

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Центр _____ ННЦЗФН
(повна назва)

Кафедра _____ Штучного інтелекту
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти _____ перший (бакалаврський)

_____ Розробка інтелектуальної системи моніторингу тривожності на основі
аналізу тексту користувача
(тема)

Виконав:
здобувач _____ четвертого року навчання,
групи _____ ІТШЗ-21-1

_____ Богдан Семененко
(власне ім'я, прізвище)

Спеціальність _____ 122 Комп'ютерні науки
(код і повна назва спеціальності)

Тип програми _____ освітньо-професійна
Освітня програма _____ Штучний інтелект
(повна назва освітньої програми)

Керівник _____ ас. Марія Погурська
(посада, власне ім'я, прізвище)

Допускається до захисту

Завідувач кафедри ШІ _____
(підпис)

_____ Олег ЗОЛОТУХІН
(власне ім'я, прізвище)

2025 р.

Харківський національний університет радіоелектроніки

Центр _____ ННЦЗФН _____

Кафедра _____ Штучного інтелекту _____

Рівень вищої освіти _____ перший (бакалаврський) _____

Спеціальність _____ 122 Комп'ютерні науки _____
(код і повна назва)

Тип програми _____ освітньо-професійна _____

Освітня програма _____ Штучний інтелект _____
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____

(підпис)

«_____» _____ 20__ р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві _____ Семененку Богдану Сергійовичу _____
(прізвище, ім'я, по батькові)

1. Тема роботи _____ Розробка інтелектуальної системи моніторингу тривожності на основі аналізу тексту користувача _____

затверджена наказом університету від 7 травня 2025 р. № 80Стз

2. Термін подання студентом роботи до екзаменаційної комісії 24 червня 2025 р.

3. Вихідні дані до роботи Науково-технічні публікації з тематики обробки природної мови та психології тривожності, документація бібліотек Python (transformers, torch, scikit-learn, spacy), набір даних Sentiment Analysis for Mental Health для тренування та тестування системи _____

4. Перелік питань, що потрібно опрацювати в роботі _____

1) Аналіз предметної галузі та постановка задачі дослідження _____

2) Теоретичні основи аналізу тексту для виявлення емоційного стану _____

3) Проектування та реалізація системи моніторингу тривожності _____

4) Експериментальна перевірка та оцінка результатів _____

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Строк / терміни виконання етапів роботи	Примітка
1	Отримання завдання на кваліфікаційну роботу	05.05.2025	виконано
2	Аналіз предметної галузі та постановка задачі	06.05.2025–10.05.2025	виконано
3	Теоретичне дослідження методів аналізу тексту	11.05.2025–15.05.2025	виконано
4	Вибір архітектури, підбір даних, підготовка середовища	16.05.2025–20.05.2025	виконано
5	Розробка класифікаційної моделі та донавчання	21.05.2025–26.05.2025	виконано
6	Реалізація Telegram-бота та інтеграція компонентів	27.05.2025–01.06.2025	виконано
7	Побудова модуля симптомів, GPT-відповідей та звітності	02.06.2025–06.06.2025	виконано
8	Проведення експериментів, тестування, оформлення результатів	07.06.2025–13.06.2025	виконано
9	Перевірка пояснювальної записки на академічну добросовісність	16.06.2025	виконано
10	Нормоконтроль	17.06.2025–18.06.2025	виконано
11	Завершення оформлення, підготовка до захисту	19.06.2025–22.06.2025	виконано
12	Захист перед ЕК	24.06.2025	

Дата видачі завдання 5 травня 2025 р.

Здобувач 
(підпис)

Керівник роботи _____
(підпис)

ас. Марія Погурська
(посада, власне ім'я, прізвище)

РЕФЕРАТ

Пояснювальна записка: 74 с., 33 рис., 1 дод., 27 джерел.

АНАЛІЗ ТЕКСТУ, ОБРОБКА ПРИРОДНОЇ МОВИ,
ПСИХОЕМОЦІЙНИЙ СТАН, САМОСПОСТЕРЕЖЕННЯ,
ТРАНСФОРМЕРНА МОДЕЛЬ, ТРИВОЖНІСТЬ, GPT, TELEGRAM-БОТ.

Об'єкт дослідження – інтелектуальні методи оцінки психоемоційного стану людини.

Предмет дослідження – технології обробки текстових повідомлень з метою виявлення ознак тривожності, побудови емоційної динаміки та формування інтерпретованого зворотного зв'язку.

Мета роботи – розробити інтелектуальну систему, яка на основі аналізу тексту користувача визначає рівень тривожності, відслідковує зміни емоційного стану з часом та формує короткі підтримувальні пояснення і звіти для користувача.

Методи дослідження – методи обробки природної мови (NLP), класифікація текстів за допомогою трансформерної моделі RoBERTa, rule-based виявлення симптомів, генерація пояснень за допомогою моделі GPT, візуалізація даних та інтеграція через Telegram API.

У роботі розглянуто проблему автоматизованого моніторингу тривожності на основі текстових повідомлень користувача. Запропоновано інтелектуальну систему, що поєднує класифікацію рівня тривожності, виявлення типових симптомів, побудову графіків зміни психоемоційного стану та генерацію коротких пояснень на базі GPT. Реалізовано Telegram-інтерфейс, збереження історії в базі даних, формування текстових і PDF-звітів. Результати підтверджують доцільність використання таких рішень для індивідуального самопостереження та первинної емоційної самодіагностики.

ABSTRACT

Bachelor's thesis contains: 74 pp., 33 fig., 1 ann., 27 references.

ANXIETY, EMOTIONAL STATE, GPT, NATURAL LANGUAGE PROCESSING, SELF-MONITORING. TELEGRAM BOT, TEXT ANALYSIS, TRANSFORMER MODEL.

Object of the research – intelligent methods for assessing a person's emotional and psychological state.

Subject of the research – technologies for analyzing user-generated text messages to detect signs of anxiety, track emotional dynamics, and generate interpretable feedback.

The aim of the work – to develop an intelligent system that analyzes user messages to assess anxiety levels, track changes in emotional state over time, and generate concise supportive explanations and reports.

Research methods – natural language processing (NLP), text classification using the RoBERTa transformer model, rule-based symptom detection, generation of explanations with GPT, data visualization, and Telegram API integration.

This work addresses the problem of automated anxiety monitoring based on users' text messages. An intelligent system is proposed that integrates anxiety classification, detection of typical symptoms, visualization of emotional state dynamics, and generation of GPT-based explanatory feedback. A user-friendly Telegram interface was developed, along with data storage and report generation in text and PDF formats. The results confirm the feasibility of such approaches for individual self-monitoring and preliminary emotional self-assessment.

ЗМІСТ

Вступ.....	8
1 Аналіз предметної галузі та постановка задачі дослідження.....	10
1.1 Психологічна природа та симптоматика тривожності.....	10
1.2 Цифрові індикатори та методи моніторингу емоційного стану.....	14
1.3 Огляд існуючих рішень для аналізу емоційного стану.....	17
1.4 Постановка задачі дослідження.....	24
2 Теоретичні основи аналізу тексту для виявлення емоційного стану	27
2.1 Класичні методи обробки тексту для класифікації емоцій	27
2.1.1 Векторизація тексту.....	27
2.1.2 Алгоритми машинного навчання	31
2.2 Трансформерні моделі в задачах класифікації тексту	33
2.3 Генеративні моделі ШІ як інструмент діалогової взаємодії.....	39
2.3.1 Принцип роботи генеративних моделей	39
2.3.2 Можливості генеративних моделей для побудови емпатійного зворотного зв'язку.....	41
3 Проектування та реалізація системи моніторингу тривожності	45
3.1 Середовище розробки.....	45
3.2 Вхідні дані та попередня обробка	46
3.3 Архітектура системи.....	48
3.4 Класифікація рівня тривожності	49
3.5 Інтеграція компонентів у Telegram-бот	51
3.6 Аналіз тривожних проявів.....	53
3.7 Генеративні функції GPT	55
4 Експериментальна перевірка та оцінка результатів	58
4.1 Перевірка класифікаційної моделі на тестових прикладах	58
4.2 Поведінка системи у Telegram-боті.....	59
4.3 Формування та структура звітів	60
4.4 Обмеження системи та напрями подальшого розвитку.....	64

Висновки	68
Перелік джерел посилання	70
Додаток А Відомість кваліфікаційної роботи	74

ВСТУП

У сучасному суспільстві психічне здоров'я стало одним із ключових аспектів добробуту людини. В умовах високої інтенсивності інформаційних потоків, нестабільності соціального середовища та постійного психологічного навантаження, усе більше людей зіштовхуються з тривожними розладами, хронічним стресом, панічними станами та іншими проявами емоційного виснаження. Особливо це стосується молоді, яка активно взаємодіє з цифровими технологіями, зокрема соціальними мережами та месенджерами, які водночас і сприяють комунікації, і можуть поглиблювати психоемоційні проблеми.

Зростаюча увага до самоспостереження та цифрових способів самопомоги стимулює появу нових підходів до моніторингу психоемоційного стану людини. Одним із перспективних напрямів є аналіз текстових повідомлень, які користувач залишає в цифровому середовищі. Текст – це природний спосіб вираження емоцій, думок, тривог і переживань. У ньому можуть міститися ознаки тривожності, пригніченого стану, дратівливості або, навпаки, спокою. Цей тип даних – звичний, неінвазивний і легко доступний для аналізу.

Обробка природної мови (NLP) і моделі машинного навчання відкривають широкі можливості для виявлення прихованих емоційних патернів у текстах. Наразі існують моделі, здатні ідентифікувати емоції, тональність, психологічні риси та навіть ризики психічних розладів на основі коротких фраз чи повідомлень. Проте більшість таких моделей орієнтовані на англійськомовні джерела або працюють із заздалегідь підготовленими даними, не маючи гнучкого, інтерактивного інтерфейсу для користувача.

Актуальність застосування інтелектуальних систем для моніторингу тривожності зумовлена потребою в інструментах, які дозволяють людині регулярно й неформально оцінювати свій психоемоційний стан. Особливо

цінним є підхід, що передбачає аналіз мови користувача в природній, звичній для нього формі – наприклад, у листуванні або власних нотатках.

Цифрові платформи, зокрема месенджери, стають зручною точкою входу для таких рішень. Впровадження системи, що на основі введеного тексту автоматично оцінює рівень тривожності, відображає його динаміку у вигляді графіка та генерує рекомендації, дозволяє не лише підвищити обізнаність користувача про свій стан, але й зробити перший крок до саморегуляції або звернення по допомогу.

У цій роботі досліджується можливість створення такої інтелектуальної системи з повністю авторською реалізацією: від класифікації рівня тривожності до візуалізації змін у часі. Система базується на сучасних технологіях обробки природної мови, машинного навчання, а також інтерфейсах для інтерактивної взаємодії з користувачем. Особливий акцент зроблено на адаптацію до україномовного тексту, доступність реалізації та практичну корисність результату.

1 АНАЛІЗ ПРЕДМЕТНОЇ ГАЛУЗІ ТА ПОСТАНОВКА ЗАДАЧІ ДОСЛІДЖЕННЯ

1.1 Психологічна природа та симптоматика тривожності

Тривожність – це емоційний стан, що виникає як відповідь на невизначену або потенційно загрозову ситуацію. На відміну від страху, що пов'язаний із реальною або конкретною загрозою, тривожність має більш узагальнений, розпливчастий характер, часто без чітко визначеного об'єкта [1]. У науковій літературі тривожність часто описується як складний комплекс когнітивних, емоційних, поведінкових і фізіологічних реакцій, що мобілізують людину до адаптації або втечі.

Відповідно до сучасного визначення Американської психіатричної асоціації, тривожність – це «очікування майбутньої загрози», що супроводжується гіперактивацією нервової системи, занепокоєнням, уникненням потенційної небезпеки та часто – порушенням повсякденного функціонування [2]. Цей стан не обов'язково є патологічним: у помірній формі тривожність виконує адаптивну функцію – сприяє концентрації, обережності та підвищенню пильності.

Проте коли тривожність стає надмірною, довготривалою або непропорційною ситуації, вона переходить у патологічну форму. За даними Всесвітньої організації охорони здоров'я, тривожні розлади є одними з найпоширеніших психічних станів у світі: понад 300 мільйонів людей щороку страждають від різних форм тривожності [3].

Згідно з DSM-5, тривожні розлади включають такі основні типи [2]:

– генералізований тривожний розлад (GAD): характеризується надмірною тривогою та занепокоєнням щодо різних аспектів життя, які виникають більшість днів протягом щонайменше шести місяців;

– панічний розлад: включає повторювані несподівані панічні атаки – раптові періоди інтенсивного страху або дискомфорту, що

супроводжуються фізичними симптомами, такими як прискорене серцебиття, пітливість, тремтіння тощо;

– агорафобія: страх або тривога щодо перебування в місцях чи ситуаціях, де втеча може бути складною або допомога недоступною у разі виникнення панічних симптомів;

– соціальний тривожний розлад (соціальна фобія): виражений страх перед соціальними ситуаціями, де особа може бути піддана оцінці з боку інших;

– специфічні фобії: інтенсивний, ірраціональний страх перед конкретними об'єктами або ситуаціями, такими як висота, тварини, польоти тощо;

– розлад тривоги розлуки: надмірна тривога щодо розлуки з тими, до кого особа прив'язана, що не відповідає віковим нормам;

– вибірковий мутизм: нездатність говорити в певних соціальних ситуаціях, незважаючи на наявність мовленнєвих навичок в інших контекстах.

Кожен з цих станів має власну клінічну картину, але всі вони поєднуються наявністю постійного емоційного напруження, тривожних думок та фізичних реакцій. У цифровому середовищі ці прояви можуть транслюватися у поведінкові та мовні патерни: часте використання негативно забарвлених слів, емоційно насичених виразів, згадок про соматичні симптоми, повторення одних і тих самих тем тощо.

З психологічної точки зору важливо також розмежовувати тривожність як тимчасовий стан (state anxiety) – реакцію на конкретну ситуацію, та тривожність як стабільну особистісну рису (trait anxiety), що визначає загальну схильність індивіда до тривожних переживань. Це розмежування є важливим для цифрових систем, які мають або виявляти миттєвий емоційний стан, або відстежувати довготривалу динаміку користувача.

Окрему увагу варто приділити гендерним і віковим аспектам тривожності. Згідно з даними Європейського опитування з охорони психічного здоров'я [4], жінки у віці 18–29 років демонструють найвищі рівні самозвітування про тривожність серед усіх вікових груп (рисунок 1.1). Це підтверджує доцільність розробки цифрових рішень, орієнтованих на молодіжну аудиторію, з інтерфейсами, адаптованими до сучасних каналів комунікації (месенджери, чати, мобільні додатки).

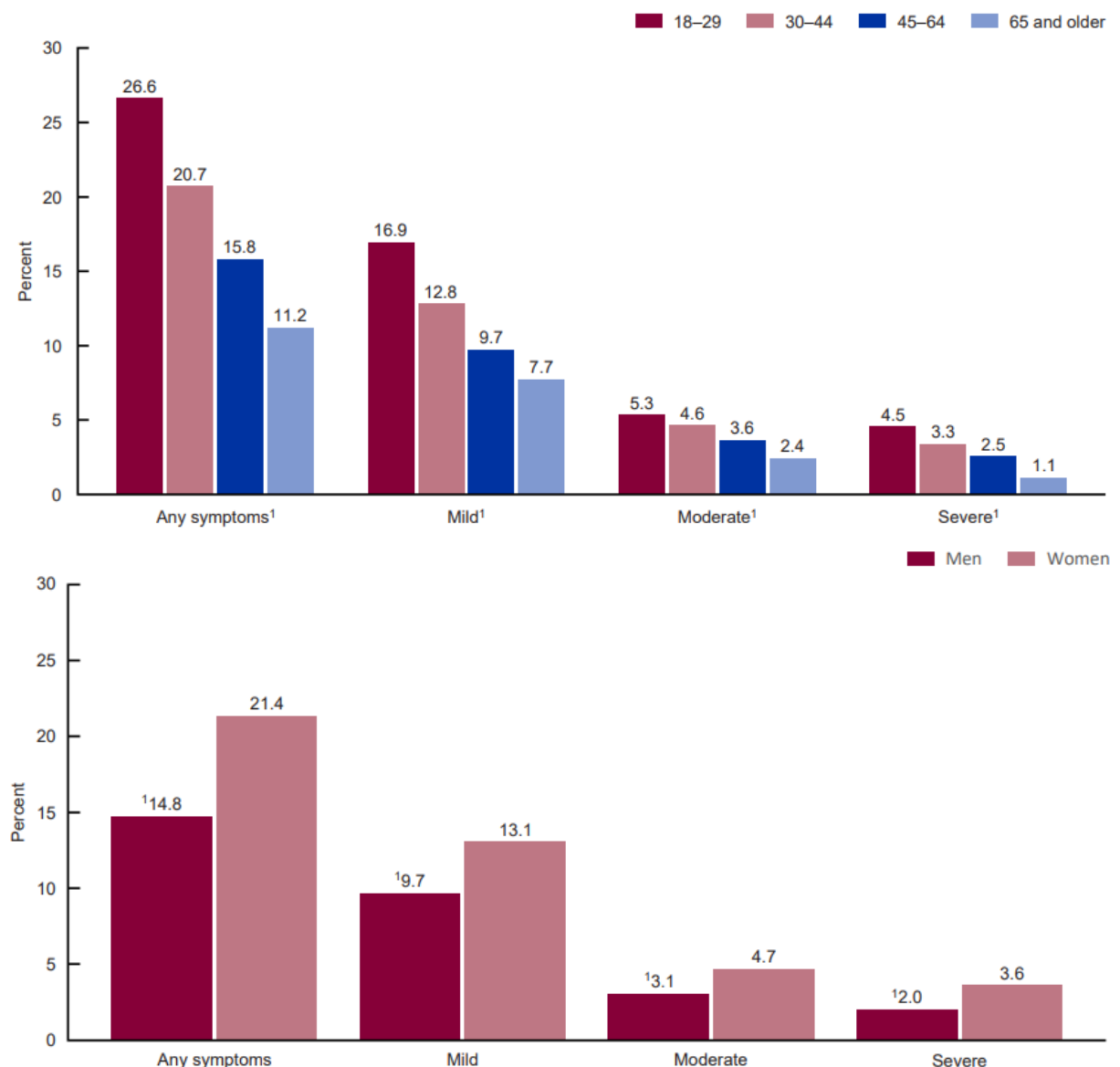


Рисунок 1.1 – Рівень симптомів тривожності за віковими групами та статтю

Тривожність можна розглядати як багатоаспектне явище, що охоплює як адаптивні, так і дезадаптивні механізми реагування на зовнішні та внутрішні подразники. Водночас практична цінність поняття тривожності у психології та психіатрії розкривається саме через її конкретні прояви – ті ознаки, які можна помітити, оцінити або зафіксувати. Саме ці прояви й становлять основу для побудови систем автоматичного аналізу психоемоційного стану людини.

