

ChatGPT У ПОРІВНЯННІ З АНАЛОГАМИ

Вінницький національний технічний університет

Анотація:

Публікація присвячена дослідженню можливостей аналогів ChatGpt в області підтримки корисної людино-машинної комунікації. Проведено трьох етапний експеримент, який дозволив визначити переваги та недоліки ChatGpt у порівнянні з відомими аналогами Perplexity, ChatSonic та Bart для різних типів завдань.

Ключові слова: ChatGpt, Perplexity, ChatSonic, Bart

Annotation:

The paper is devoted to the study of the capabilities of ChatGpt analogs in the field of supporting useful human-machine communication. A three-stage experiment was conducted to determine the advantages and disadvantages of ChatGpt in comparison with well-known analogs such as Perplexity, ChatSonic, and Bart for different types of tasks.

Keywords: ChatGpt, Perplexity, ChatSonic, Bart

Вступ

Розвиток методів штучного інтелекту створює безліч можливостей для взаємодії між людьми та комп'ютерами. Одним із суттєвих досягнень в цій області є модель генерації тексту ChatGPT (Generative Pre-trained Transformer), розроблена компанією OpenAI. ChatGPT вражає своєю здатністю створювати змістовне та природне мовлення, що наближається до рівня, здатного заперечити опоненту. Іншими словами, підтримується майже повноцінна розмова із штучним інтелектом, як з живою людиною. Однак, чи є аналоги цієї моделі? Чи існують інші системи, які можуть забезпечити подібну якість та продуктивність? У даному дослідженні буде розглянуто три відомі аналоги ChatGPT: ChatSonic, Perplexity та Bart. Кожен з них має свої унікальні особливості, які варто дослідити та порівняти.

Розуміння цих аналогів допоможе нам розширити перспективу відповідних досліджень та відкрити нові можливості у сфері розробки інтелектуальних діалогових систем. Задача дослідження полягає у порівнянні обраних аналогів та визначенні, як вони співвідносяться з ChatGPT у контексті генерації тексту та розв'язання формальних завдань.

Основна частина

Все більше компаній та дослідників займаються розробкою та вдосконаленням моделей штучного інтелекту для генерації тексту та моделювання діалогів. Одним із найвідоміших представників цієї області є ChatGPT, який вражає своєю здатністю створювати природні та змістовні відповіді. Однак, науковці та розробники продовжують шукати аналоги, які можуть покращити ці моделі та відкрити нові можливості. У даній частині роботи буде розглянуто три таких аналоги: ChatSonic, Perplexity та Bart. Кожен з них має свої унікальні особливості та можливості, які варто вивчити та порівняти з ChatGPT. Аналіз цих аналогів допоможе нам краще зрозуміти сучасні тенденції у розробці діалогових систем та відкрити нові горизонти для подальшого вдосконалення штучного інтелекту.

ChatSonic – це один з найближчих аналогів ChatGPT, розроблений з метою покращення генерації тексту та відтворення природного мовлення. Він використовує схожі архітектурні принципи та методи навчання, що дозволяють йому досягти високої якості та зрозумілості в діалогах з користувачами. ChatSonic відзначається здатністю створювати змістовні відповіді та вірно розуміти контекст, що дозволяє покращити якість діалогу та забезпечити більш природну взаємодію [1].

Perplexity виступає не як окрема модель, а як метрика оцінки якості моделей генерації тексту. Ця метрика вимірює здатність моделі передбачати наступне слово в тексті. Чим менше значення перплексії,

тим краще модель встигає передбачити наступне слово, що свідчить про її високу якість та здатність генерувати змістовний текст. Використання перплексії дозволяє порівнювати різні моделі та оцінювати їхню ефективність в генерації тексту [2].

Bart (Bidirectional and AutoRegressive Transformers) є ще одним аналогом ChatGPT, який поєднує принципи двостороннього та авторегресивного моделювання. Його архітектура дозволяє краще моделювати контекст та залежності між словами, що сприяє покращенню зрозумілості та якості згенерованого тексту у діалогових системах. Bart володіє вражаючими можливостями у створенні змістовного та природного мовлення, що дозволяє використовувати його для різних сфер застосування, включаючи автоматизовану підтримку клієнтів та генерацію вмісту [3].

В процесі проведення експериментів було здійснено покрокове дослідження кожного з аналогів ChatGpt. Отож, розглянемо отримані результати для кількох запланованих для експерименту завдань.

Самим першим завданням ChatGpt та його аналогів було написання есе на складну тему: «Ієрархія потреб за Маслоу – чи актуальна ще в 21 столітті». Варто відмітити, що всі результати на формальному рівні не були правильними, оскільки ні один з аналогів не написав структурного есе, що відповідає стандартним вимогам до побудови. Було написано звичайний твір, в якому майже не була розкрита тема. ChatGPT написав величезний твір, в якому однакова інформація була перефразована по декілька разів, але основна тема так і не була розкрита. ChatSonic написав приблизно такий самий твір, як і ChatGpt, але з меншою кількістю інформації, що повторювалась. В Bart взагалі виникла проблема з написанням. Останній аналог – Perplexity – найкраще впорався із завданням, оскільки написав більш-менш правильне есе, в якому не було зайвої інформації, а також було дотримано деякі стандарти есе. Проте, в усіх перерахованих випадках глибина викладення психологічного вмісту ієрархії потреб за Маслоу виявилася мінімальною, тобто придатною лише для поверхневого знайомства з обраною тематикою.

