УДК 004.85:159.9

ПОРІВНЯЛЬНИЙ АНАЛІЗ ВЕЛИКИХ МОВНИХ МОДЕЛЕЙ ДЛЯ ЧАТ-АСИСТЕНТА У СФЕРІ МЕНТАЛЬНОГО ЗДОРОВ'Я

Добриніна А.О.
e-mail: anastasiia.dobrynina@nure.ua
Науковий керівник – к.т.н., доц. Білова Т.Г.
Харківський національний університет радіоелектроніки, каф. КМІТ
м. Харків, Україна

This work presents a comparative analysis of modern large language models for the development of a chatbot assistant in the field of mental health. The study examines the ability of LLMs to interpret user queries expressed in natural language and generate responses that assist users in selecting appropriate mental health specialists. Several leading models are compared according to criteria relevant for dialogue systems, including general intelligence, instruction-following ability, response style and naturalness, and API cost efficiency. Based on this analysis, a suitable model is selected for implementing a prototype chatbot system designed to assist users in navigating mental health services and identifying appropriate specialists.

Проблема доступності кваліфікованої психологічної допомоги залишається актуальною у світі. За даними ВООЗ, понад 1.1 мільярда людей страждають від психічних розладів [1], однак значна частина з них не звертається по допомогу через недостатню поінформованість про відповідних фахівців. Особливо актуальною вона є для України, де внаслідок війни зросла потреба у психологічній підтримці населення.

Впровадження інтелектуальних систем для вирішення проблеми ускладняється тим, що користувачі описують свій стан у вільній формі без використання медичної термінології, що потребує коректної інтерпретації змісту повідомлення. Крім того, використання систем на основі штучного інтелекту (ШІ) у сфері охорони здоров’я вимагає врахування етичних аспектів. Тому подібні системи повинні функціонувати як Decision Support Systems і не встановлювати клінічних діагнозів, а допомагати користувачу зорієнтуватися у виборі відповідного фахівця.

Чат-асистент на основі великих мовних моделей (LLM) дозволяє автоматизувати первинний аналіз запитів користувачів і підбір відповідного спеціаліста. Основними задачами системи є інтерпретація текстових повідомлень користувача, розмежування компетенцій психолога і психіатра та формування інформаційної відповіді з поясненням можливого психоемоційного стану і рекомендацією щодо звернення до відповідної категорії фахівців.

Метою роботи є проведення порівняльного аналізу сучасних LLM для вибору оптимального рішення при розробці чат-асистента у сфері ментального здоров’я. Практичним контекстом дослідження є система

автоматизованого підбору спеціалістів на основі скарг користувача. Оскільки архітектура системи передбачає використання LLM-компонентів – для інтерпретації запиту та генерації відповіді – вибір моделі безпосередньо впливає на точність класифікації, якість відповідей та безпечність взаємодії з користувачем.

Станом на початок 2026 року серед LLM-моделей виокремлюються Gemini 3 Pro, GPT-5.4 та Claude Opus 4.6, які посідають перші позиції за зведеним індексом інтелекту. Ці моделі належать до так званих frontier-моделей – найбільш потужних систем ШІ, що демонструють високі результати у задачах обробки природної мови, логічного міркування та генерації тексту. Саме тому ці три моделі були обрані як об'єкти порівняльного аналізу: вони представляють провідні розробки трьох найбільших компаній у сфері ШІ – Anthropic, Google та OpenAI – і є доступними для комерційного використання через публічний API. Кожна з них активно використовується у реальних продуктах і має задокументовані показники продуктивності, що робить їх порівняння обґрунтованим і практично значущим.

Для визначення найбільш придатної моделі для реалізації чат-асистента проведено порівняльний аналіз за чотирма критеріями: рівень загального інтелекту моделі (MMLU) [2], здатність до виконання інструкцій (IFEval) [2], стиль і природність відповідей у діалозі з користувачем та економічна ефективність використання API. Обрані критерії відповідають ключовим вимогам до чат-асистентів: коректній інтерпретації запитів, виконанню системних інструкцій та природності діалогу. Порівняльний аналіз характеристик розглянутих моделей наведено в таблиці 1.