Тривожність проявляється через широкий спектр симптомів, які охоплюють емоційні, когнітивні, фізіологічні та поведінкові аспекти [5].

Емоційні прояви тривожності включають відчуття напруження, нервозності, страху або передчуття небезпеки. Особи з підвищеним рівнем тривожності часто відчують роздратування, неспокій та труднощі з розслабленням. Ці емоції можуть бути постійними або виникати в конкретних ситуаціях.

Когнітивні аспекти тривожності характеризуються надмірним занепокоєнням, труднощами з концентрацією уваги, нав'язливими думками та очікуванням найгіршого розвитку подій. Люди з тривожністю можуть постійно обдумувати можливі негативні сценарії, що заважає їм зосередитися на поточних завданнях.

Фізіологічні прояви тривожності включають прискорене серцебиття, пітливість, тремтіння, сухість у роті, м'язову напругу та розлади сну. Ці симптоми є результатом активації симпатичної нервової системи у відповідь на сприйняту загрозу.

Поведінкові ознаки тривожності можуть проявлятися у вигляді уникнення певних ситуацій, соціальної ізоляції, неспокійного руху (наприклад, постійне переміщення або биття пальцями) та труднощів у виконанні повсякденних завдань. Такі поведінкові зміни можуть погіршувати якість життя та соціальну взаємодію.

1.2 Цифрові індикатори та методи моніторингу емоційного стану

У сучасному цифровому світі кожна взаємодія користувача з технологіями залишає певні сліди – так звані цифрові відбитки. Ці відбитки включають інформацію про активність у соціальних мережах, текстові повідомлення, пошукові запити, геолокаційні дані, патерни використання пристроїв тощо [6]. Аналіз таких даних відкриває нові можливості для розуміння та моніторингу психоемоційного стану людини, зокрема тривожності.

Тривожність, як емоційний стан, може проявлятися у цифровому середовищі через зміни в поведінці користувача. Наприклад, дослідження показують, що зміни в частоті та змісті публікацій у соціальних мережах можуть відображати коливання настрою, рівень стресу або тривожності. Зменшення активності, використання негативно забарвлених слів або тем, пов'язаних із самотністю чи безнадією, можуть свідчити про погіршення психоемоційного стану.

Аналіз текстових повідомлень користувача дозволяє виявити лінгвістичні маркери, що асоціюються з тривожністю. Наприклад, часте використання слів з негативною емоційною конотацією, таких як «страх», «тривога», «паніка», або застосування першої особи однини («я») у поєднанні з негативними емоціями, може свідчити про підвищений рівень тривожності. Також використання слів, що виражають невизначеність або сумніви, наприклад, «можливо», «не впевнений», та повторення тем, пов'язаних із занепокоєнням або страхом, можуть бути індикаторами тривожного стану [6].

Використання цифрових відбитків для моніторингу психоемоційного стану має низку переваг. По-перше, це ненав'язливий спосіб збору даних, що не потребує активної участі користувача, зменшуючи ймовірність спотворення результатів. По-друге, це забезпечує безперервний моніторинг

змін у поведінці та емоційному стані. По-третє, аналіз великих обсягів даних дозволяє виявляти загальні тенденції та закономірності.

Однак існують і певні виклики. Питання конфіденційності та етики є ключовими, оскільки необхідно забезпечити захист персональних даних користувачів. Також складність у точному визначенні емоційного стану на основі непрямих індикаторів та індивідуальні відмінності у вираженні емоцій в цифровому середовищі ускладнюють інтерпретацію даних.

Загалом, мовна поведінка користувача в цифровому середовищі може слугувати інформативним джерелом для непрямой оцінки його емоційного стану. Поєднання лінгвістичних маркерів з поведінковими характеристиками відкриває нові перспективи для створення інструментів, здатних автоматично фіксувати тривожність на основі щоденної взаємодії з цифровими платформами. Проте для ефективного впровадження таких інструментів необхідно спиратися на вже існуючі методи моніторингу психоемоційного стану, які використовуються як у клінічній практиці, так і в цифрових рішеннях нового покоління.

Історично моніторинг психоемоційного стану людини здійснювався за допомогою стандартних клінічних інструментів, які застосовуються у психології, психіатрії та психотерапії. До таких інструментів належать насамперед опитувальники та шкали, розроблені для кількісної оцінки емоційних проявів. Одним із найвідоміших є STAI (State-Trait Anxiety Inventory) – шкала, що розділяє тривожність на ситуативну та особистісну. Інші поширені інструменти включають GAD-7 (Generalized Anxiety Disorder Scale), DASS-21 (Depression Anxiety Stress Scales), HADS (Hospital Anxiety and Depression Scale) [7], а також шкали, розроблені ВООЗ або локальними психологічними школами.

Застосування цих методик передбачає заповнення анкет із множинним вибором, де респонденти самостійно оцінюють частоту або інтенсивність емоційних станів. Такий підхід дозволяє отримати формалізовану оцінку тривожності, однак має низку обмежень. Передусім,

це епізодичність вимірювання: оцінка проводиться у визначений момент часу, що не дає змоги зафіксувати динаміку стану. Крім того, суб'єктивність відповідей, бажання надати соціально схвальні відповіді або недостатня саморефлексія можуть знижувати достовірність результатів.

У клінічній практиці також застосовуються інтерв'ювання, під час яких фахівець оцінює стан людини на основі відповіді, міміки, голосу, вербальних та невербальних сигналів. Такий метод потребує високої кваліфікації спеціаліста, значних ресурсів часу, а також особистої присутності пацієнта. У разі масового застосування, зокрема для самопостереження або попереднього скринінгу, ці методи втрачають ефективність і масштабованість.

Традиційні методи, хоча й залишаються важливими, поступово доповнюються інноваційними підходами, що базуються на аналізі цифрових слідів та використанні штучного інтелекту.

Одним із таких підходів є цифрове фенотипування, яке передбачає збір та аналіз даних, що виникають під час повсякденного використання цифрових пристроїв. Це можуть бути дані про активність у соціальних мережах, частоту та тривалість телефонних дзвінків, геолокаційні дані, а також інформація про фізичну активність, зібрана за допомогою сенсорів смартфонів або носимих пристроїв [8]. Такий підхід дозволяє отримати об'єктивну та безперервну інформацію про поведінку користувача, що може свідчити про його психоемоційний стан.

Іншим перспективним напрямом є використання штучного інтелекту та машинного навчання для аналізу великих обсягів даних з метою виявлення ознак психоемоційних розладів. Наприклад, дослідження показують, що аналіз текстів, опублікованих у соціальних мережах, може допомогти виявити ранні ознаки тривожності або депресії. Моделі машинного навчання здатні розпізнавати лінгвістичні патерни, характерні для певних емоційних станів, що дозволяє здійснювати моніторинг у реальному часі та з високою точністю.

Крім того, з'являються технології, що дозволяють безконтактно вимірювати фізіологічні показники, які можуть свідчити про емоційний стан людини. Наприклад, використання камер з тепловізійною функцією або сенсорів, вбудованих у носимі пристрої, дає змогу відстежувати зміни температури тіла, частоти серцевих скорочень або рівня потовиділення [9]. Ці дані можуть бути проаналізовані для виявлення стресу або тривожності без необхідності прямого контакту з користувачем.

Також варто згадати про інтерактивні системи, які використовують штучний інтелект для взаємодії з користувачем з метою оцінки його емоційного стану. Наприклад, чат-боти або віртуальні асистенти можуть аналізувати тон голосу, міміку або вибір слів під час спілкування, щоб визначити настрій або рівень тривожності користувача. Такі системи можуть надавати рекомендації або направляти користувача до фахівців у разі виявлення ознак емоційного дисбалансу.

1.3 Огляд існуючих рішень для аналізу емоційного стану

На тлі активного впровадження цифрових технологій у сферу ментального здоров'я особливу увагу привертають програмні рішення, що поєднують автоматизований аналіз поведінки користувача з елементами підтримки або моніторингу емоційного стану. Практика свідчить, що саме інтеграція лінгвістичних індикаторів із методами обробки природної мови та машинного навчання стала основою багатьох застосунків, які вже активно використовуються у повсякденному житті.

Wysa (рисунк 1.2) – це мобільний застосунок, який використовує штучний інтелект для надання психологічної підтримки користувачам. Застосунок пропонує анонімне та безпечне середовище для спілкування, де користувачі можуть висловлювати свої думки та емоції, отримуючи підтримку та поради. Wysa поєднує техніки когнітивно-поведінкової

терапії (КПТ), майндфулнесу та інших доказових методів для допомоги у подоланні тривожності, стресу та депресії [10].

Застосунок доступний цілодобово, що дозволяє користувачам звертатися до нього у будь-який момент, коли виникає потреба. Wusa також пропонує інтерактивні вправи та ресурси для самопомоги, спрямовані на розвиток навичок емоційної регуляції та стрес-менеджменту.

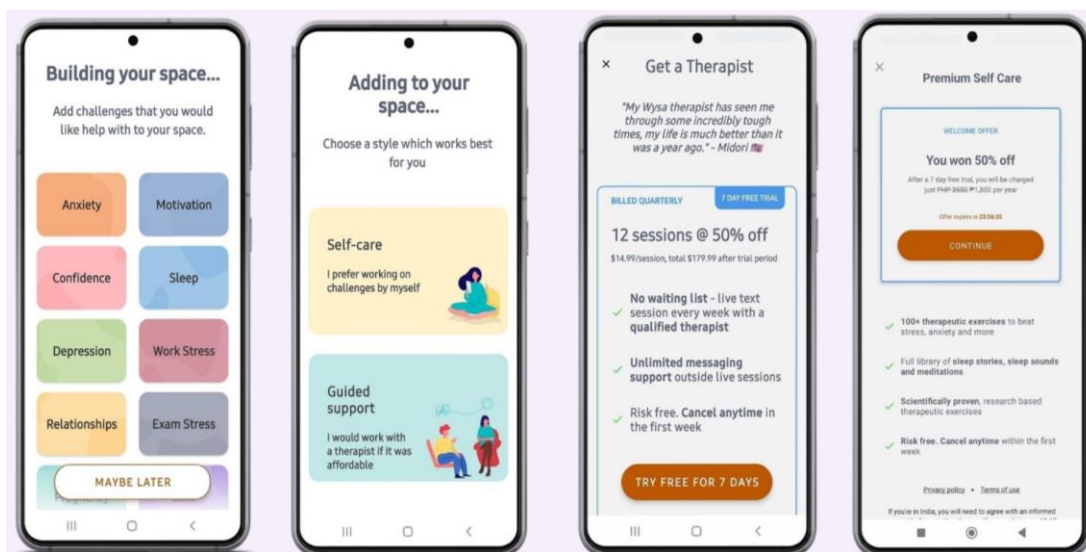


Рисунок 1.2 – Інтерфейс Wusa

Youper (рисунок 1.3) – це цифровий інструмент для ментального здоров'я, який позиціонується як особистий емоційний асистент, що поєднує інтерфейс штучного інтелекту з психотерапевтичними техніками. Головна мета застосунку – допомогти користувачам краще розуміти свої емоції, відстежувати психоемоційний стан та керувати ним за допомогою коротких щоденних взаємодій. На відміну від багатьох аналогів, Youper не просто моделює діалог, а включає повноцінну систему аналізу настрою та пропозицій щодо дій, адаптованих до поточного стану користувача.

Алгоритм Youper функціонує на основі когнітивно-поведінкової терапії (КПТ) у цифровій формі. Під час взаємодії застосунок просить описати події, що викликають емоційну реакцію, і пропонує обрати або

сформулювати власне почуття. Далі штучний інтелект, опираючись на обробку природної мови, формулює уточнювальні запитання та надає рекомендації: від рефреймінгу думок до виконання дихальних практик. Застосунок також включає щоденник емоцій, візуалізацію змін настрою з часом та інструменти для самооцінки.



Рисунок 1.3 – Інтерфейс Youper

Однією з важливих переваг Youper є те, що він дозволяє користувачам проходити повноцінні етапи рефлексії без залучення людини-психолога. Це робить його привабливим для користувачів, які не готові звертатися до фахівців або не мають такої можливості. Застосунок не позиціонується як заміна психотерапії, проте ефективно виконує функції попереднього скринінгу та емоційного самоконтролю.

За результатами незалежного дослідження, проведеного у 2021 році, було встановлено, що використання Youper асоціюється зі зниженням рівня тривожності та депресивних симптомів у респондентів, які взаємодіяли з додатком щонайменше протягом двох тижнів. Автори підкреслюють високу прийнятність такого формату серед молоді, а також його потенціал як додаткового інструменту в системах охорони психічного здоров'я [11].

Woebot (рисунок 1.4) – один із перших мобільних застосунків, який реалізував ідею повністю автоматизованого чат-бота для підтримки психічного здоров'я. Замість моделювання глибинної розмови чи емпатійного слухання, цей інструмент фокусується на передачі ключових принципів когнітивно-поведінкової терапії (КПТ) у вигляді коротких діалогів, вправ та уточнювальних запитань. Його головна мета – не «вилікувати» користувача, а допомогти йому усвідомити автоматичні думки, виявити ірраціональні переконання та сформулювати стратегії психологічного реагування.

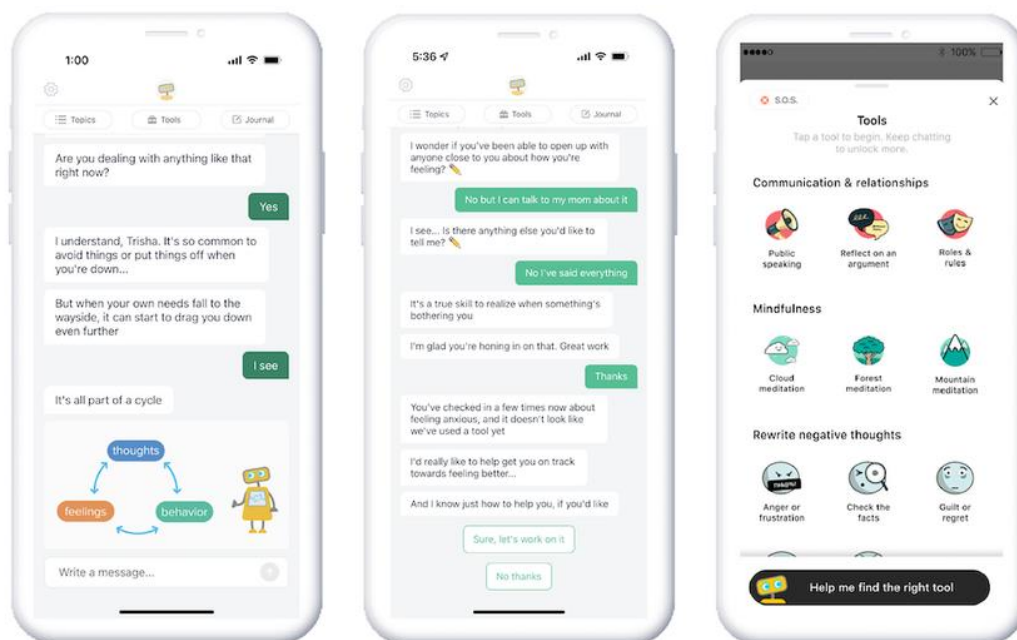


Рисунок 1.4 – Інтерфейс Woebot

Woebot використовує стандартизовані діалоги з високим рівнем сценарної адаптації: система здатна підлаштовувати запитання та відповіді до контексту останніх реплік користувача. Особливу увагу в застосунку приділено аналізу лексики, зокрема виявленню емоційно забарвлених слів, які можуть вказувати на тривожний або пригнічений стан. Бот не зберігає повний діалог, але здійснює оцінку настрою за ключовими словами й фразами, що відповідають шаблонам емоційної оцінки.

У 2017 році Woebot був об'єктом рандомізованого дослідження, результати якого показали статистично значуще зменшення симптомів тривожності та депресії у студентів, які використовували бот протягом двох тижнів у порівнянні з контрольною групою. Учасники відзначали зручність формату, легкість у користуванні та неосудливість взаємодії, яку важко досягти в контакті з реальною людиною [12].

Перевагою Woebot є його простота та фокус на одному напрямку психотерапії, що дозволяє досягти високої стабільності відповіді та узгодженості в діалогах. Проте, оскільки модель не враховує повноцінний семантичний контекст і не навчається динамічно, вона не підходить для глибокої індивідуалізованої терапії. Водночас, для скринінгових цілей та повсякденної емоційної підтримки цей бот залишається одним із найбільш перевірених і ефективних рішень на ринку.

