

АНАЛІЗ РОБОТИ З ШТУЧНИМИ ІНТЕЛЕКТАМИ

Д.Є. Проценко

Харківський національний університет радіоелектроніки

Україна, 61166, Харків, пр. Науки 14

E-mail: dmytro.protsenko@nure.ua

Анотація: У роботі основну увагу було приділено аналізу, порівнянню штучних інтелектів та підключенню безкоштовної моделі у свій проект. Було розглянуті популярні моделі ШІ, порівняно функціонал та популярність серед користувачів.

Ключові слова: штучний інтелект, ШІ, підключення.

COMPARISON OF METHODS OF INTERACTION WITH ASSISTANTS

D. Protsenko

Kharkiv National University of Radio Electronics

Ukraine, 61166, Kharkiv, Nauky av., 14

E-mail: dmytro.protsenko@nure.ua

Annotation: The article focused on on the analyzing, comparing artificial intelligences, and integrating a free model into your project. Popular AI models were reviewed; their functionality and popularity among users were compared.

Key words: artificial intelligence, AI, connectivity.

На сьогоднішній день існує багато різновидів штучного інтелекту, які застосовуються у різних сферах. Наприклад найбільш популярний ChatGPT є чат-ботом, який працює на базі generative pre-trained transformer – тобто він навчається за допомогою великої кількості інформації зі статей, книг, інтернета та інших ресурсів. Він може аналізувати питання і шукати відповідь з різних джерел.

ChatGPT я і інші ШІ приймають та виводять інформацію за допомогою токенів. Це працює таким чином: коли користувач пише запит у чат, ШІ розбиває його на токени – фрагменти тексту, які можуть бути цілим словом, частиною слова або розділовим знаком. Таким чином, модель «розбиває» речення на токени, аналізує їх і потім і шукає відповідь. Знайшовши результат, він виводить його тільки в зворотньому порядку: токени аналізуються, будується речення і виводиться відповідь. Зазвичай кількість токенів є обмеженою, наприклад модель GPT-3.5 має пам'ять на 4 – 16 тисяч токенів, а модель GPT-4 має пам'ять до 128 тисяч токенів. Можемо помітити, що кількість токенів стало більшим у 8 раз. Це пов'язано із тим що GPT-4 — це новіша, потужніша архітектура, яка має більше пам'яті й ефективніше працює з токенами.

Через швидкий розвиток сфери штучного інтелекту, великі компанії почали створювати власні ШІ для виконання певних робіт через гнучкість: Grok, Gemini, Copilot, Claude, DeepSeek – такі самі чат боти як і раніше згаданий ChatGPT, але кожний такий чат має певні обмеження та налаштування на відповідь через закони країн, політичний стан і моральні принципи. ШІ від Nvidia створювався не для спілкування, а для роботи із рендером зображення у реальному часі, Midjourney – для створення зображень з тексту, Suna – для створення музики і пісень і багато інших. Також різні розробники створюють ШІ для обробки зображень з камери і звуку з мікрофону. Наприклад настільний робот ЕМО від LivingAI (рис. 1)[5] має вбудований нейромережевий процесор, який відповідає за аналіз емоцій користувача і його голос. Завдяки цим функціям, робот може підлаштовуватись під настрій власника, реагувати на тон і гучність голосу, бути більш гнучким у поведінці ніж звичайні запрограмовані роботи на певний перелік команд.



Рисунок 1 – Робот ЕМО від LivingAI

За аналізом сайту Artificial Analysis (сайт який займається порівнянням актуальних ШІ), станом на жовтень 2025, маємо результати зображені на графіках (рис.2–3)[2].

