher%20J.%20Pal%20-%20Data%20Mining_%20Practical%20Machine%20Learning%20Tools%20and%20Techniques-Morgan%20Kaufmann%20Publisher.pdf](http://alvarestech.com/temp/Industria4.0/2019/(Morgan%20Kaufmann%20Series%20in%20Data%20Management%20Systems)%20Ian%20H.%20Witten,%20Eibe%20Frank,%20Mark%20A.%20Hall,%20Christopher%20J.%20Pal%20-%20Data%20Mining_%20Practical%20Machine%20Learning%20Tools%20and%20Techniques-Morgan%20Kaufmann%20Publisher.pdf) (дата зверення 18.05.2026)

46. Sarkar D. Text Analytics with Python. 2nd ed. Berkeley : Apress, 2019. URL: <https://link.springer.com/book/10.1007/978-1-4842-4354-1> (дата зверення 18.05.2026)

47. Jurafsky D. Natural Language Processing. Stanford University Lecture Notes, 2022. URL: <https://web.stanford.edu/class/cs224n/> (дата зверення 18.05.2026)

48. IBM Cloud Education. What is Sentiment Analysis? URL: <https://www.ibm.com/topics/sentiment-analysis> (дата зверення 23.05.2026)

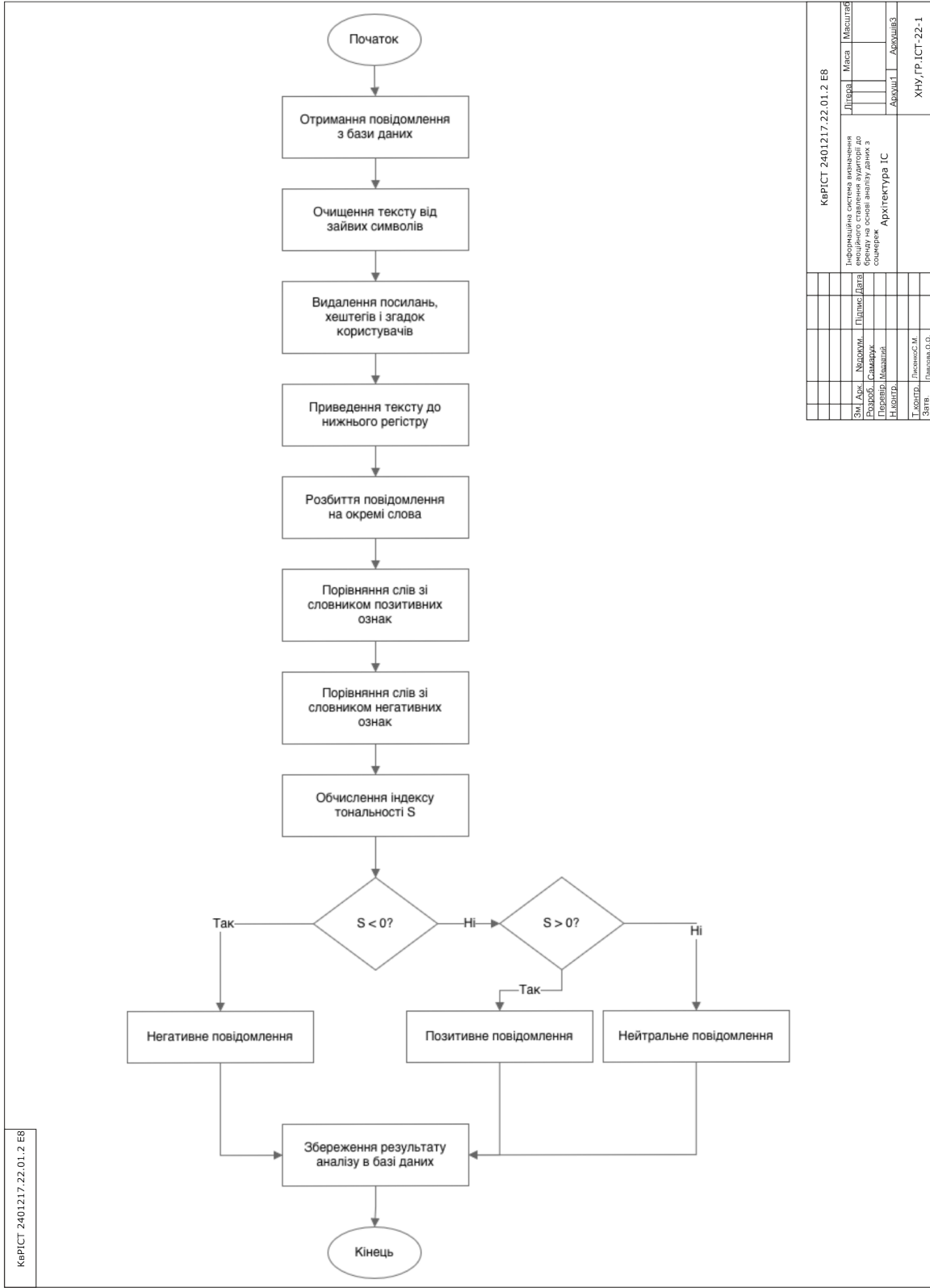
49. Oracle Documentation. MySQL 8.0 Reference Manual. URL: <https://dev.mysql.com/doc/> (дата зверення 22.05.2026).