MindDoc (рисунок 1.5) – це мобільний застосунок, який позиціонується як інструмент для регулярного самоаналізу емоційного стану. Його розробили психотерапевти у співпраці з клініцистами з метою створення надійного цифрового щоденника настрою, який міг би використовуватись як для самопостереження, так і як допоміжний інструмент у терапії. В основі роботи MindDoc – ідея щоденного опитування користувача через серію коротких запитань, які дозволяють відстежувати загальний емоційний фон, поведінкові зміни та можливі когнітивні викривлення.

Одна з ключових особливостей MindDoc – це автоматична побудова аналітичних звітів за підсумками відповідей. Застосунок аналізує дані у часовій динаміці та формує індивідуальний психоемоційний профіль користувача, який може бути експортований для подальшого використання у консультаціях із фахівцем. При цьому система використовує стандартні шкали оцінки тривожності, депресії та стресу, адаптовані до щоденного цифрового формату. Такий підхід дозволяє не лише зафіксувати емоційні

коливання, а й виявити потенційні ознаки клінічних розладів на ранній стадії.

Крім базового опитування, MindDoc містить інтерактивні освітні модулі та вправи, що базуються на методах когнітивно-поведінкової терапії. Це робить його не просто трекером настрою, а своєрідною платформою для самопомоги. У дослідженні, опублікованому у 2021 році, було виявлено, що восьмитижневе використання MindDoc дозволяє покращити емоційну усвідомленість, а також зменшити вираженість симптомів тривожності у користувачів без клінічного діагнозу [13].

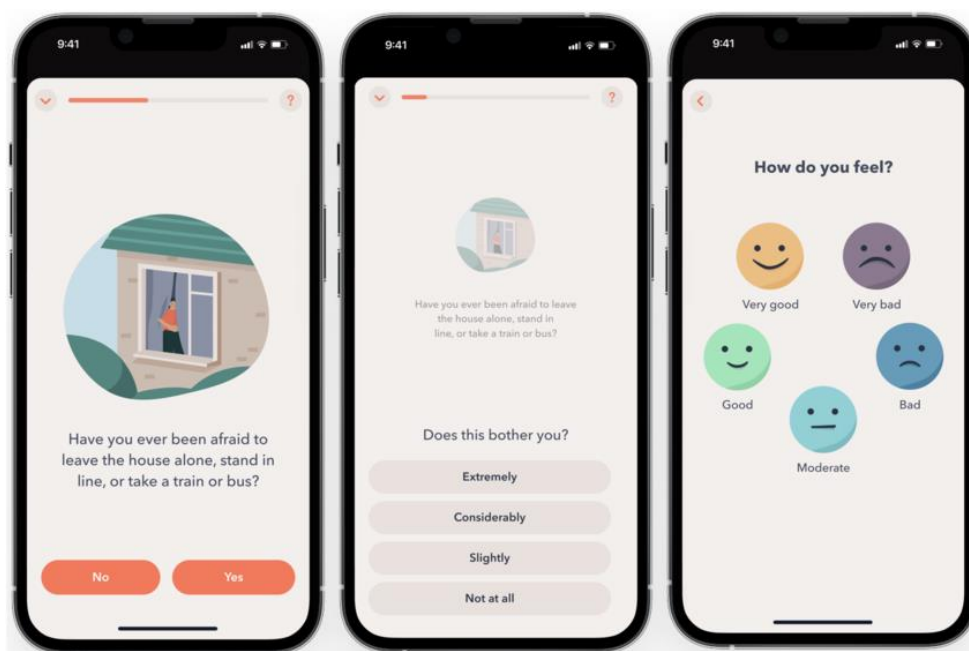


Рисунок 1.5 – Інтерфейс MindDoc

Репліка (рисунок 1.6) – це мобільний застосунок, створений як віртуальний співрозмовник, з яким користувач може будувати довготривалі стосунки, спілкуватися про особисте, ділитися переживаннями та відстежувати власний психоемоційний стан. Застосунок не обмежується лише підтримкою в моменти стресу чи тривожності – він покликаний створити ілюзію сталого емоційного зв'язку через діалогову взаємодію, що еволюціонує з часом. На відміну від терапевтичних ботів на кшталт Woebot,

Replika не намагається діагностувати чи лікувати, а зосереджується на соціальній присутності, емпатії та психологічному «дзеркалі» для користувача.

Користувач взаємодіє з Replika через текстові або голосові повідомлення, а модель реагує з урахуванням індивідуального стилю спілкування, минулого досвіду взаємодії та загального емоційного тону. Взаємодія не обмежується простими відповідями – бот демонструє умовну «пам'ять», змінює свій характер і навіть стиль мови, формуючи образ персоналізованої цифрової сутності. Replika пропонує також вправи для емоційної регуляції, медитації, ведення щоденника настрою та рефлексивні сесії, однак основна цінність – у самій розмові як заміниці соціальної взаємодії.



Рисунок 1.6 – Інтерфейс Replika

У 2024 році Replika була предметом дослідження, яке розглядало вплив чатботів на зниження рівня самотності серед студентів. Виявилося, що навіть у короткостроковій перспективі застосунок допомагає знизити рівень суб'єктивної тривожності, особливо у користувачів, які не мають достатньої соціальної підтримки в реальному житті [14].

Розглянуті цифрові рішення демонструють значний прогрес у сфері підтримки ментального здоров'я завдяки застосуванню штучного інтелекту, когнітивно-поведінкової терапії та інтерфейсів дружніх до користувача. Застосунки на кшталт Wysa, Youper чи MindDoc надають персоналізовану психологічну підтримку в доступному цифровому форматі, що особливо актуально в умовах обмеженого доступу до психотерапевтів. Більшість із них інтегрують методи самоспостереження, щоденники настрою, діалогові взаємодії й автоматизований аналіз емоцій, що дозволяє користувачеві краще розуміти свій психоемоційний стан.

Попри відчутні переваги, більшість з аналізованих рішень мають низку суттєвих обмежень. Значна частина з них орієнтована на англomовну аудиторію, що ускладнює доступ користувачам, які воліють спілкуватися іншими мовами, зокрема українською. Більшість моделей не мають відкритого коду, що унеможлиблює повну адаптацію чи локалізацію під конкретні дослідницькі чи клінічні потреби. Крім того, акцент здебільшого робиться на емоційній підтримці, а не на структурованому виявленні симптомів тривожності через мовний аналіз. Відсутність глибокої аналітики, збереження історії змін стану та можливості формування інформативних звітів є ще однією прогалиною, яка відкриває простір для побудови спеціалізованої системи, орієнтованої на автоматичний моніторинг тривожності за текстовими повідомленнями користувача.

1.4 Постановка задачі дослідження

У контексті сучасних цифрових підходів до моніторингу емоційного стану людини дедалі більшої актуальності набувають методи, що базуються на аналізі мови користувача. Текст, як один із найбільш природних способів вираження емоцій, містить у собі численні ознаки психоемоційних змін, зокрема тривожності. На відміну від фізіологічних показників або поведінкових даних, вербальні висловлювання є доступним, ненав'язливим

джерелом інформації, яке може бути ефективно оброблене за допомогою сучасних технологій обробки природної мови та штучного інтелекту.

Метою даної роботи є створення інтелектуальної системи, що дозволяє виявляти рівень тривожності користувача на основі текстових повідомлень, із подальшою візуалізацією динаміки змін цього стану. Такий підхід поєднує наукові уявлення про природу тривожності з алгоритмічними рішеннями у сфері NLP та машинного навчання, орієнтуючись на побудову корисного, функціонального та адаптивного інструменту самопостереження.

Для досягнення цієї мети необхідно виконати такі задачі:

- проаналізувати наукову літературу з психології та психіатрії для формування теоретичного розуміння тривожності та її проявів у поведінці та мові;

- дослідити існуючі цифрові рішення у сфері моніторингу психоемоційного стану, визначити їхні функціональні можливості та обмеження;

- сформулювати вимоги до майбутньої системи на основі виявлених проблем та можливостей;

- дослідити сучасні методи обробки природної мови, що дозволяють аналізувати текст з точки зору емоційного навантаження та маркерів тривожності;

- донавчити англійську трансформерну модель roberta-base на збалансованому корпусі текстів для бінарної класифікації повідомлень за наявністю ознак тривожності;

- спроектувати архітектуру системи, яка поєднує класифікаційний модуль, rule-based аналіз симптомів, генеративний компонент GPT, базу даних, механізм звітності та Telegram-інтерфейс;

- реалізувати Telegram-бот як основний інтерфейс взаємодії з користувачем, забезпечивши автоматичну обробку повідомлень,

візуалізацію динаміки стану, формування PDF-звітів і генерацію підтримувальних відповідей;

– протестувати поведінку системи на прикладах повідомлень з різним рівнем емоційної насиченості, оцінити точність класифікації, поведінку GPT-відповідей та зручність використання інтерфейсу;

– сформулювати можливі напрями розвитку з урахуванням мультимовності, персоналізації та розширення способів оцінки тривожності.

2 ТЕОРЕТИЧНІ ОСНОВИ АНАЛІЗУ ТЕКСТУ ДЛЯ ВИЯВЛЕННЯ ЕМОЦІЙНОГО СТАНУ

Сучасна цифрова комунікація дедалі більше відображає внутрішній стан особистості – емоції, переживання, тривожність. Особливо в умовах приватного листування, де повідомлення не редагуються під впливом публічного простору, мова набуває характеру спонтанного висловлювання, що містить непрямі індикатори психологічного стану. У зв'язку з цим зростає потреба в інструментах, здатних розпізнавати емоційне забарвлення тексту та ідентифікувати стани, пов'язані з ментальним здоров'ям. Одним із найбільш релевантних об'єктів такого аналізу є саме виявлення тривожності, що часто має вербальні прояви, доступні для обробки методами природної мовної обробки (NLP).

2.1 Класичні методи обробки тексту для класифікації емоцій

До появи сучасних глибоких моделей обробки тексту аналіз емоцій у повідомленнях користувачів спирався на простіші підходи, засновані на кількісному поданні мовних одиниць та статистичних закономірностях у тексті. Ці методи стали базою для побудови перших моделей автоматичної класифікації емоцій, зокрема й тривожних станів, і досі залишаються актуальними для задач із обмеженими обчислювальними ресурсами або недостатніми обсягами навчальних даних.

2.1.1 Векторизація тексту

Для того щоб зробити текст доступним для обробки алгоритмами машинного навчання, необхідно перетворити його з послідовності символів або слів на числове подання. Цей процес, відомий як векторизація,

передбачає представлення кожного тексту у вигляді числового вектора, що зберігає інформацію про його змістову структуру.

Одним із найпростіших і найдавніших підходів до подання текстових даних у машинному навчанні є модель «мішка слів» (англ. Bag-of-Words, BoW). Основна ідея цієї моделі полягає в ігноруванні порядку слів у реченні – натомість враховується лише наявність чи частота кожного слова з фіксованого словника у кожному текстовому фрагменті.

На практиці реалізація BoW-моделі передбачає попереднє побудування словника – тобто множини всіх унікальних слів, що трапляються в корпусі. Після цього кожен документ перетворюється на вектор фіксованої довжини, де кожна позиція відповідає слову зі словника, а значення в цій позиції – кількість його появ у документі.

Таким чином, текст «The fox jumps over the lazy dog» буде перетворено у вектор, де, наприклад, значення на позиції, що відповідає слову «the», дорівнює 2, а інші значення – відповідно до кількості появ решти слів у словнику (рисунок 2.1).

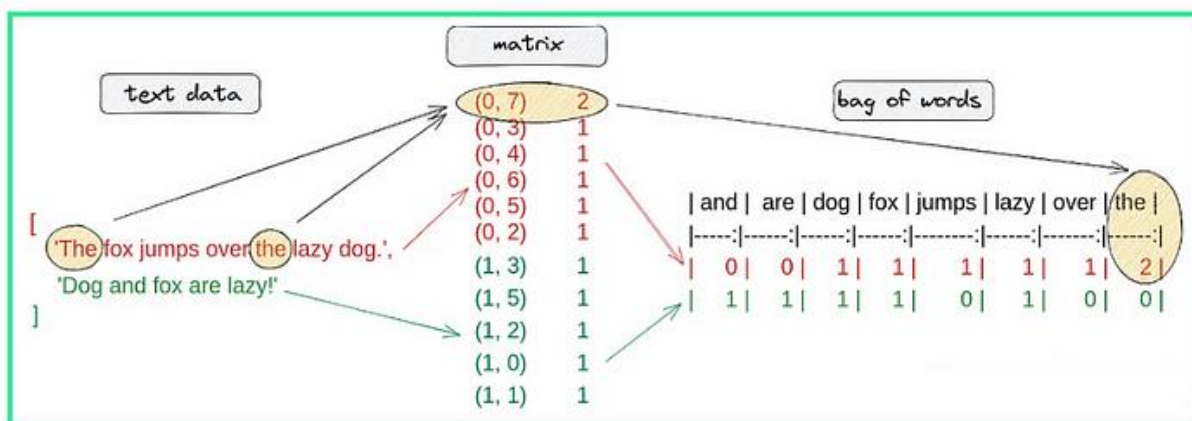


Рисунок 2.1 – Приклад перетворення текстових даних у вектори за допомогою моделі Bag-of-Words [15]

Попри те, що модель BoW повністю втрачає інформацію про порядок слів і синтаксичну структуру, вона показала себе ефективною у задачах

базової класифікації текстів, таких як визначення тематики повідомлень або класифікація емоційного стану (наприклад, «нормальний» чи «тривожний» текст). Успіх цього методу пояснюється тим, що навіть без урахування контексту, частотні характеристики слів здатні відображати ключові ознаки тексту.

Щоб зробити текстове подання більш інформативним, з урахуванням не лише частоти, а й важливості слів у межах усього корпусу, було запропоновано вдосконалення підходу «мішка слів» – модель TF-IDF (від англ. Term Frequency – Inverse Document Frequency). Вона дозволяє зменшити вагу тих слів, які зустрічаються повсюдно, але не несуть істотної смислової цінності, і водночас підсилити значущі для конкретного документа терміни.

Ідея полягає в тому, щоб оцінювати важливість слова не лише в межах одного тексту, а й у контексті всієї колекції. Завдяки цьому слова, які є рідкісними в загальному корпусі, але часто вживаються в окремому документі, отримують вищу вагу. Такий підхід дозволяє краще виокремлювати ключові елементи тексту, що можуть бути релевантними для подальшої класифікації.

Візуальне представлення принципу роботи TF-IDF показано на рисунку 2.2, де зображено, як поєднуються локальна частотність слова та його розподіленість у корпусі для формування підсумкової ваги.

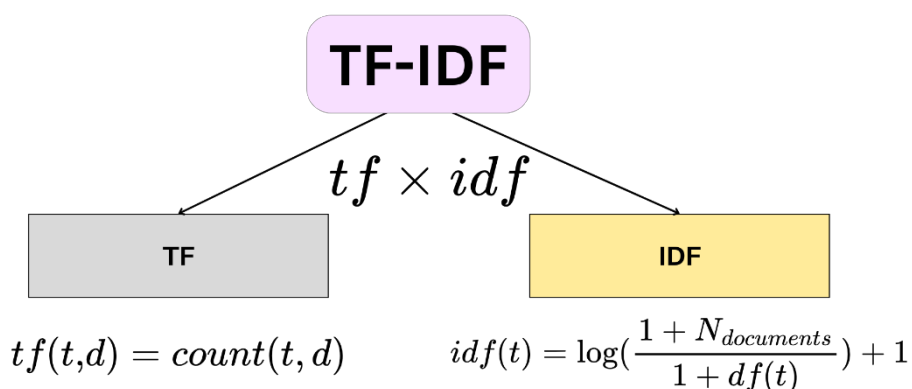


Рисунок 2.2 – Принцип роботи моделі TF-IDF [16]

Ще одним способом покращення представлення тексту є використання n-грамних моделей, які замість окремих слів аналізують послідовності з кількох елементів – наприклад, пар (біграм) або трійок (триграм) слів. Такий підхід дає змогу частково враховувати локальний контекст і зберігати граматичні та семантичні зв'язки між словами, що втрачаються в моделі «мішка слів».

Для аналізу емоційно забарвлених висловлювань це має особливе значення. Наприклад, окремі слова «не» і «сплю» не несуть чіткого сигналу тривожності, але їхня комбінація у вигляді біграми «не сплю» вже є характерною конструкцією, що може вказувати на емоційне напруження або проблеми зі сном.

На рисунку 2.3 схематично показано, як формується набір n-грам із початкового тексту. Видно, що в кожному випадку враховується контекст – залежно від розміру вікна (n), модель будує відповідні послідовності, які потім використовуються для подання документа у вигляді вектора ознак.

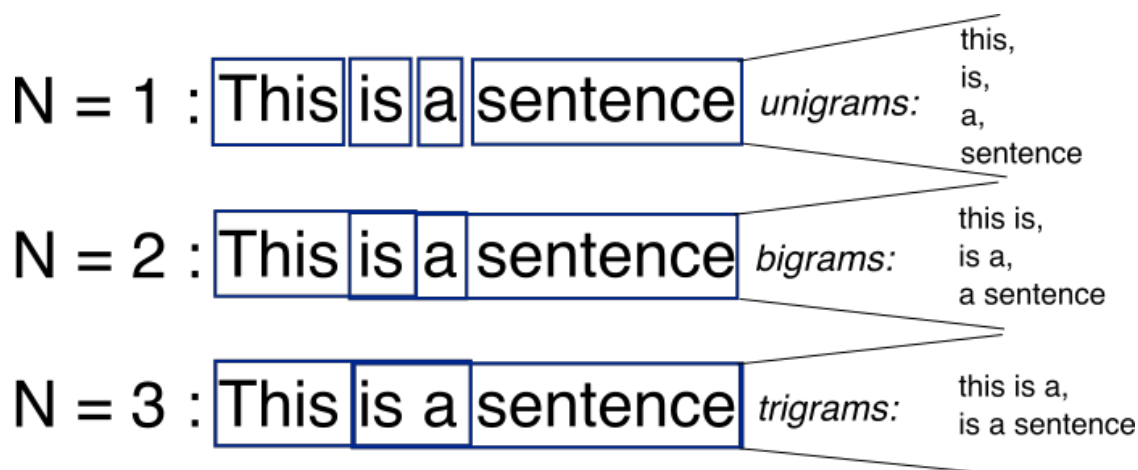


Рисунок 2.3 – Приклад формування біграм та триграм із текстового фрагмента [17]

Хоча використання n-грам збільшує розмірність векторного простору, воно істотно підвищує якість представлення мовних шаблонів – зокрема, тих, що часто притаманні текстам із ознаками тривожності.