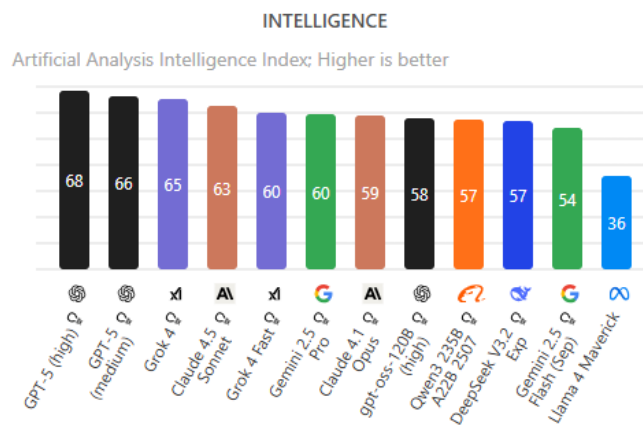


Рисунок 2 – Результати порівняння ШІ за інтелектуальним індексом

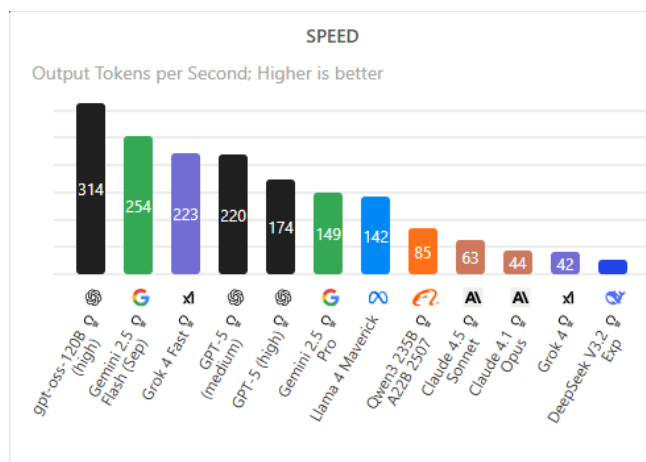


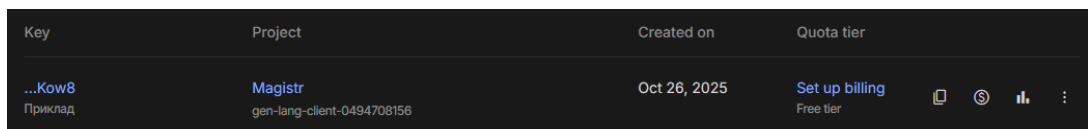
Рисунок 3 – Результати порівняння ШІ за швидкістю

Роблячи підсумки цих результатів, можемо зрозуміти, що за результатами порівняння рівня інтелектуального індексу (рис.2) 1 та 2 місце займає дві моделі ChatGPT 5 версії, а 3 місце – Grok 4. Проте, якщо брати аналіз швидкості генерації відповіді (рис.3), тоді маємо такі результати: 1 місце – ChatGPT gpt-oss-mini 120B, 2 місце – Gemini 2.5 Flash (модель, орієнтована на швидкі відповіді із збереженням їх високої якості), 3 місце – Grok 4 Fast (пришвидшена

модель Grok 4, як і Gemini 2.5 Flash). Також є багато інших показників чат-ботів які можна порівняти, але через їхній швидкий розвиток результати аналізу постійно змінюються.

Під час аналізу результатів порівнянь ШІ, згадувалась модель ChatGPT gpt-oss-mini 120B. Gpt-oss – серія відкритих моделей від компанії Open AI, яка дозволяє модифікувати і використовувати модель у власних роботах. Завдяки цьому, різні компанії можуть додавати штучний інтелект до своїх проєктів у комерційних цілях. Проте швидкість обробки запитів може бути низькою через обмеження виводу токенів для економії.

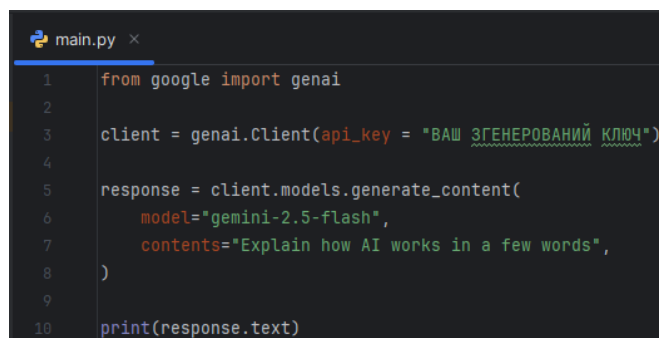
Деякі компанії також надають можливість ознайомитись із штучним інтелектом у власних проєктах. Однією з таких компаній є Google яка надає можливість згенерувати API ключ для під'єднання моделі штучного інтелекту до проєкту. Якщо зайти на їх сторінку із налаштуваннями підключень ШІ до проєктів, можна знайти усі необхідні інструкції. Наприклад, нам треба підключити Google Gemini, модель gemini-2.5-flash. По-перше, створюємо API ключ, вводим назву проєкту де буде застосовуватись згенерований ключ. В результаті отримаємо пункт, де написана уся інформація про створений ключ (рис. 4). Варто зазначити, що у розділі Quota tier написано, що ми маємо безкоштовний план користування. Це означає, що ми можемо тестувати модель у проєкті, проте через велику кількість запитів, відповіді можуть обмежуватись по швидкості або по кількості.



Key	Project	Created on	Quota tier
...Kow8 Приклад	Magistr gen-lang-client-0494708156	Oct 26, 2025	Set up billing Free tier

Рисунок 4 – Інформація про створений API ключ

Наступним етапом підключення ШІ є встановлення необхідної бібліотеки. Google Gemini підтримує такі мови програмування як: Python, JavaScript, GO, Java, REST. Розберемо підключення на прикладі мови програмування Python. У терміналі на пристрої прописуємо бібліотеку для встановлення, а саме: `pip install google-gemai`. Завдяки ній, ми зможемо працювати із моделлю у нашому проєкті. Далі, необхідно під'єднати і налаштувати модель безпосередньо у проєкт. На сайті вже є приклад коду, в якому можемо задавати питання асистенту і отримувати відповідь, для цього просто необхідно додати наш згенерований API ключ (рис. 5).



```
main.py x
1 from google import genai
2
3 client = genai.Client(api_key = "ВАШ ЗГЕНЕРОВАНИЙ КЛЮЧ")
4
5 response = client.models.generate_content(
6     model="gemini-2.5-flash",
7     contents="Explain how AI works in a few words",
8 )
9
10 print(response.text)
```