Таблиця 1 – Порівняння характеристик великих мовних моделей

Критерій	Anthropic Claude Opus 4.6	Google Gemini 3 Pro	OpenAI GPT-5.4
Загальний інтелект (MMLU)	≈91% [3]	≈91.8% [4]	≈91% [5]
Виконання інструкцій (IFEval)	≈94% [3]	≈85% [4]	≈92% [5]
Стиль відповідей	Більш природний, експресивний стиль, високий рівень адаптації до користувача [3]	Більш аналітичний та структурований стиль [4]	Більш сухий та академічний стиль, зосереджений на технічній точності [5]
Ціна (за 1 млн токенів)	≈$5 (input) / $25 (output) [3]	≈$2 (input) / $12 (output) [4]	≈2.50 (input)/ 15.00 (output) [5]

Таким чином, усі розглянуті моделі є придатними для реалізації чат-асистента у сфері ментального здоров'я, однак мають різні профілі

переваг та особливості використання у прикладних діалогових системах. Claude Opus 4.6 вирізняється більш природним стилем відповідей та адаптацією до користувача, GPT-5.4 демонструє високий рівень виконання інструкцій і точність генерації відповідей, тоді як Gemini 3 Pro забезпечує конкурентний рівень загального інтелекту та вигідне співвідношення якості і вартості використання API.

Водночас жодна з моделей не має абсолютної переваги за всіма критеріями, що обумовлює доцільність їх комбінованого використання залежно від конкретного завдання системи. Вибір моделі або їх поєднання визначається насамперед пріоритетами проєктування: якістю діалогу, точністю класифікації запитів або економічною ефективністю розгортання.

З урахуванням наведених у таблиці 1 характеристик для реалізації прототипу системи доцільно обрати модель Gemini 3 Pro. У запропонованій архітектурі система може використовувати каскадний підхід: для інтерпретації запиту користувача та визначення можливого психоемоційного стану доцільно застосовувати модель Gemini. На етапі формування текстової відповіді система також може використовувати Gemini, а за потреби – модель Claude, яка вирізняється більш природним стилем генерації. Такий підхід дозволяє підвищити якість діалогу з користувачем, зберігаючи простоту інтеграції системи.

Отже, проведений порівняльний аналіз показав, що моделі Claude, GPT та Gemini демонструють подібний рівень інтелектуальних можливостей, однак відрізняються за виконанням інструкцій, стилем відповідей та вартістю використання API. Для реалізації чат-асистента у сфері ментального здоров'я базовою моделлю доцільно обрати Gemini 3 Pro, яка забезпечує достатній рівень інтелектуальних можливостей та економічну ефективність використання. Ефективність роботи системи також значною мірою залежить від коректного формування промпта та контексту взаємодії з моделлю.

Список використаних джерел:

1. Mental disorders : вебсайт. URL: https://www.who.int/news-room/fact-sheets/detail/mental-disorders (дата звернення 11.03.2026).

2. Open LLM Leaderboard – Hugging Face : вебсайт. URL: https://huggingface.co/docs/leaderboards/en/open_llm_leaderboard/about (дата звернення 11.03.2026).

3. Claude Opus 4.6 Model Overview : вебсайт. URL: https://automatio.ai/uk/models/claude-opus-4-6 (дата звернення 11.03.2026).

4. Gemini 3 Pro Model Overview : вебсайт. URL: https://automatio.ai/uk/models/gemini-3-pro (дата звернення 11.03.2026).

5. GPT-5.4 Model Overview : вебсайт. URL: https://automatio.ai/uk/models/gpt-5-4 (дата звернення 11.03.2026).