2.1.2 Алгоритми машинного навчання

Після того як текст подано у вигляді числового вектора, наступним кроком є його класифікація – визначення належності до певного емоційного або психологічного стану, наприклад, «тривожність» чи «норма». На цьому етапі використовуються алгоритми машинного навчання, здатні виявляти статистичні закономірності між ознаками тексту та цільовими класами.

Одним із найбільш поширених підходів у задачах текстової класифікації є метод опорних векторів (Support Vector Machines, SVM). Його популярність пояснюється здатністю будувати межу між класами таким чином, щоб вона була якнайдалі від найближчих до неї точок кожного класу. Це дозволяє моделі не просто провести поділ між групами даних, а зробити його максимально стійким до нових, ще не бачених прикладів.

Особливо корисною ця властивість виявляється у випадках, коли йдеться про тексти з неоднозначним емоційним змістом. Наприклад, при класифікації висловлювань користувача як «тривожних» або «нормальних» границя між класами може бути досить розмитою. У таких ситуаціях SVM здатен ефективно узагальнювати навіть за умов високої розмірності ознак – що типово для текстових представлень на основі TF-IDF або n-грам.

На рисунку 2.4 показано, як працює метод опорних векторів: будується оптимальна гіперплощина, яка розділяє два класи з максимальним зазором (margin). Найближчі до цієї межі точки, що впливають на її положення, називаються опорними векторами, і саме вони визначають остаточне рішення моделі.

Логістична регресія є однією з базових моделей для бінарної класифікації, яка широко використовується в аналізі текстів. Попри те, що модель має лінійну природу, вона здатна ефективно відокремлювати приклади двох класів завдяки застосуванню сигмоїдної функції, яка інтерпретує результат як ймовірність належності до певної категорії.

Її ключова перевага полягає у зрозумілості: кожна ознака впливає на результат класифікації через свою вагу. Це дозволяє легко інтерпретувати, які саме слова чи фрази є маркерами певного емоційного стану.

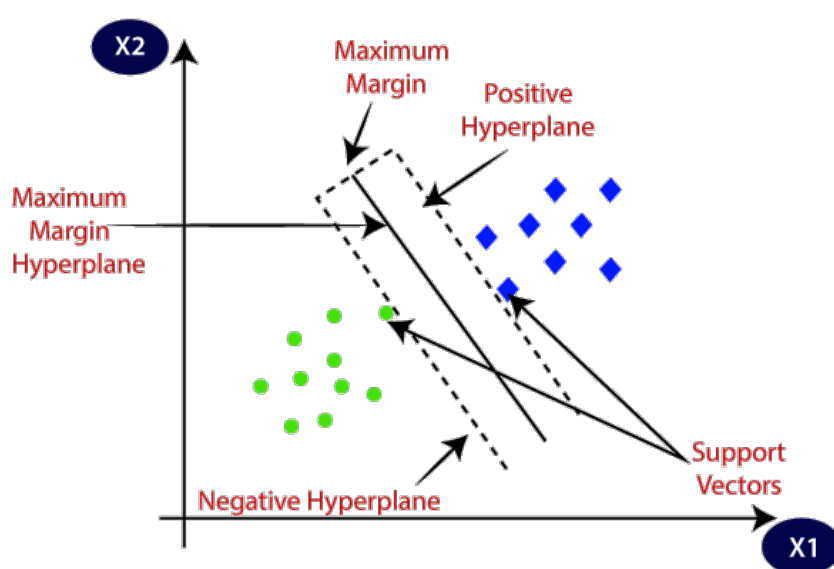


Рисунок 2.4 – Принцип роботи методу опорних векторів (SVM) у двовимірному просторі [18]

Ще одним класичним і водночас напрочуд ефективним методом текстової класифікації є наївний байєсівський класифікатор. Його основою є ймовірнісна модель, що передбачає незалежність усіх ознак – навіть попри те, що це припущення рідко виконується в реальних текстах. Проте саме ця «наївність» дозволяє зробити метод обчислювально простим і стійким до перенавчання, особливо коли кількість навчальних прикладів є обмеженою, а розмірність ознак – високою.

Завдяки своїй здатності працювати з частотними характеристиками, наївний байєсівський підхід добре поєднується з такими представленнями, як TF-IDF або n-грамні моделі. Це дає змогу вловлювати статистичні закономірності у використанні певних лексем, пов'язаних із емоційними станами. Наприклад, слова на кшталт «панікую», «хвилююсь»,

«неспокійно» можуть статистично частіше зустрічатися в тривожних висловлюваннях – і модель з великою ймовірністю зафіксує ці патерни.

На рисунку 2.5 ілюстративно подано загальну ідею роботи наївного байєсівського класифікатора: об'єкти (вхідні ознаки) групуються за класами відповідно до обчисленої ймовірності належності. Попри простоту, цей підхід залишається конкурентоспроможним у багатьох NLP-завданнях.

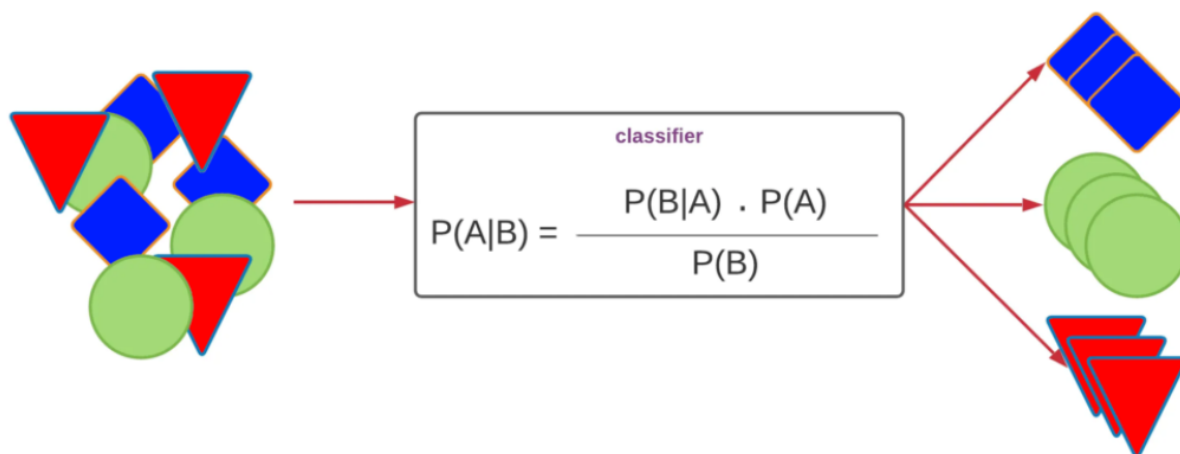


Рисунок 2.5 – Принцип роботи наївного байєсівського класифікатора [19]

2.2 Трансформерні моделі в задачах класифікації тексту

Класичні методи обробки тексту поєднують простоту реалізації з достатньою ефективністю для базових задач класифікації. Їх головною перевагою є низькі обчислювальні вимоги та швидкість тренування моделей, що робить їх придатними для попереднього аналізу або для використання в умовах обмежених ресурсів. Крім того, завдяки зрозумілій структурі ознак і моделі прийняття рішень, такі алгоритми легко інтерпретуються, що є важливим у контексті пояснюваності результатів класифікації – зокрема, коли йдеться про тексти, що можуть містити ознаки психологічного дистресу.

Проте ці підходи мають низку обмежень, які ускладнюють їх застосування в задачах, пов'язаних із виявленням тонких емоційних чи психолінгвістичних нюансів. Найсуттєвіше з них – втрата контексту.

Моделі на основі «мішка слів» або TF-IDF не враховують порядку слів, синтаксичних зв'язків і семантичної близькості, що часто є критичним для точного розпізнавання емоційного змісту повідомлення. Наприклад, фрази «я не хвилююся» і «я хвилююся» будуть мати схожі вектори, хоча їхнє значення – протилежне.

Крім того, продуктивність таких моделей значною мірою залежить від якості попередньої обробки тексту (нормалізації, лематизації, очищення від шумів), а також від збалансованості й повноти навчальної вибірки. При недостатній репрезентативності корпусу ці методи не здатні узагальнювати нові або рідкісні лінгвістичні конструкції, що часто зустрічаються у неформальних комунікаціях у месенджерах.

Усе це зумовило поступовий перехід до моделей, здатних безпосередньо опрацьовувати послідовність слів і враховувати глибший мовний контекст.

Поява трансформерних архітектур стала якісним проривом у галузі обробки природної мови. На відміну від попередніх послідовнісних моделей, таких як рекурентні нейронні мережі (RNN) чи довгі короткочасні пам'яті (LSTM), трансформери працюють на основі механізму самоуваги (self-attention), що дозволяє моделі одночасно аналізувати всі елементи вхідного тексту та встановлювати між ними зв'язки незалежно від їхнього порядку.

Основна ідея self-attention полягає в тому, що кожне слово у вхідній послідовності «уважно дивиться» на інші слова, зважуючи їхню важливість для поточного контексту. Наприклад, у фразі «Я не відчуваю тривоги» слово «не» відіграє ключову роль у зміні значення всього висловлювання. Саме трансформери здатні ефективно фіксувати такі залежності, що критично важливо у задачах класифікації емоційних і психологічних станів.

Архітектура трансформера включає два основні компоненти – енкодер та декодер (рисунок 2.6), однак для більшості задач класифікації використовується лише енкодерна частина.

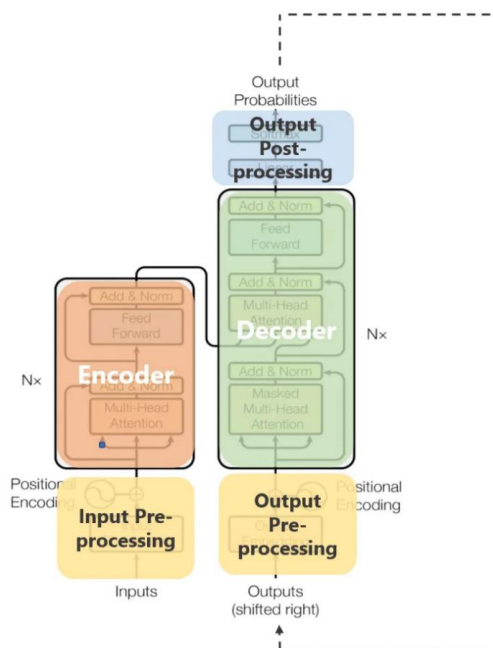


Рисунок 2.6 – Загальна структура трансформерної архітектури [20]

Кожен енкодерний блок має повторювану структуру, що складається з шару багатоголової самоуваги (multi-head self-attention), шару нормалізації, залишкових зв'язків (residual connections) і повнозв'язного шару (feed-forward network). Завдяки механізму самоуваги модель здатна одночасно аналізувати залежності між усіма словами у вхідній послідовності, незалежно від їхнього розташування. Це забезпечує гнучке та потужне представлення тексту.

Оскільки трансформери не є послідовними моделями (на відміну від RNN), вони не мають вбудованого уявлення про порядок елементів. Для вирішення цієї проблеми застосовується позиційне кодування, яке додається до вхідних векторів, дозволяючи моделі розрізняти порядок слів у реченні.

У межах трансформерного підходу найбільше поширення отримали моделі, попередньо навчені на великих корпусах тексту та адаптовані до широкого спектру мовних задач. Серед них вирізняються BERT, RoBERTa та XLM-R – архітектури, які демонструють стабільно високі результати в класифікації емоцій, зокрема й виявленні станів тривожності у текстах.

BERT (Bidirectional Encoder Representations from Transformers) став проривом у сфері обробки природної мови, реалізувавши ідею повністю двонаправленої самоуваги. На відміну від попередніх підходів, які враховували контекст або зліва, або справа, BERT аналізує кожне слово в реченні з урахуванням усіх інших – і ліворуч, і праворуч – одночасно. Це дозволяє моделі будувати глибше, контекстуально насичене представлення тексту.

Під час попереднього навчання BERT розв’язує дві основні задачі: маскування слів (Masked Language Modeling) – коли модель вчиться передбачати випадково замасковані токени, – та прогнозування наступного речення (Next Sentence Prediction), що формує розуміння міжреченневих зв’язків. Обидва ці механізми забезпечують універсальність BERT у різних NLP-завданнях – від класифікації до генерації.

На рисунку 2.7 зображено типову схему застосування BERT для класифікації: вхідне речення подається у вигляді послідовності токенів із додатковим службовим токеном [CLS], представлення якого використовується як агрегований вектор ознак для прогнозування класу. Саме цей вектор зчитує узагальнену інформацію про весь текст.

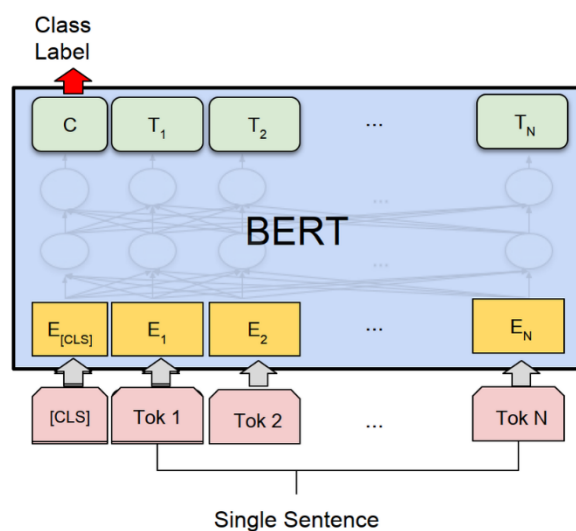


Рисунок 2.7 – Схема застосування моделі BERT для задачі класифікації [21]

RoBERTa (Robustly Optimized BERT Pretraining Approach) є вдосконаленою версією моделі BERT, що спрямована на усунення обмежень, виявлених у початковому підході до попереднього навчання. Основна ідея RoBERTa полягає не у зміні архітектури, а в поліпшенні тренувального процесу: було усунуто задачу передбачення наступного речення, збільшено розмір навчального корпусу, подовжено час тренування та введено динамічне маскування токенів, яке змінюється при кожній епохі навчання.

Завдяки таким змінам RoBERTa демонструє вищу точність і стабільність у більшості NLP-завдань, зокрема на невеликих або незбалансованих вибірках. Крім того, модель краще справляється з варіативністю синтаксису та стилістичних форм, що є типовим для неформального листування у месенджерах чи соціальних мережах.

На рисунку 2.8 наведено схему роботи RoBERTa при обробці двох послідовностей. Як і в BERT, використовується службовий токен [CLS], представлення якого подається на вхід класифікаційного шару. Проте на відміну від BERT, модель не має окремої задачі розпізнавання зв'язку між реченнями, що спрощує застосування у задачах, де важливий контекст всередині одного повідомлення.

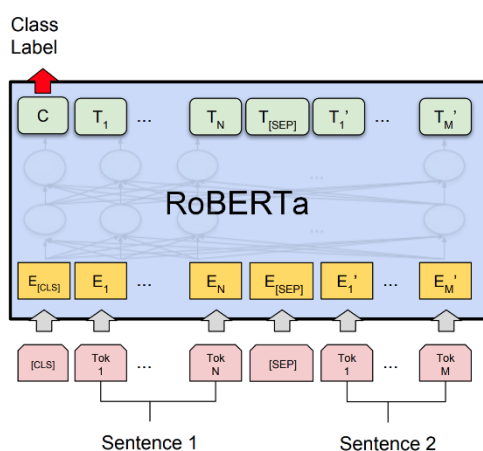


Рисунок 2.8 – Схема подання вхідних послідовностей у моделі RoBERT [22]

XLM-R (Cross-lingual RoBERTa) – це мультимовна трансформерна модель, розроблена на базі архітектури RoBERTa, яка здатна обробляти десятки мов, включно з українською. На відміну від моделей, орієнтованих лише на англійську, XLM-R тренується на великомасштабному корпусі CommonCrawl із понад 100 мовами, що дозволяє їй ефективно узагальнювати знання та переносити контекстуальне розуміння між мовами.

Оскільки створення окремих моделей для кожної мови є трудомістким, XLM-R дає змогу застосовувати одну спільну модель без потреби в додатковому попередньому навчанні для кожної конкретної мови. При цьому вона демонструє високу якість семантичного подання, навіть для мов із обмеженими ресурсами. Модель уніфікує обробку та забезпечує глибоке розуміння навіть коротких або неформальних висловлювань, характерних для повсякденного спілкування.

На рисунку 2.9 подано типову схему застосування XLM-R для задач класифікації: вхідний текст токенізується, передається до трансформерного енкодера, а фінальне представлення службового токена [CLS] використовується для прогнозування мітки класу через softmax-шар.