Рисунок 5 – Приклад коду підключення моделі

Завдяки цьому простому прикладу коду ми можемо перевірити, чи працює модель у нашому проєкті. На запит "Explain how AI works in a few words", що перекладається як "Поясни як працює ШІ у декількох словах", наша програма дає коротку відповідь: "AI learns patterns from data to make decisions or predictions." - "Штучний інтелект вивчає закономірності з даних, щоб приймати рішення або робити прогнози." Але через гнучкість, нашу програму можна налаштувати як завгодно: зробити чат-бота у реальному часі або голосового асистента як у кіно, обмежити кількість токенів для економії і багато іншого.

Таким чином, компанії надають можливість ознайомитись із моделями ШІ не тільки як користувач, але і як розробник: під'єднати до власного проєкту, розширити функціонал, додати

інтерактивності і т.д. До того ж, все більше корпорацій надає пробний безкоштовний доступ до своїх ШІ для ознайомлення, а вже потім користувачі приймають рішення, платити за них чи ні.

ВИСНОВКИ. За результатами аналізу ринку штучних інтелектів від різних компаній можна зрозуміти, що наразі існує великий вибір моделей, які націлені під різні задачі. Через швидкий розвиток сфери, порівняння швидкості, якості і енергоефективності може постійно змінюватись, через актуальність ринку і його постійне вдосконалення, тому що компанії розуміють наскільки це потужний інструмент. ШІ може не просто відповідати на запитання, він може генерувати медіа, прискорювати обробку даних, робити її більш злагодженою.

До того ж, варто зазначити, що деякі великі корпорації надають можливість безкоштовно ознайомитись із моделями і використати їх у власних проектах. Завдяки такому підходу, різні невеликі студії, розробники або ентузіасти можуть навчитися використовувати цей інструмент у власних розробках.

ЛІТЕРАТУРА

1. Google AI Studio. Google AI Studio. URL: <https://aistudio.google.com/?project=gen-lang-client-0494708156>.
2. Comparison of AI Models across Intelligence, Performance, Price | Artificial Analysis. AI Model & API Providers Analysis | Artificial Analysis. URL: <https://artificialanalysis.ai/models>.
3. The Complete Guide: How to Use ChatGPT API in Application. UI UX Design Agency for SaaS, Fintech & AI | Adam Fard UX Studio. URL: <https://adamfard.com/blog/how-to-use-chatgpt-api#:~:text=Educational%20Tools:%20Educators%20and%20e,provide%20an%20engaging%20learning%20experience>.
4. Gemini API | Google AI for Developers. Google AI for Developers. URL: <https://ai.google.dev/gemini-api/docs>.
5. EMO - LivingAI. LivingAI. URL: <https://living.ai/emo/>.
6. openai/gpt-oss-120b · Hugging Face. Hugging Face – The AI community building the future. URL: https://huggingface.co/openai/gpt-oss-120b?utm_source=chatgpt.com.
7. FaTPay. Understanding Token Limits in OpenAI's GPT Models. Medium. URL: <https://medium.com/coinmonks/understanding-token-limits-in-openais-gpt-models-37fbe67c89f4>.
8. GitHub - openai/gpt-oss: gpt-oss-120b and gpt-oss-20b are two open-weight language models by OpenAI. GitHub. URL: <https://github.com/openai/gpt-oss>.
9. Nevliudov, I., Yevsieiev, V., Maksymova, S., Gopejenko, V., & Kosenko, V. (2025). Development of mathematical support for adaptive control for the intelligent gripper of the collaborative robot manipulator. *Advanced Information Systems*, 9(3), 57-65.
10. Невлюдов, І., Клименко, О., Євсєєв, В., & Максимова, С. (2025). IMPROVEMENT OF THE ENCODING INFORMATION METHOD FOR PHARMACEUTICAL PRODUCTS QR-CODES DURING SORTING ON A ROBOTIC CONVEYOR LINE. *Вісник Національного технічного університету «ХПІ»*. Серія: Технології в машинобудуванні, (1 (11)), 128-134.
11. Невлюдов, І., Євсєєв, В., Максимова, С., & Артюх, Р. (2025). Математична модель адаптивного ієрархічного високорівневого керування триланкового колаборативного робота-маніпулятора. *Сучасний стан наукових досліджень та технологій в промисловості*, (2 (32)), 58-68.
12. Demska, N., Yevsieiev, V., Maksymova, S., & Ababneh, J. (2025). DECISION-MAKING MODEL FOR CONTROLLING A COLLABORATIVE ROBOT-MANIPULATOR BASED ON THE SENSOR FUSION METHOD AND CNN APPROACH TO RULE FORMATION. *Multidisciplinary Journal of Science and Technology*, 5(6), 846-859.

Науковий керівник: Бронніков Артем Ігорович, к.т.н., доцент кафедри КІТАР.