50. Scikit-learn Documentation. URL: <https://scikit-learn.org/stable/> (дата зверення 11.05.2026).

ДОДАТОК В

(обов'язковий)

Копія креслення «Алгоритм роботи ІС»



КвРІСТ 2401217.22.01.2.E8		Дата	Масштаб
Зм. Арк.	Всього Арк.	Підпис	Дата
Розроб.	Складено	Перевір.	Модифік.
Н. Кольца	І. Лещенко	С.М. Зайва	І. Лещенко
Інформаційна система вивчення речовинного складу аудиторії до складу системи аналізу даних з соціальних мереж		Архитектура ІС	
ХНУ «Львівський національний університет імені Г.С.Сковороди»		ХНУ. ГР. ІСТ-22-1	

КвРІСТ 2401217.22.01.2.E8

Протокол аналізу звіту подібності експертом

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

Автор: Владислав САМАРУК

Співавтор:

Назва: Інформаційна система визначення емоційного ставлення аудиторії до бренду на основі аналізу даних соцмереж

Експерт: Дмитро МЕДЗАТИЙ

Підрозділ: Кафедра комп'ютерної інженерії та інформаційних систем

Коефіцієнт подібності 1: 6.62%

Коефіцієнт подібності 2: 1.47%

Мікропробіли: 12

Заміна букв: 0

Інтервали: 0

Білі знаки: 0

Дата створення звіту: 2026-06-02 02:56:18.0

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 процедур. Таким чином робота не приймається.

Обґрунтування:

2026-06-02

Дата

Доцент Андрій Нічепорук

експерт

Anti-Plagiarism (<http://ap.km.ua>) v-15.701

Максимальне співпадіння з одним документом 13.0%

Словники перевірки: en_US, ru_RU, ua_UA. Помилки в документах: 9%

ID: 273039 Назва: БКР Інформаційна система визначення емоційного ставлення аудиторії до бренду на основі аналізу даних соцмереж Додано в БД: 2026-06-01 Автора: Владислав САМАРУК Керівники: Дмитро МЕДЗАТІЙ Консультанти: Опоненти:	Документ		Сумарний збіг по Базі Даних	
	Символи	Лексеми	Символи	Лексеми
	104580	880	16994 (16%)	164 (19%)

Джерело плагіату

ID	Опис	Наявність плагіату в документі	
		Символи	Лексеми
269471	Назва: Звіт з ПДП Інформаційна система визначення емоційного ставлення аудиторії до бренду на основі аналізу даних Додано в БД: 2026-02-24 Автора: В. В. Самарука Керівники: Павлова О.О. Консультанти: Опоненти:	13190 (13.0%)	109 (12.0%)

РЕЦЕНЗІЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ

Дипломник: Самарук Владислав Володимирович

Тема: Інформаційна система визначення емоційного ставлення аудиторії до бренду на основі аналізу даних соцмереж

Спеціальність: 126 «Інформаційні системи та технології»

Обсяг кваліфікаційної роботи:

Кількість листів креслень 3 Кількість сторінок записки 74

1. Короткий зміст роботи та прийнятих рішень: Метою кваліфікаційної роботи є розробка інформаційної системи визначення емоційного ставлення аудиторії до бренду на основі аналізу даних соціальних мереж. У роботі розглянуто процеси збору, попередньої обробки, збереження та аналізу текстових повідомлень, а також реалізовано програмні засоби для класифікації повідомлень за тональністю.

2. Висновок про відповідність роботи дипломному завданню: Робота повністю відповідає поставленому завданню. У межах кваліфікаційної роботи виконано аналіз предметної області, визначено вимоги до інформаційної системи, розроблено її архітектуру, спроектовано базу даних та реалізовано основні програмні модулі.

3. Характеристика виконання кожного розділу, ступінь використання останніх досягнень науки і техніки і передових методів роботи: У першому розділі кваліфікаційної роботи проведено дослідження предметної області, проаналізовано сучасні підходи до моніторингу соціальних мереж, визначення тональності тексту та оцінювання емоційного ставлення аудиторії до бренду. Розглянуто методи обробки природної мови, алгоритми класифікації текстових даних, а також особливості використання інформаційних систем у сфері цифрового маркетингу та репутаційного аналізу.

У другому розділі виконано проектування інформаційної системи визначення емоційного ставлення аудиторії до бренду. Визначено функціональні та нефункціональні вимоги до системи, розроблено її загальну архітектуру, описано основні програмні модулі, спроектовано структуру бази даних та визначено логіку

взаємодії між компонентами системи. Особливу увагу приділено процесам збору повідомлень із соціальних мереж, очищення текстових даних, визначення тональності повідомлень і збереження результатів аналізу.