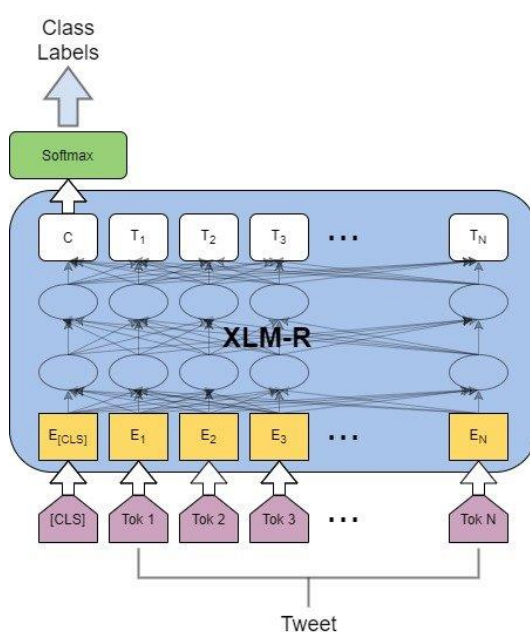


Рисунок 2.9 – Схема класифікації тексту за допомогою моделі XLM-R [23]

Хоча трансформерні моделі на кшталт BERT чи XLM-R спочатку створювались як універсальні мовні представлення, їхня ефективність у конкретних прикладних задачах, зокрема – виявленні емоцій, значною мірою залежить від подальшої адаптації. Така адаптація передбачає донавчання моделі на тематичних корпусах із міченими прикладами, де текст пов'язано з певними емоційними станами, наприклад: страхом, тривожністю, пригніченістю, гнівом або нейтральністю.

Процес адаптації складається з двох основних етапів: по-перше, підготовка відповідного датасету, який включає короткі, неформальні висловлювання (наприклад, повідомлення з форумів чи соціальних мереж); по-друге – тонке налаштування (*fine-tuning*) трансформерної моделі на цьому корпусі з використанням цільової розмітки. Такий підхід дозволяє моделі не лише зберігати загальні мовні знання, але й навчатися виявляти специфічні патерни, характерні для певного емоційного стану.

2.3 Генеративні моделі ШІ як інструмент діалогової взаємодії

У системах, що орієнтовані не лише на пасивне виявлення емоційного стану користувача, а й на надання зворотного зв'язку, дедалі більшого поширення набувають генеративні мовні моделі. Завдяки здатності створювати зв'язні, контекстуально релевантні та стилістично адаптивні тексти, такі моделі можуть виконувати роль співрозмовника або асистента, який реагує на емоційно забарвлені повідомлення емпатійно й делікатно.

2.3.1 Принцип роботи генеративних моделей

Генеративні мовні моделі, такі як GPT, працюють на основі ймовірнісного передбачення наступного слова в послідовності, враховуючи як локальний, так і глобальний контекст запиту. Цей підхід дозволяє формувати відповіді, що стилістично й тематично узгоджені з початковим

повідомленням користувача, навіть якщо воно сформульоване неформально або має приховане емоційне навантаження.

У центрі механізму генерації тексту лежить трансформерна архітектура, натренована на великих обсягах текстових даних. Цей механізм працює в авторегресивному режимі, тобто генерує текст поступово – токен за токеном, кожного разу спираючись на попередній контекст.

На кожному кроці модель обчислює логіти – числові значення для всіх можливих токенів словника, які потім перетворюються у ймовірності за допомогою функції softmax. Далі здійснюється вибір: або детермінований (наприклад, greedy decoding – вибір токена з найвищою ймовірністю), або стохастичний (sampling), зокрема:

- Top-k sampling – вибір із k найбільш імовірних токенів;
- Top-p (nucleus) sampling – вибір із мінімальної підмножини токенів, сума ймовірностей яких перевищує поріг p;
- Temperature scaling – регулювання «гостроти» розподілу для керування варіативністю.

Це дає змогу моделі бути як точною, так і гнучкою: повторна генерація на одній і тій самій основі може призвести до різних, але семантично обґрунтованих результатів. Така властивість особливо цінна для задач, пов'язаних із відповіданням користувачеві в емоційно насичених або невизначених ситуаціях – наприклад, у системах підтримки, ведення діалогу або навчальних платформах. Залежно від обраної стратегії вибору токенів, відповідь може бути лаконічною чи розгорнутою, нейтральною чи емпатичною, стандартною або варіативною за змістом.

На рисунку 2.10 зображено типовий крок генерації: модель на основі попередніх токенів (наприклад, «this is») обчислює ймовірності для наступного слова, після чого відбувається вибір наступного токена згідно з заданою стратегією.

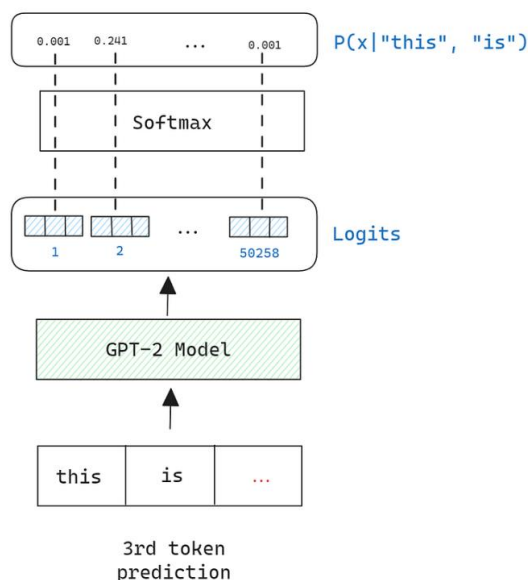


Рисунок 2.10 – Схема покрокового передбачення наступного токена в трансформерній моделі (GPT) [24]

2.3.2 Можливості генеративних моделей для побудови емпатійного зворотного зв'язку

Однією з ключових переваг генеративних мовних моделей у контексті моніторингу тривожності є здатність формувати емпатійно забарвлені відповіді, які не лише реагують на зміст повідомлення, а й враховують емоційний стан співрозмовника. Це особливо важливо в умовах, де відсутній безпосередній людський контакт, а користувач звертається до цифрового сервісу у момент психологічного напруження.

Моделі, як-от GPT-4, здатні адаптувати тональність відповіді залежно від інструкції або контексту: вони можуть бути стриманими, підбадьорливими, нейтральними або уважними. Наприклад, на повідомлення типу «не знаю, скільки ще витримаю», модель може згенерувати репліку на зразок: «Те, що ти це сказав – вже важливо. Дякую за довіру». Така відповідь не замінює професійної допомоги, однак здатна створити відчуття присутності, прийняття й тимчасової підтримки.

Для досягнення керованої поведінки генеративної моделі застосовується *prompt-інжиніринг* – процес формулювання спеціального текстового запиту (*prompt*), який визначає стиль, інтонацію, зміст або роль, у межах яких має відповідати модель. Це не є простою передачею інструкцій; йдеться про ретельно сформульований вхід, який використовується як контекст для генерації, впливаючи на структуру, тональність і характер відповіді.

У найпростішому випадку промпт може містити фрази на кшталт: «Відповідай коротко, спокійно, доброзичливо. Не давай порад. Просто визнай емоційний стан». Таке формулювання задає чітку поведінкову рамку, яка зберігається незалежно від того, яким є початковий запит користувача – формальним, розмитим чи емоційно насиченим. Завдяки цьому забезпечується послідовність та стабільність реакцій, що особливо важливо у системах з людським інтерфейсом: чат-ботах, віртуальних асистентах або навчальних застосунках.

На рисунку 2.11 представлено ключові складові *prompt-інжинірингу*. До них належать:

- *Token* – базові текстові одиниці, з якими працює модель;
- *Prompt* – сформульований запит, що ініціює генерацію;
- *Context* – додаткова інформація, яка підсилює або уточнює відповідь;
- *Priming* – приховані або завуальовані інструкції, що задають стиль чи рамку поведінки;
- *Fine-tuning* – не *prompt-орієнтований*, а параметричний спосіб контролю, який однак іноді комбінується з інженерією запитів.

Цей підхід дозволяє отримувати прогнозовані, адаптивні та інтерпретовані результати без потреби у зміні самої моделі, що робить *prompt-інжиніринг* одним із ключових інструментів у практичному застосуванні великих мовних моделей.

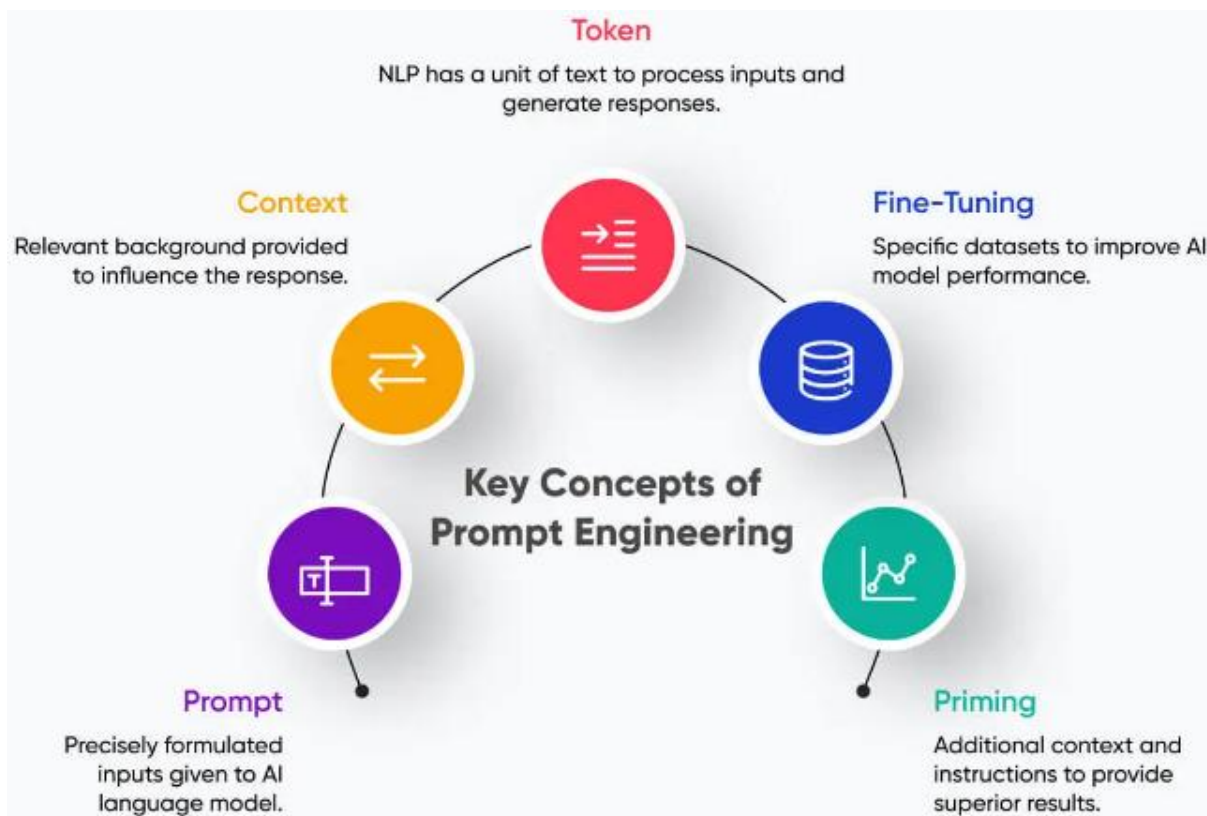


Рисунок 2.11 – Основні компоненти prompt-інжинірингу в роботі з мовними моделями [25]

У межах prompt-інжинірингу розрізняють кілька стратегій взаємодії з мовною моделлю залежно від того, як саме формується вхідний запит і яку інформацію він містить. Найбільш поширеними є три підходи: zero-shot, few-shot та instruction-style prompting.

Zero-shot prompting. Цей підхід передбачає відсутність прикладів у запиті. Моделі просто формулюється завдання у вигляді інструкції. У цьому режимі модель покладається виключно на своє попереднє навчання. Такий підхід зручний завдяки своїй простоті, але його ефективність залежить від того, наскільки чітко сформульоване завдання та наскільки модель вже знайома з подібними інструкціями.

Few-shot prompting. На відміну від zero-shot, у цьому випадку в prompt вбудовуються кілька демонстрацій того, як саме слід виконувати завдання. Такий формат дозволяє моделі краще вловити шаблон або інструкцію через

приклади, навіть без додаткового навчання. Цей підхід особливо корисний у задачах класифікації, узагальнення або стилізації відповідей.

Instruction-style prompting. Instruction prompting вирізняється тим, що запит формулюється як відкрита інструкція, а не задача з прикладами. Цей формат особливо добре працює з мовними моделями, попередньо навченими в інструкційній парадигмі (instruction-tuned LLMs, наприклад GPT-3.5-turbo, GPT-4, FLAN-T5 тощо), «які розуміють» і виконують команди природною мовою.

У багатьох реальних сценаріях ці стратегії комбінуються, що дозволяє досягати більш гнучкої й керованої поведінки генеративних систем.

Крім того, генеративні моделі характеризуються низьким порогом входу для користувача: взаємодія з ними не потребує спеціальних команд, структурованих запитів чи технічних знань. Людина може висловлюватись у звичній для себе формі – вільно, емоційно, фрагментарно – і все одно отримує релевантну відповідь. У таких випадках модель не намагається аналізувати чи оцінювати, а радше демонструє присутність, уважність і базову підтримку. Саме ця форма «неінвазивного» реагування забезпечує високу психологічну прийнятність, що є критично важливою у ситуаціях підвищеної емоційної чутливості.

Таким чином, генеративний підхід формує нову якість текстової взаємодії, у якій на перший план виходять не лише смислові відповідності, а й інтонаційна м'якість, ритмічна плавність та емоційна коректність відповіді. Це дає змогу моделі адаптуватися до комунікативного стилю користувача, підвищуючи суб'єктивне відчуття зрозумілості, підтримки й персоналізованої присутності.

3 ПРОЄКТУВАННЯ ТА РЕАЛІЗАЦІЯ СИСТЕМИ МОНІТОРИНГУ ТРИВОЖНОСТІ

У процесі створення інтелектуальної системи моніторингу тривожності ключовим завданням було забезпечити поєднання точного автоматичного аналізу з максимально зручною формою взаємодії. Результатом стала повноцінна програмна реалізація, що об'єднує класифікаційну модель, лінгвістичні механізми та генеративні компоненти у межах простого користувацького інтерфейсу.

3.1 Середовище розробки

Розробку системи здійснено мовою програмування Python 3.10 у віртуальному середовищі, сформованому засобами `venv`. Для створення Telegram-бота, що виступає основним інтерфейсом користувача, використано асинхронну архітектуру на базі бібліотеки `python-telegram-bot`. Збереження даних реалізовано за допомогою вбудованої бази `sqlite3`, формування графіків – засобами `matplotlib`, а створення PDF-звітів – за допомогою бібліотеки `fpdf`.

Класифікаційна модель базується на архітектурі `roberta-base`, завантаженої з бібліотеки `transformers` і інтегрованої із фреймворком `torch`. Підготовка та балансування корпусу здійснювалися із використанням `datasets` та модулів `scikit-learn`. Переклад україномовних повідомлень англійською реалізовано засобами `deep_translator`, а лінгвістичний аналіз тексту з метою виявлення типових симптомів тривожності – з використанням `srasu`.

Генеративна логіка системи забезпечується через доступ до моделі GPT за допомогою API бібліотеки `openai`. Вона використовується для побудови пояснень до окремих повідомлень користувача, а також для формування підсумкових висновків у звітах.

Допоміжні функції реалізовані з використанням стандартних бібліотек Python (os, datetime, re, uuid, json, traceback), що забезпечують обробку даних, генерацію унікальних ідентифікаторів, роботу з форматами дати та обробку помилок.

3.2 Вхідні дані та попередня обробка

Для побудови класифікатора тривожності було використано відкритий датасет Sentiment Analysis for Mental Health, розміщений на платформі Kaggle [26]. Початковий корпус містить 51 074 унікальних текстових повідомлення, кожне з яких має одну або кілька міток, що вказують на психоемоційний стан (наприклад, anxiety, depression, adhd, bipolar, normal тощо).

У межах цього проєкту було обрано лише два класи – anxiety та normal, як такі, що дозволяють найбільш чітко розмежувати тривожний стан і його відсутність. Після відфільтрування записів з відповідними мітками було збережено 3888 прикладів класу anxiety та 16 351 приклад класу normal. Типові приклади текстів кожного класу наведено на рисунку 3.1.

■ Приклади з навчального набору:

label	text
anxiety	Can penicillin cause headaches? I've been on penicillin to treat strep throat for a couple days now...
anxiety	Have my body forgotten how to breath properly? I have a constant feeling of tightness in my chest. N...
anxiety	Booked a health test - total regression Hi, I've experienced health anxiety on and off for most of m...
normal	Where are these many previews of children?
normal	Most furious to choose food randomly near grab 30 min looking for it, then it's closed like it doesn...
normal	ha the biggest headache but a second job woo

Рисунок 3.1 – Типові приклади текстів з навчального набору

Для уникнення зміщення моделі у бік переважаючого класу було здійснено балансування: випадкова підмножина normal була приведена до

кількості прикладів класу anxiety. У результаті утворено збалансований корпус із загальною кількістю 7776 прикладів – по 3888 з кожного класу.

Отриманий набір було випадковим чином поділено у співвідношенні 80:20 на тренувальну та валідаційну частини. Графік розподілу класів у тренувальній вибірці наведено на рисунку 3.2.

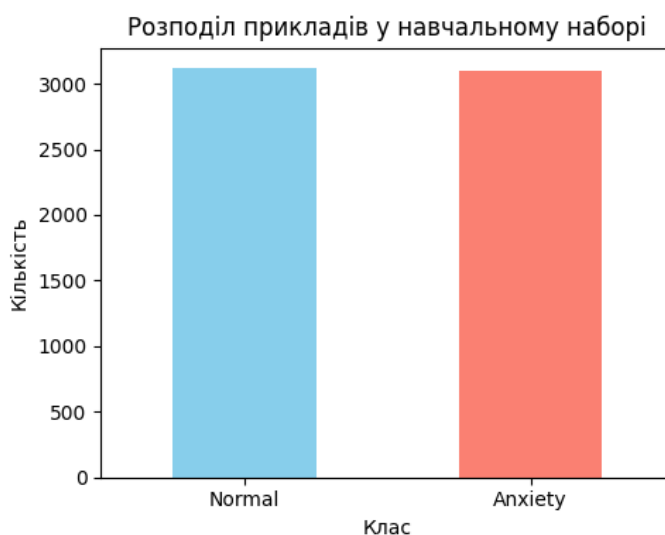


Рисунок 3.2 – Розподіл прикладів у навчальному наборі

Під час підготовки корпусу не виконувалась жодна текстова нормалізація, така як зниження регістру, лематизація або видалення пунктуації. Це пояснюється тим, що в подальшому планувалося використання англійської трансформерної моделі, яка має власний токенизатор, здатний обробляти текст у первинному вигляді. Токенізація виконувалась на етапі навчання, із застосуванням RobertaTokenizer з моделі roberta-base. Попередньо зчитані з диска файли було перетворено у формат Dataset з бібліотеки Hugging Face, після чого тексти було токеновано з фіксованою максимальною довжиною 128 токенів. Отримані числові представлення тексту (індекси токенів, маски уваги та мітки класів) використовувалися безпосередньо у навчальному циклі, без додаткового збереження на диск.

3.3 Архітектура системи

Реалізована система побудована за модульною архітектурою, що поєднує інтерфейс взаємодії з користувачем у вигляді Telegram-бота, класифікаційний модуль на основі трансформерної моделі, допоміжні компоненти для лінгвістичного аналізу та генерації тексту, а також засоби збереження та звітності. Усі обчислення виконуються локально, без використання зовнішніх серверів або хмарних інтерфейсів, що забезпечує автономність і конфіденційність обробки.

Загальна схема взаємодії компонентів подана на рисунку 3.3.

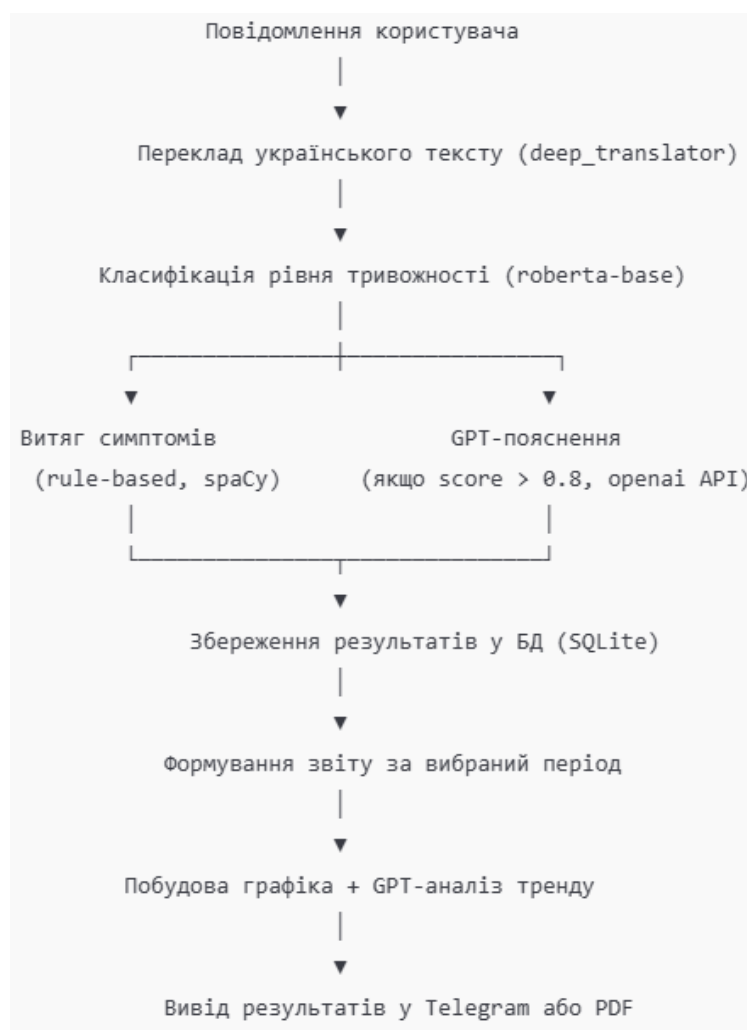


Рисунок 3.3 – Архітектура реалізованої системи моніторингу тривожності

Користувач взаємодіє із системою шляхом надсилання повідомлень у Telegram. Далі запускається послідовна обробка, яка включає переклад вхідного тексту, класифікацію рівня тривожності, виявлення характерних симптомів, за потреби – формування генеративного пояснення, а також збереження результатів у базу даних. На основі накопичених даних користувач може отримати звіт за вибраний період із графічною візуалізацією й додатковим GPT-коментарем щодо динаміки тривожності. Поділ системи на незалежні функціональні блоки забезпечує гнучкість подальшого розвитку: кожен з компонентів – класифікація, аналіз, генерація, збереження або звітність – може бути доопрацьований або замінений без зміни загальної логіки роботи системи.

3.4 Класифікація рівня тривожності

Для реалізації класифікаційного модуля системи було обрано попередньо навчену англomовну модель roberta-base [27], розроблену на основі трансформерної архітектури. Цей варіант було обрано з огляду на високу якість представлення коротких текстів та здатність моделі зберігати контекст у межах обмеженої довжини вхідної послідовності.

Модель має ≈ 124 млн параметрів, включає 12 шарів, 12 голів самоуваги та прихований розмір 768 (рисунок 3.4).

Layer (type (var_name))	Input Shape	Output Shape	Param #	Trainable
RobertaForSequenceClassification (RobertaForSequenceClassification)	--	[1, 2]	--	True
└RobertaModel (roberta)	[1, 128]	[1, 128, 768]	--	True
└┬RobertaEmbeddings (embeddings)	--	[1, 128, 768]	39,000,576	True
└┬┬RobertaEncoder (encoder)	[1, 128, 768]	[1, 128, 768]	85,054,464	True
└┬RobertaClassificationHead (classifier)	[1, 128, 768]	[1, 2]	--	True
└┬┬Dropout (dropout)	[1, 768]	[1, 768]	--	--
└┬┬┬Linear (dense)	[1, 768]	[1, 768]	590,592	True
└┬┬┬Dropout (dropout)	[1, 768]	[1, 768]	--	--
└┬┬┬Linear (out_proj)	[1, 768]	[1, 2]	1,538	True
Total params: 124,647,170				
Trainable params: 124,647,170				
Non-trainable params: 0				
Total mult-adds (M): 124.65				
Input size (MB): 0.00				
Forward/backward pass size (MB): 106.96				
Params size (MB): 498.59				
Estimated Total Size (MB): 605.55				

Рисунок 3.4 – Архітектура класифікаційної моделі roberta-base

Підготовлений і збалансований корпус даних було токенізовано з використанням `RobertaTokenizer`, з обмеженням максимальної довжини до 128 токенів. Навчання здійснювалося із використанням бібліотеки `Hugging Face Transformers` та фреймворку `PyTorch`. Для організації тренувального циклу застосовувався клас `Trainer`, із явно заданими метриками `accuracy`, `precision`, `recall` та `f1`. Функція втрат – `CrossEntropyLoss`, оптимізатор – `AdamW`. Навчання тривало 3 епохи на центральному процесорі, при розмірі батчу 8 та фіксованій швидкості навчання $2 \cdot 10^{-5}$. Модель, яка показала найкраще значення F1-міри на валідаційній вибірці, була збережена для подальшого використання у Telegram-боті.

Отримані результати свідчать про високу ефективність класифікації. На тестовій вибірці модель досягла точності 99%, значення F1-міри – 0.99 для обох класів, при збереженому балансі `precision` та `recall` (рисунок 3.5). Матриця помилок (рисунок 3.6) демонструє лише 22 хибних класифікації зі 1556 прикладів: 10 `false positives` та 12 `false negatives`. Такий результат свідчить про високу узгодженість моделі з розміткою та достатню чутливість до контекстуально виражених проявів тривожності.

	precision	recall	f1-score	support
Normal	0.98	0.99	0.99	769
Anxiety	0.99	0.98	0.99	787
accuracy			0.99	1556
macro avg	0.99	0.99	0.99	1556
weighted avg	0.99	0.99	0.99	1556

Рисунок 3.5 – Основні метрики якості класифікації

Додатково було здійснено аналіз хибних класифікацій (рисунок 3.7). У прикладах `false positives` спостерігаються фрази, що містять загальні переживання або неструктуровані емоційні висловлювання без чітких маркерів тривоги. Натомість помилки типу `false negatives` частіше містять

непрямі згадки про тривожні стани – через згадування ліків, сторонніх осіб або ситуацій, що потребують інтерпретації. Це свідчить про те, що модель у деяких випадках не розпізнає опосередковані або контекстуально складні прояви тривожності.

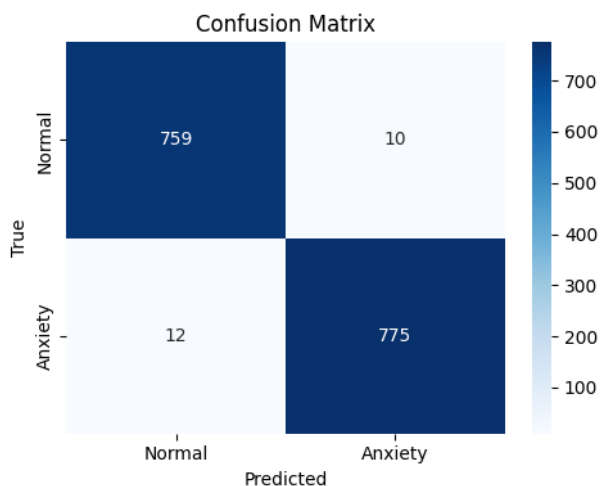


Рисунок 3.6. – Матриця помилок на тестовій вибірці

● False Positives (нормальні, але визнані тривожними):

text	true	predicted
dangermouse i think maybe you should get a couple more hour of sleep hon how productive can you be r...	0	1
If this is all I'm going to experience side-effect wise, then I regret not starting earlier. I reall...	0	1
Work! Isn't it normal if you are bored looking for work.. I don't think so all day long I just stay ...	0	1

● False Negatives (тривожні, але визнані нормальними):

text	true	predicted
Lamictal ? Anyone with good reviews about lamictal?	1	0
"Buy Friends" comment? My mother doesn't have friends. She said if she won a lot of money it would b...	1	0
Tonight he continues to chatter about money, I'm so confused what to do. Once he said don't worry ab...	1	0

Рисунок 3.7 – Приклади хибних класифікацій: false positives та false negatives

3.5 Інтеграція компонентів у Telegram-бот

Telegram-бот виступає єдиною точкою взаємодії користувача з усіма модулями системи. Він реалізує не лише інтерфейс введення й виведення, а й координує повний цикл обробки повідомлення – від моменту надсилання

тексту до генерації звіту чи GPT-пояснення. Роботу бота реалізовано з використанням бібліотеки `python-telegram-bot` (версія 20), що дозволяє побудувати асинхронну архітектуру з підтримкою кнопкової взаємодії, обробників команд, контексту користувача та інлайн-відповідей.

Основний сценарій виглядає так: після запуску бота або натискання кнопки «Запис» користувач надсилає текстове повідомлення. Його одразу зберігає система, при цьому не відображаючи жодних технічних деталей – тільки повідомлення про успішне збереження. Далі текст автоматично перекладається з української на англійську (`deep_translator`) і подається на вхід класифікаційній моделі `roberta-base`. Отриманий `score` фіксується разом з повідомленням у локальній базі даних `SQLite`. Крім того, викликається модуль аналізу симптомів, а у випадку високого рівня `score` (понад 0.8) – додатково формується GPT-пояснення. Усі ці дії відбуваються автоматично, без потреби ручного підтвердження. Загальний сценарій взаємодії користувача з ботом, включно з надсиланням повідомлення, підтвердженням збереження та автоматичним GPT-поясненням, подано на рисунку 3.8.

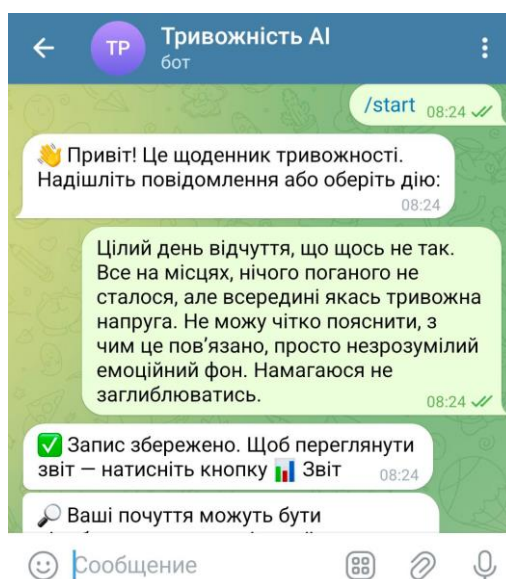


Рисунок 3.8 – Початок взаємодії з ботом: повідомлення, збереження, GPT-пояснення

Збереження здійснюється у таблицю з полями `user_id`, `original_text`, `translated_text`, `score`, `symptoms`, `timestamp`. Користувач безпосередньо не бачить цих параметрів – система працює у форматі емоційного щоденника, де кожен запис зберігається у фоні.

За запитом користувача бот формує узагальнений звіт за день, тиждень, місяць, весь час або вручну обраний діапазон. Звіт містить кількість записів, середній рівень тривожності, список найчастіших симптомів, графік зміни `score` у часі та, за потреби, GPT-аналіз. Побудова графіка виконується засобами `matplotlib`, а форматування звіту – у вигляді текстового повідомлення або PDF-документа (`fpdf`). Користувач може обрати між стандартним PDF-звітом і розширеним варіантом, який містить усі повідомлення за вказаний період.

Завдяки такій побудові Telegram-бот координує всю логіку системи: запуск класифікації, виклик GPT, обробку кнопок, збереження в БД, побудову графіків і генерацію звітів. Це забезпечує цілісну інтеграцію всіх компонентів, зручність використання і готовність до розгортання без необхідності зовнішнього серверного середовища.

3.6 Аналіз тривожних проявів

На додаток до класифікації повідомлення як тривожного або нейтрального, реалізована система виконує поглиблений аналіз змісту повідомлення з метою виявлення характерних симптомів тривожності. На відміну від основної моделі, цей модуль не використовує машинного навчання, а базується на словниковому підході із фіксованими шаблонами. Його мета – виявити, які саме прояви тривожного стану присутні в тексті: когнітивні, емоційні, фізіологічні чи поведінкові.

Словник шаблонів було сформовано вручну на основі аналізу корпусу навчальних даних, з урахуванням лексем, які найчастіше траплялись у

прикладях класу anxiety. Відібрані слова й фрази було згруповано за чотирма типами тривожних проявів:

- когнітивні: пов’язані з мисленням, надмірною концентрацією або нав’язливими думками (наприклад, «can’t focus», «overthinking», «confused»);
- емоційні: відображають настрій або внутрішній стан (наприклад, «panic», «fear», «nervous»);
- фізіологічні: включають згадки про тілесні симптоми (наприклад, «tired», «headache», «chest pain»);
- поведінкові: стосуються реакцій уникнення, зниження активності чи ізоляції (наприклад, «hide», «stay home», «cancel plans»).

Для виявлення відповідних фрагментів текст обробляється модулем `sparse`: здійснюється лематизація, зниження реєстру, виділення речень. Далі до кожного речення застосовуються регулярні вирази, що відповідають шаблонам з попередньо створеного словника (рисунку 3.9). Якщо виявлено збіг, симптом додається до списку виявлених, разом із його категорією.

```

SYMPTOM_PATTERNS = {
  "емоційні": [
    "fear", "panic", "anxiety", "worry", "nervous", "scared", "afraid", "stressed"
  ],
  "фізичні": [
    "headache", "fatigue", "tired", "sweat", "chest pain", "palpitations",
    "shortness of breath", "dizzy", "nausea", "heart racing", "pressure",
    "can't sleep", "insomnia", "chest hurts"
  ],
  "когнітивні": [
    "can't concentrate", "thoughts", "looping", "rumination", "overthinking",
    "can't think", "confused", "can't focus"
  ],
  "поведінкові": [
    "avoid", "hide", "cancel plans", "stay home", "skip", "check again"
  ]
}

```

Рисунок 3.9 – Словник тривожних симптомів, структурований за категоріями

Цей модуль викликається автоматично для кожного повідомлення, незалежно від його класифікаційного результату. Виявлені симптоми не відображаються користувачу безпосередньо, проте зберігаються у базу даних і використовуються під час побудови щоденникового звіту. У підсумку користувач бачить найбільш частотні симптоми за період, згруповані за типами (рисунок 3.10).

Це дозволяє не лише оцінити загальний рівень тривожності, а й відстежувати, як саме вона проявляється – через фізичне виснаження, когнітивні труднощі, емоційні стани чи зміну поведінки.

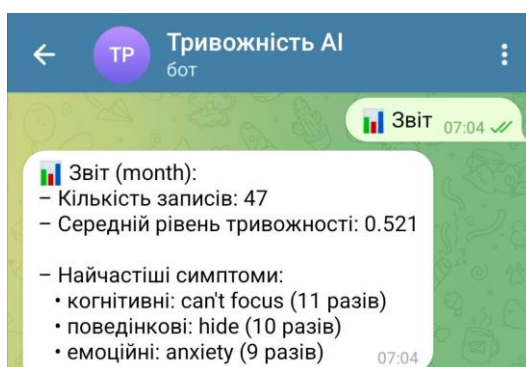


Рисунок 3.10 – Приклад підсумкового виводу симптомів у Telegram-боті

3.7 Генеративні функції GPT

Генеративна модель GPT інтегрована в систему для формування персоналізованих пояснень та інтерпретацій результатів аналізу. Її використання не впливає на класифікаційне рішення, однак відіграє критичну роль у підвищенні інтерпретованості та суб'єктивного сприйняття результатів користувачем. Застосування великої мовної моделі в цьому контексті обґрунтовано тим, що класичні правила або заготовлені фрази не здатні враховувати змістовні особливості конкретного повідомлення або змінювати тон відповіді залежно від ситуації.