У третьому розділі кваліфікаційної роботи виконано програмну реалізацію інформаційної системи. Реалізовано модулі попередньої обробки тексту, аналізу емоційного забарвлення повідомлень, роботи з базою даних та подання результатів користувачу. Також проведено тестування основних функціональних можливостей системи, зокрема перевірено коректність обробки позитивних, негативних і нейтральних повідомлень, правильність збереження результатів та зручність відображення аналітичної інформації.

4. Позитивні сторони роботи: до позитивних сторін роботи можна віднести актуальність теми, практичну спрямованість розробки, логічну структуру кваліфікаційної роботи, використання сучасних підходів до обробки текстових даних і можливість практичного застосування системи для моніторингу репутації бренду та підтримки прийняття управлінських рішень у сфері цифрового маркетингу.

5. Негативні сторони роботи: до недоліків роботи можна віднести недостатньо глибоке дослідження можливостей використання складніших моделей машинного навчання та обмежений обсяг тестового набору даних для перевірки точності визначення тональності повідомлень.

6. Оцінка графічного оформлення та пояснювальної записки роботи: Пояснювальна записка оформлена коректно, згідно з чинними вимогами до оформлення кваліфікаційних робіт. Графічні матеріали є зрозумілими, відповідають змісту роботи та відображають основні архітектурні й функціональні рішення, прийняті під час розробки інформаційної системи.

7. Відгук про роботу в цілому: Робота виконана на належному науково-технічному рівні.

8. Інші зауваження: _____

9. Оцінка дипломної роботи: добре /В/85

Рецензент (прізвище, ім'я, по батькові, посада, місце роботи) _____

Олександр Михайлович, доцент КНУ, ХНУ

"3"

6

2026 р.

(підпис)

Зав. кафедри КІС
д-р. філософії Ользі ПАВЛОВІЙ

Самарук Владислав Володимирович

ПІБ здобувача вищої освіти

ІСТ, 4 курсу, групи ІСТ-22-1

ЗАЯВА

З правилами чинного Положення про систему забезпечення академічної доброчесності у Хмельницькому національному університеті, згідно з яким виявлення академічного плагіату є підставою для відмови в допуску кваліфікаційної роботи до захисту і застосування заходів академічної відповідальності, ознайомлений (а). Про використання спеціалізованих програмних засобів (СПЗ) StrikePlagiarism та Anti-Plagiarism для перевірки кваліфікаційних робіт здобувачів вищої освіти на наявність академічного плагіату оповіщений (а). Надаю університету право на передачу моєї роботи для обробки та збереження в базах даних СПЗ і використання роботи для виявлення академічного плагіату в інших роботах, які перевіряються СПЗ.

Також надаю свою згоду на обробку й збереження університетом моєї роботи в Інституційному репозитарії Хмельницького національного університету.

Робота надається для перевірки в електронному варіанті. Електронна версія моєї роботи збігається (ідентична) з друкованою.

31 травня 2026 року

РІШЕННЯ ЕКСПЕРТНОЇ КОМІСІЇ

КАФЕДРИ КОМП'ЮТЕРНОЇ ІНЖЕНЕРІЇ ТА ІНФОРМАЦІЙНИХ СИСТЕМ ПРО ДОПУСК КВАЛІФІКАЦІЙНОЇ РОБОТИ ДО ЗАХИСТУ

Назва кваліфікаційної роботи Інформаційна система визначення емоційного ставлення аудиторії до бренду на основі аналізу даних соцмереж

Автор Владислав САМАРУК

Освітня програма Інформаційні системи та технології

Рівень вищої освіти перший (бакалаврський)

Спеціальність 126 Інформаційні системи та технології

Науковий керівник: к.т.н., доцент Дмитро МЕДЗАТИЙ

На основі аналізу кваліфікаційної роботи на дотримання вимог академічної доброчесності (у т.ч. відсутності ознак академічного плагіату) з урахуванням результатів перевірки роботи спеціалізованим програмним засобом(ами) 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.	відповідає
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	

Підтвердження:

Запозичення, виявлені в роботі, є законними і не є плагіатом, оскільки:

- 1) усі запозичення фрагментарні, або мають належним чином оформленні посилання;
- 2) окремі виявлені збіги є випадковими фразами або виразами, про що свідчить посилання системи на збіг з джерелами на один фрагмент речення;
- 3) всі зафіксовані системою ознаки модифікації тексту відносяться до комбінування латинських символів зі україномовними скороченнями індексів в формулах, що не є модифікацією тексту.
- 4) значна частина знайденого плагіату відноситься до списку використаних джерел

Сумарний обсяг всіх запозичень, визначений системою виявлення збігів/ ідентичності/схожості StrikePlagiarism, складає 6.62%; та системою Anti-Plagiarism складає 13%, що, з урахуванням наведених обґрунтувань, відповідає характеру наукового дослідження і свідчить на користь кваліфікаційної роботи.

01.06.2026

Завідувач кафедри

Гарант освітньої програми

Керівник кваліфікаційної роботи

Підпис

Підпис

Підпис

Ольга ПАВЛОВА

Ім'я, ПРІЗВИЩЕ

Єлизавета ГНАТЧУК

Ім'я, ПРІЗВИЩЕ

Дмитро МЕДЗАТИЙ

Ім'я, ПРІЗВИЩЕ