Перше застосування GPT реалізовано як генерація емпатичного пояснення до повідомлення користувача, якщо модель класифікує його як тривожне з високою впевненістю ($\text{score} > 0.8$). Такий поріг дозволяє уникнути зайвих викликів генеративної моделі для повідомлень із неоднозначною емоційною забарвленістю, а також знижує навантаження на API без втрати цінності зворотного зв'язку.

Пояснення формулюється українською мовою у вигляді звертання на «Ви», з емпатичним, м'яким тоном і без формального аналізу. Для цього до моделі передається промпт із чітко визначеними інструкціями, зокрема: «сформулюйте коротке пояснення для людини, яка це написала... не більше двох речень, не аналізуйте, не повторюйте фрази користувача, будьте доброзичливі» тощо. Приклад такої відповіді наведено на рисунку 3.11. Відповідь додається до повідомлення лише після його збереження у базу даних.

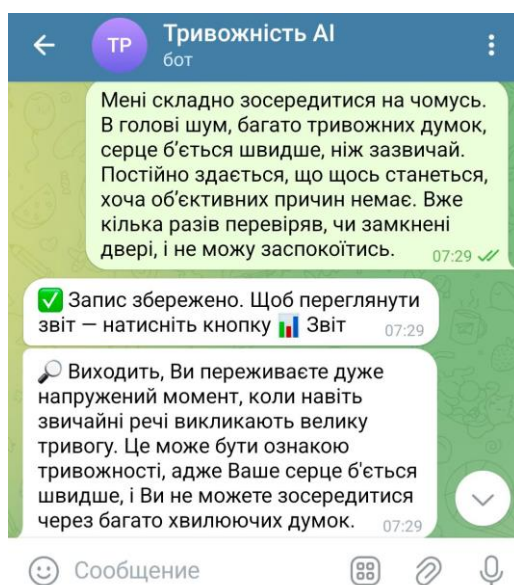


Рисунок 3.11 – GPT-пояснення до повідомлення з високим рівнем тривожності

Друге застосування GPT – генерація підсумкового аналітичного коментаря у звіті. Після побудови графіка тривожності за вибраний період і

підрахунку основних показників система автоматично формує текстовий запит до моделі, який містить: кількість повідомлень, середній score, найчастіші симптоми, а також щоденні середні значення тривожності. GPT на основі цих даних генерує короткий огляд, у якому аналізується динаміка (покращення, погіршення, стабільність) та окреслюються можливі причини – знову ж таки, в максимально підтримувальному, а не діагностичному тоні (рисунок 3.12).

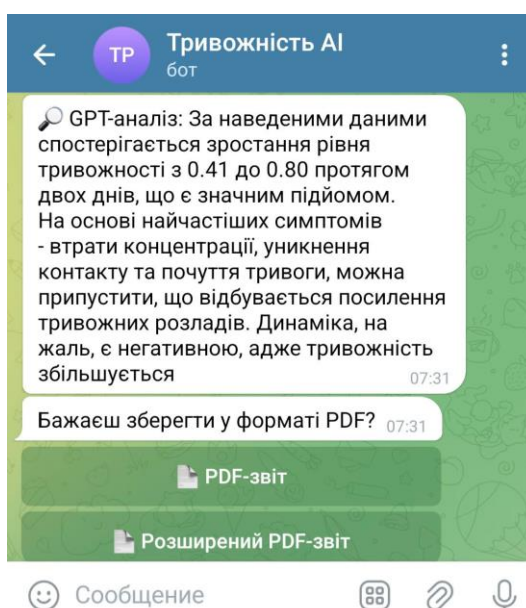


Рисунок 3.12 – GPT-аналіз динаміки тривожності у межах звіту

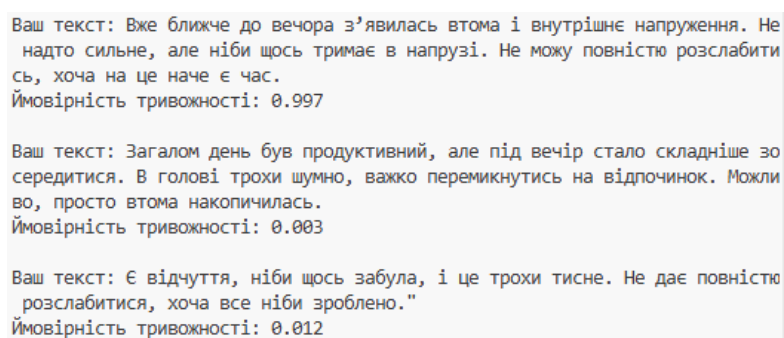
Обидва генеративні сценарії реалізовано за допомогою API бібліотеки OpenAI. У роботі використовувалась модель GPT-4o, з параметрами температури у межах 0.5–0.6 та обмеженням довжини відповіді. Для користувача цей функціонал виглядає як природне доповнення до класифікаційного висновку – не у формі медичного пояснення, а у вигляді короткого, психологічно комфортного тексту, що покликається підтримати, а не інтерпретувати. У випадку недоступності сервісу GPT виводиться службове повідомлення про неможливість створити пояснення.

4 ЕКСПЕРИМЕНТАЛЬНА ПЕРЕВІРКА ТА ОЦІНКА РЕЗУЛЬТАТІВ

4.1 Перевірка класифікаційної моделі на тестових прикладах

Перед інтеграцією моделі у Telegram-бот було здійснено окреме тестування класифікатора roberta-base на прикладах текстів, що відрізняються за рівнем вираженості тривожних проявів. Це дозволило оцінити поведінку моделі при класифікації типових повідомлень, а також переконатися, що повернене значення score справді корелює з емоційним тоном тексту.

На рисунку 4.1 наведено три фрагменти, подані на вхід моделі за допомогою окремого скрипта infer.py. У першому прикладі описано виражене відчуття напруги, труднощі з розслабленням і фонову тривогу. Модель повернула $score = 0.997$, що відповідає високій впевненості в тому, що повідомлення є тривожним. У другому випадку текст виглядає більш збалансовано, з частиною нейтральної лексики, однак містить згадки про шум у голові, втому, труднощі з перемиканням уваги. Незважаючи на це, модель класифікує повідомлення як нейтральне з дуже низьким значенням $score (0.003)$. Третій приклад – дещо двозначний: у ньому є фрази про тиск і напругу, але загальний тон залишається стриманим. Модель і тут не виявляє достатньо маркерів тривожності й повертає $score = 0.012$.



```
Ваш текст: Вже ближче до вечора з'явилась втома і внутрішнє напруження. Не надто сильне, але ніби щось тримає в напрузі. Не можу повністю розслабитися, хоча на це наче є час.  
Ймовірність тривожності: 0.997  
  
Ваш текст: Загалом день був продуктивний, але під вечір стало складніше зосередитися. В голові трохи шумно, важко перемикнути на відпочинок. Можливо, просто втома накопичилась.  
Ймовірність тривожності: 0.003  
  
Ваш текст: Є відчуття, ніби щось забула, і це трохи тисне. Не дає повністю розслабитися, хоча все ніби зроблено."  
Ймовірність тривожності: 0.012
```

Рисунок 4.1 – Результати класифікації трьох прикладів повідомлень: score для кожного випадку

Усі ці приклади демонструють, що модель працює як «рішучий фільтр»: у випадках, коли повідомлення не містить явних маркерів тривожності, вона повертає score, близький до нуля. Водночас емоційно насичені повідомлення ідентифікуються з високою впевненістю. Це свідчить про добру узгодженість моделі з навчальними даними і коректну поведінку на нових прикладах.

4.2 Поведінка системи у Telegram-боті

На цьому етапі проекту було протестовано функціональність системи в інтегрованому середовищі Telegram-бота. Головна мета полягала у перевірці того, як класифікаційна модель, генеративні функції, аналіз симптомів і база даних узгоджено взаємодіють у межах одного сценарію.

Загальний алгоритм роботи виглядає наступним чином: після запуску команди /start користувач надсилає текстове повідомлення. Бот автоматично ініціалізує повний цикл обробки: переклад вхідного тексту з української на англійську (deep_translator), класифікацію з використанням моделі roberta-base, виявлення симптомів за допомогою rule-based модуля (symptom_extractor.py), збереження результату у базу даних (sqlite3) та, за потреби, генерацію пояснення через GPT.

Усі повідомлення обробляються у фоновому режимі, а користувач отримує лише стисле підтвердження про успішне збереження. Якщо score перевищує встановлений поріг (0.8), бот додатково надсилає коротке GPT-пояснення, сформульоване у доброзичливій формі. Воно не дублює зміст повідомлення, не містить оцінних суджень, а лише пояснює, чому описані стани можуть свідчити про тривожність. Зовнішньо це виглядає як емпатичне продовження діалогу, не порушуючи логіки щоденникового режиму.

Усі додаткові аналітичні дії, включно з фіксацією score та симптомів, залишаються прихованими для користувача – вони використовуються лише

при формуванні звіту. Такий підхід дозволяє не перевантажувати інтерфейс зайвими технічними деталями, зберігаючи емоційно безпечний стиль взаємодії. При цьому GPT-реакції з'являються лише тоді, коли класифікатор має високу впевненість у тривожному контексті.

Приклад повної взаємодії представлено на рисунку 4.2: користувач надсилає повідомлення, що містить опис складного емоційного стану, система зберігає його у базу, і на основі високого score автоматично формує підтримувальне GPT-пояснення. Жодних технічних термінів не виводиться, інтерфейс залишається мінімалістичним та зрозумілим.

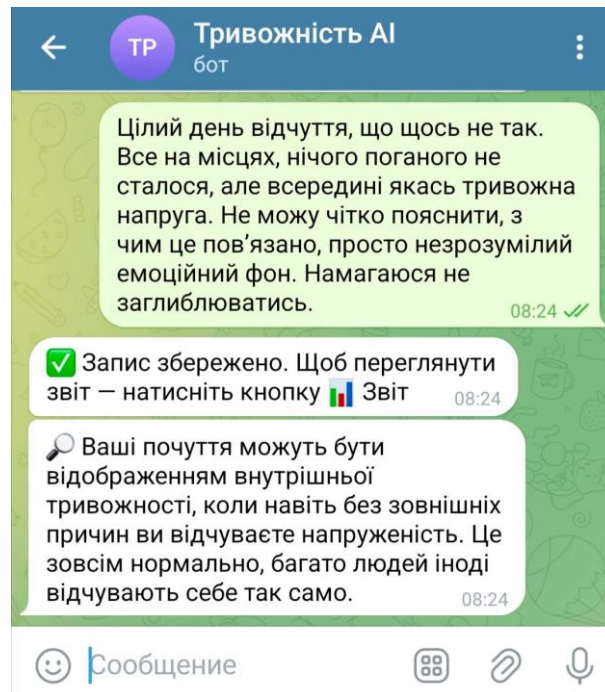


Рисунок 4.2 – Взаємодія з ботом: повідомлення, підтвердження збереження, GPT-пояснення

4.3 Формування та структура звітів

Ключовим елементом реалізованої системи є механізм створення звітів, що дозволяє користувачеві отримати агреговану картину змін емоційного стану на основі власних повідомлень. Ця функція доступна

безпосередньо в Telegram-боті та працює за запитом – як через кнопки з фіксованими періодами, так і за допомогою ручного введення діапазону дат (рисунок 4.3).

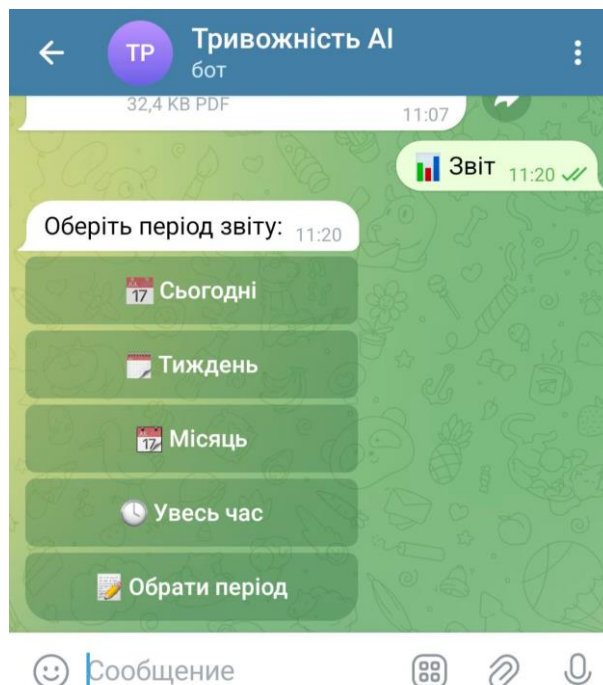


Рисунок 4.3 – Інтерфейс вибору періоду

Після вибору періоду бот здійснює вибірку з бази даних усіх повідомлень користувача за вказаний інтервал. Зібрана інформація агрегується та подається у вигляді текстового звіту, що включає:

- кількість записів;
- середній рівень тривожності (score);
- список найпоширеніших симптомів, згрупованих за категоріями;
- генеративний GPT-коментар щодо загальної динаміки (за наявності достатніх даних).

Такий формат дозволяє не лише оцінити загальний стан, а й побачити його прояви в деталях. GPT-висновок формується автоматично на основі узагальнених значень score та виявлених симптомів – він описує динаміку змін і пропонує м'яке інтерпретаційне пояснення (рисунок 4.4).

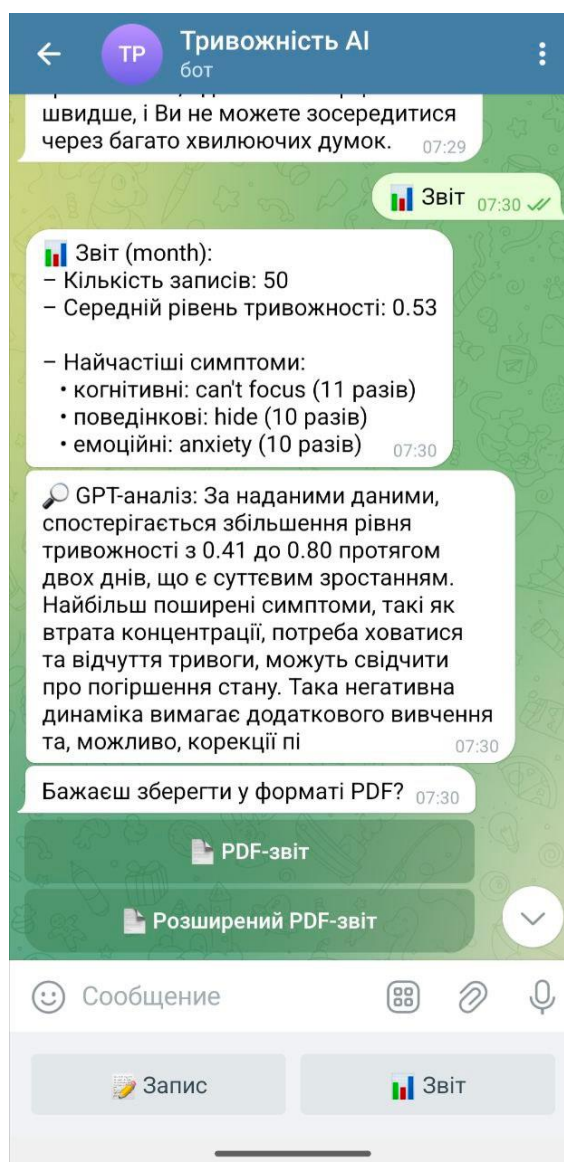


Рисунок 4.4 – Структура Telegram-звіту

Для зручності користувача звіт також може бути збережено у форматі PDF. Передбачено два варіанти:

– стислий PDF-звіт – містить ключові статистичні показники (період, кількість повідомлень, середній рівень тривожності, симптоми) без детальної розшифровки (рисунок 4.5);

– розширений PDF-звіт – містить також повний список повідомлень користувача за період, кожне з яких супроводжується індивідуальним значенням score. Це дає змогу аналізувати зміну емоційного стану у динаміці та повертатися до окремих записів за потреби.

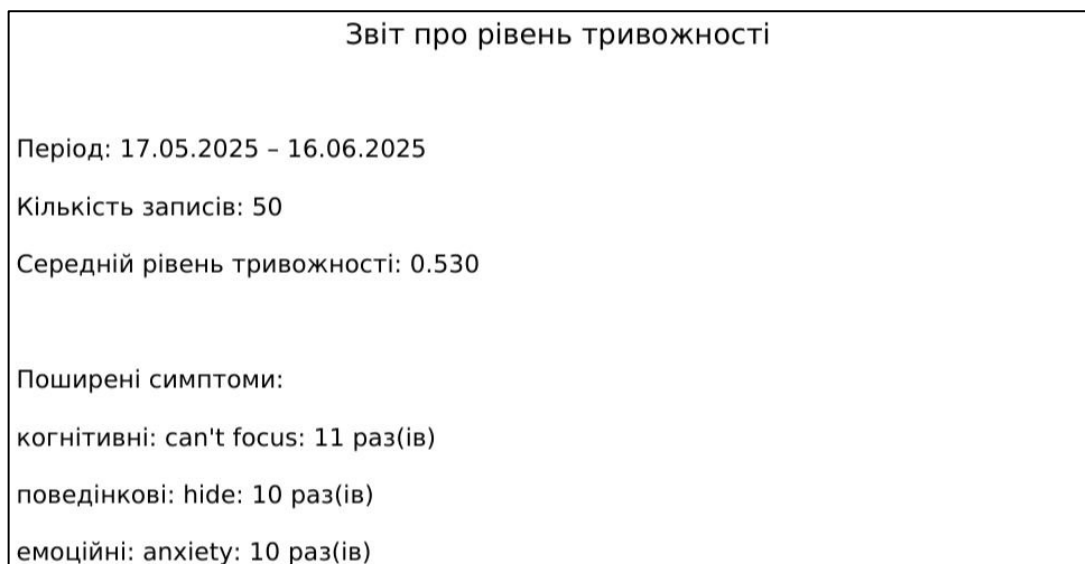


Рисунок 4.5 – Стислий PDF-звіт

Обидва варіанти PDF-звіту формуються автоматично після натискання відповідної кнопки та надсилаються користувачу безпосередньо в чат. Такий механізм трансформує накопичені повідомлення в структуровану аналітичну форму, зручну як для особистого використання, так і для потенційного звернення до фахівця.

Крім основних функцій, у процесі реалізації було також додано механізм побудови графіка рівня тривожності за часовими мітками. Цей графік виводиться разом зі звітом у Telegram і дозволяє візуально оцінити динаміку емоційного стану користувача (рисунок 4.6). Візуалізацію реалізовано засобами бібліотеки `matplotlib` з автоматичним формуванням графіка на основі щоденних або погодинних `score`.

Побудова графіка є особливо корисною в контексті самоспостереження, оскільки дає змогу побачити зміни емоційного стану не лише чисельно, а й у наочній формі. За візуальною кривою користувач може відстежити, в які дні чи години спостерігалось загострення або стабілізація тривожності, не аналізуючи кожне повідомлення окремо.

Оскільки графік будується динамічно, його можна отримати як за короткий період (наприклад, «Сьогодні»), так і за тривалу історію записів.

Такий підхід робить систему не лише технічно функціональною, а й доступною для повсякденного використання без потреби у спеціальній підготовці або аналітичних навичках.

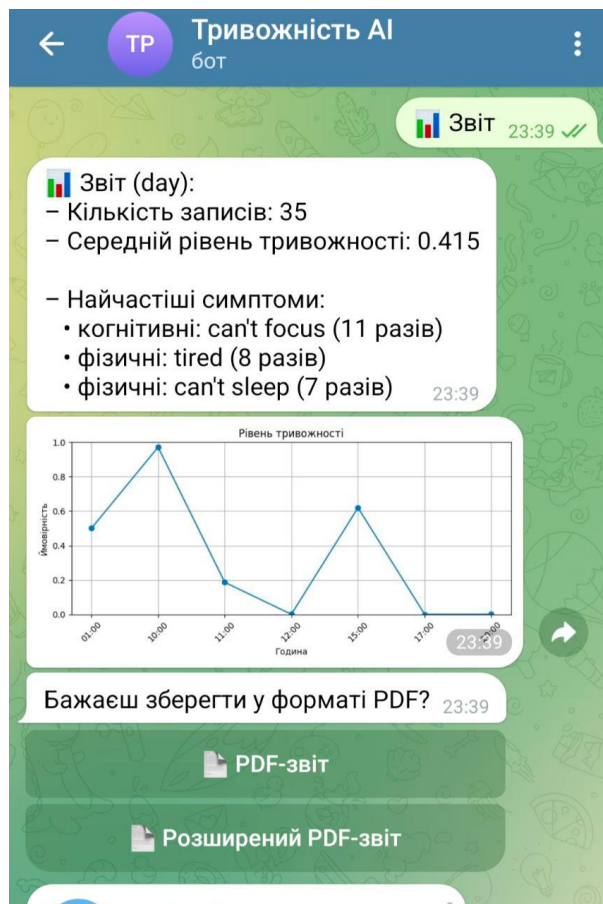


Рисунок 4.6 – Візуалізація динаміки тривожності

4.4 Обмеження системи та напрями подальшого розвитку

Попри успішну реалізацію функціональних компонентів, розроблена система моніторингу тривожності має низку обмежень, які необхідно враховувати при оцінюванні її довготривалої ефективності та можливостей масштабування.

Одним з ключових технічних обмежень є використання англомовної трансформерної моделі roberta-base, що потребує попереднього машинного перекладу вхідних повідомлень з української мови. Хоча переклад

здійснюється автоматично й із достатньою точністю, цей підхід не гарантує збереження всіх семантичних і стилістичних нюансів тексту, зокрема емоційного забарвлення, модальності чи контексту. Втрата таких деталей може впливати на коректність класифікації, особливо у випадках з неявно вираженими симптомами або з використанням розмовної лексики.

Окремої уваги потребує обмеження класифікаційної моделі в частині обробки контексту. Система аналізує кожне повідомлення ізольовано, без урахування попередніх записів користувача. Це унеможливорює відстеження послідовності станів, накопичення симптомів або виявлення закономірних змін у динаміці психоемоційного фону. Таким чином, реалізований підхід забезпечує лише фіксацію миттєвого емоційного стану, але не дозволяє здійснювати довгостроковий моніторинг або прогнозування. Крім того, відсутність механізмів автоматичного навчання на нових даних означає, що модель не адаптується до мовних особливостей конкретного користувача, не оновлює свої уявлення про лексичні патерни тривожності та не враховує змін у стилі мовлення з часом.

Ще одним обмеженням є застосування rule-based підходу для виявлення симптомів. Незважаючи на те, що такий підхід дозволяє чітко контролювати категорії ознак (когнітивні, емоційні, фізіологічні, поведінкові), він залишається жорстким і не масштабованим: шаблони формуються вручну, не враховують синонімії, граматичних конструкцій та багатозначності. Система не розпізнає перефразовані або метафоричні вирази, не враховує заперечення, модальні конструкції або інтенсивність стану. Як наслідок, реальні прояви тривожності можуть бути втрачені, якщо не збігаються із заздалегідь визначеними шаблонами. Водночас система не дозволяє персоналізувати список ознак або додати нові симптоми в процесі використання, що обмежує її адаптивність.

Обмеження має також генеративний компонент системи. Модель GPT-4o використовується у двох режимах: пояснення окремого повідомлення та аналіз динаміки у звіті. В обох випадках генерація

здійснюється на основі мінімального контексту: або одного тексту, або статистичних даних за період. GPT не має доступу до повної історії спілкування, не враховує попередні реакції, не адаптується до стилю відповіді, і не використовує симптоми чи лінгвістичні патерни при формуванні реплік. У результаті відповіді можуть бути повторюваними або занадто узагальненими. Крім того, виклик API залежить від зовнішнього сервісу, що створює ризики затримок, помилок або недоступності сервісу в критичні моменти.

У поточній реалізації відсутня підтримка персоналізації. Усі користувачі отримують однаковий функціонал, порогові значення класифікації, шаблони аналізу та стилі відповіді. Немає механізмів ідентифікації користувача з метою тривалого ведення історії або налаштування індивідуальних параметрів. Це обмежує потенціал системи як інструмента самостереження з урахуванням особистого психоемоційного профілю.

Нарешті, система не реалізує жодного з механізмів інтерактивного зворотного зв'язку. Користувач не має змоги позначити, чи був класифікаційний результат або GPT-відповідь доречними, корисними або навпаки – неадекватними. Така зворотна взаємодія була б цінною для подальшого аналізу точності, довіри до системи та її прийнятності в реальних умовах.

Незважаючи на наведені обмеження, реалізована система демонструє стійку роботу в умовах фіксованого сценарію використання. Усі модулі функціонально пов'язані, інтерфейс забезпечує інтуїтивну взаємодію, а результати класифікації та звітності – високу інформативність. З урахуванням відкритої архітектури, реалізовану систему можна ефективно масштабувати. Потенційні напрями подальшого розвитку включають:

- перехід до україномовної або мультимовної моделі класифікації (наприклад, XLM-R або ukr-roberta), що дозволить уникнути перекладу та зберегти емоційні особливості мови;

- впровадження адаптивного підходу до виявлення симптомів – з використанням іменованих сутностей (NER), класифікації фраз або векторного пошуку схожих патернів;

- розширення GPT-функцій на основі історії повідомлень, що дозволить формувати динамічні пояснення з урахуванням попередніх станів;

- додавання режиму регулярного анкетування (наприклад, GAD-7) для структурованої оцінки тривожності у поєднанні з аналізом тексту;

- реалізацію персонального профілю користувача з можливістю фіксувати шаблони поведінки, налаштовувати інтерфейс та вибирати стиль GPT-відповідей;

- впровадження механізму зворотного зв'язку щодо коректності класифікації та якості реакції системи.

Запропоновані напрями не лише усувають поточні обмеження, а й відкривають можливості для створення інтелектуальної адаптивної системи підтримки емоційного благополуччя, що відповідає сучасним вимогам персоналізованої взаємодії.

ВИСНОВКИ

У межах цієї кваліфікаційної роботи розроблено інтелектуальну систему моніторингу тривожності, що здійснює автоматичний аналіз текстових повідомлень користувача з метою виявлення ознак тривожного стану, побудови емоційної динаміки та формування інтерпретованого зворотного зв'язку у зручному форматі.

У теоретичній частині було досліджено психологічну природу тривожності, її прояви в мовленні та поведінці, сучасні методи моніторингу психоемоційного стану, а також технічні підходи до класифікації емоційних текстів – від класичних моделей до трансформерних архітектур. Окрему увагу приділено генеративним мовним моделям як інструментам підтримки користувача в умовах емоційного напруження.

Для виявлення тривожних повідомлень використано модель roberta-base, яка була донавчена на збалансованому підкорпусі англійських прикладів із мітками «anxiety» та «normal». Модель досягла високих результатів класифікації (точність 99%, F1-мера 0.99) та продемонструвала стійкість до різноманітних текстових формулювань. Для українськомовного вводу реалізовано автоматичний переклад із використанням бібліотеки deep_translator.

Архітектура системи охоплює класифікаційний модуль, rule-based аналіз симптомів, генеративний компонент GPT, систему збереження історії в базі даних, механізм побудови графіків та генерацію звітів (включно з PDF-версією). Уся взаємодія реалізована через Telegram-бота, що підтримує кнопкову навігацію, режим емоційного щоденника, ручний вибір періоду та автоматичне формування аналітичного підсумку.

Проведене тестування підтвердило працездатність усіх компонентів. Система коректно реагує на повідомлення з різною емоційною насиченістю, формує персоналізовані пояснення до тривожних записів та дозволяє отримати динамічний звіт із візуалізацією змін стану. Окремі обмеження,

пов'язані з використанням англомовної моделі, rule-based підходом до симптомів і відсутністю персоналізації, були виявлені та детально проаналізовані.

Отримані результати демонструють доцільність застосування трансформерних моделей і генеративних технологій для задач м'якої психологічної підтримки. Система є прикладом ефективного інструменту самоспостереження, який може бути основою для подальших розробок у напрямі адаптивних цифрових асистентів емоційного добробуту.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. What's the Difference Between Anxiety and Fear? *Psychology Today*. URL: <https://www.psychologytoday.com/us/blog/escaping-our-mental-traps/202301/whats-the-difference-between-anxiety-and-fear> (date of access: 10.04.2025).
2. What are Anxiety Disorders? *Psychiatry.org – Home*. URL: <https://www.psychiatry.org/patients-families/anxiety-disorders/what-are-anxiety-disorders> (date of access: 10.04.2025).
3. Anxiety disorders. *World Health Organization (WHO)*. URL: <https://www.who.int/news-room/fact-sheets/detail/anxiety-disorders> (date of access: 10.04.2025).
4. Mental Health, Depression, and Anxiety Among Adults in the United States: 2022 Report. *Centers for Disease Control and Prevention | CDC*. URL: <https://www.cdc.gov/nchs/data/nhsr/nhsr213.pdf> (date of access: 13.04.2025).
5. Anxiety Symptoms: Comprehensive List of Physical & Emotional Signs. URL: <https://altitudecare.net/anxiety-symptoms-comprehensive-list-of-physical-emotional-signs> (date of access: 13.05.2025).
6. What is a Digital Footprint? Invisible Online Trails. *The Online Reputation and Wikipedia Blog*. URL: <https://blog.reputationx.com/digital-footprint> (date of access: 17.05.2025).
7. Preferred Self-Administered Questionnaires to Assess Depression, Anxiety and Somatization in People With Musculoskeletal Pain – A Modified Delphi Study / L. Bijker et al. *The Journal of Pain*. 2020. Vol. 21, no. 3-4. P. 409–417. URL: <https://doi.org/10.1016/j.jpain.2019.08.006> (date of access: 17.05.2025).
8. Ethical Dimensions of Digital Phenotyping Within the Context of Mental Healthcare – *Journal of Technology in Behavioral Science*. *SpringerLink*.

URL: <https://link.springer.com/article/10.1007/s41347-024-00423-9> (date of access: 17.05.2025).

9. An Empirical Study Comparing Unobtrusive Physiological Sensors for Stress Detection in Computer Work / Akbar et al. *Sensors*. 2019. Vol. 19, no. 17. P. 3766. URL: <https://doi.org/10.3390/s19173766> (date of access: 17.05.2025).

10. AI-Led Mental Health Support (Wysa) for Healthcare Workers During COVID-19: Service Evaluation (Preprint) / C. L. Chang et al. *JMIR Formative Research*. 2023. URL: <https://doi.org/10.2196/51858> (date of access: 21.05.2025).

11. Acceptability and Effectiveness of Artificial Intelligence Therapy for Anxiety and Depression (Youper): Longitudinal Observational Study / A. Mehta et al. *Journal of Medical Internet Research*. 2021. Vol. 23, no. 6. P. e26771. URL: <https://doi.org/10.2196/26771> (date of access: 21.05.2025).

12. Fitzpatrick K. K., Darcy A., Vierhile M. Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. *JMIR Mental Health*. 2017. Vol. 4, no. 2. P. e19. URL: <https://doi.org/10.2196/mental.7785> (date of access: 21.05.2025).

13. Effects of a self-guided transdiagnostic smartphone app on patient empowerment and mental health: A randomized controlled trial (Preprint) / A. Kerber et al. *JMIR Mental Health*. 2022. URL: <https://doi.org/10.2196/45068> (date of access: 25.05.2025).

14. Author Correction: Loneliness and suicide mitigation for students using GPT3-enabled chatbots / B. Maples et al. *npj Mental Health Research*. 2024. Vol. 3, no. 1. URL: <https://doi.org/10.1038/s44184-024-00055-0> (date of access: 25.05.2025).

15. Aydin A. Bag of Words Model in NLP. *Medium*. URL: <https://ayselaydin.medium.com/4-bag-of-words-model-in-nlp-434cb38cdd1b> (date of access: 28.05.2025).

16. Learn TF-IDF | Basic Text Models. *Codefinity: Courses with certificates* | *Online Learning Platform*.

URL: <https://codefinity.com/courses/v2/c68c1f2e-2c90-4d5d-8db9-1e97ca89d15e/b31ae769-ca29-4180-8a3d-375b8a0433ca/9ab6c781-18de-4ea4-97f0-0fd16f3f269e> (date of access: 28.05.2025).

17. Jain A. N-grams in NLP. *Medium*. URL: <https://medium.com/@abhishekjainindore24/n-grams-in-nlp-a7c05c1aff12> (date of access: 28.05.2025).

18. Deshpande V. Using Support Vector Machines (SVMs) to Develop Robust and Effective Chatbots. *Medium*. URL: <https://medium.com/@vdeshpande551/using-support-vector-machines-svms-to-develop-robust-and-effective-chatbots-f4061b06235b> (date of access: 29.05.2025).

19. The ultimate guide to Naive Bayes | Machine Learning Archive. *Machine Learning Archive*. URL: <https://mlarchive.com/machine-learning/the-ultimate-guide-to-naive-bayes/> (date of access: 01.06.2025).

20. Kulshrestha R. Transformers. *Medium*. URL: <https://medium.com/data-science/transformers-89034557de14> (date of access: 01.06.2025).

21. Text-to-Ontology Mapping via Natural Language Processing with Application to Search for Relevant Ontologies in Catalysis / L. Korel et al. *Computers*. 2023. Vol. 12, no. 1. P. 14. URL: <https://doi.org/10.3390/computers12010014> (date of access: 05.06.2025).

22. Papers with Code - RoBERTa MNLI. *The latest in Machine Learning | Papers With Code*. URL: <https://paperswithcode.com/lib/allennlp/roberta-mnli> (date of access: 05.06.2025).

23. Ranasinghe T., Zampieri M. Multilingual Offensive Language Identification with Cross-lingual Embeddings. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Online. Stroudsburg, PA, USA, 2020.

URL: <https://doi.org/10.18653/v1/2020.emnlp-main.470> (date of access: 05.06.2025).

24. Winastwan R. Text Generation with GPT. *Medium*.

URL: <https://medium.com/data-science/text-generation-with-gpt-092db8205cad> (date of access: 05.06.2025).

25. How Prompt Engineering AI Works: A Step-by-Step Guide. *Free Chatbot maker | Chatbot for Website, WhatsApp | BotPenguin*.

URL: <https://botpenguin.com/blogs/how-prompt-engineering-ai-works-a-step-by-step-guide> (date of access: 06.06.2025).

26. Sentiment Analysis for Mental Health. *Kaggle*.

URL: <https://www.kaggle.com/datasets/suchintikasarkar/sentiment-analysis-for-mental-health/data> (date of access: 06.06.2025).

27. FacebookAI/roberta-base Hugging Face. *Hugging Face – The AI community building the future*. URL: <https://huggingface.co/FacebookAI/roberta-base> (date of access: 06.06.2025).