Далі було проведено перевірку отриманих текстів на помилки та плагіат, причому результат цього виявився досить непоганий: кожне есе було згенероване індивідуально для згаданого запиту, тому перевірка на плагіат показала, що копірайту не було знайдено. Перевірка на помилки також показала досить непоганий результат – було знайдено в кожному есе по одній помилці. Отож можна припустити, що чати впоралися зі своїм завданням на 3 з 5.

Наступним завданням було створення програми на мові C#, що має обраховувати детермінант матриці. При проведенні аналізу згенерованого коду Chat'ом Gpt, можна сказати, що це був досить масивний код, який потребував певного форматування, але при перевірці його у середовищі Visual Studio якихось дефектів виявлено не було. ChatSonic та Bart видали схожі програмні коди, які також працювали без помилок. Останнім аналогом, що генерував код, був Perplexity – саме він показав найкращий результат, оскільки згенерував невеликі дві програми, одна з яких працювала на бібліотеці «MathNet», що дає змогу використовувати стандартний код для обрахування детермінанту матриці.

Загалом можна зробити висновок, що ChatGpt та його аналоги досить непогано впоралися зі своїм завданням та згенерували відносно непогані програми, які працювали без будь-яких помилок (наша оцінка – 4 з 5).

Останнім завданням для тестування чатів було обрано математичні задачі. Зокрема запропоновано не дуже складне завдання для обрахування невизначеного інтегралу (рис. 1). Даний пункт можна визнати, як повний провал цього завдання, оскільки жоден із чатів не видав з першого разу правильної відповіді, а Bart взагалі не зміг упоратися із завданням.

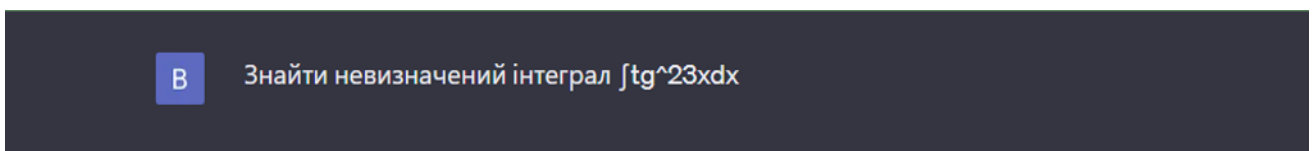


Рисунок 1 – Невизначений інтеграл

Початковим був звичайний запит для обрахування математичного прикладу, але стало зрозумілим, що такий запит не є досить точним та потребує редагування. Внаслідок уточнення та переформулювання запиту результат покращився: Perplexity видав правильну відповідь з першого разу. Тому потрібно визнати, що чати погано можуть упоратися з математичними задачами, отже потрібно сформулювати дуже точний запит для обрахування якогось нескладного завдання.

Висновки

Аналізуючи проведене дослідження, варто відзначити, що Perplexity виконував запропоновані завдання набагато краще, ніж ChatGpt. Взяти до прикладу час виконання запиту: якщо ChatGPT генерує відповідь на запит приблизно більше хвилини, Perplexity витрачає на це максимум 20 секунд, причому наводить точнішу інформацію, у якій менше повторень та помилок. ChatSonic при цьому майже повністю копіював ChatGpt, але генерував відповіді на запити швидше. Нарешті Bart проявив себе в проведеному експерименті найгірше, оскільки на 2 із трьох запитів не дав відповіді, окрім того він працює тільки через VPN та на англійській мові.

За результатами даної роботи можна зробити загальний висновок – для отримання більш коректної відповіді потрібно генерувати досить точні запити, перепитувати: «Чи це дійсно правильна відповідь», при цьому перевіряти результати власноруч, або за допомогою тих сервісів, в яких ви впевнені. Чи варта «така гра витрачених свічок» залишається риторичним питанням.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

- 1.Електронне джерело «YouTube». URL: https://www.youtube.com/watch?v=Gw6utyL5bec&ab_channel=%D0%A0%D0%BE%D0%BC%D0%B0%D0%BD%D0%9A%D0%BE%D0%BD%D0%B4%D1%80%D0%B0%D1%82%D1%8C%D0%B5%D0%B2
- 2.Електронне джерело «YouTube». URL: https://www.youtube.com/watch?v=8xyL2GHzWkA&ab_channel=Web3nity
- 3.Електронне джерело «YouTube». URL: [youtube.com/watch?v=h6rkVXxqtbk&ab_channel=HinkoK](https://www.youtube.com/watch?v=h6rkVXxqtbk&ab_channel=HinkoK)

Моїк Ігор Іванович – студент кафедри автоматизації та інтелектуальних інформаційних технологій, Вінницький національний технічний університет, м. Вінниця, e-mail: cc284692@gmail.com

Бісікало Олег Володимирович – доктор техн. наук, професор, завідувач кафедри АІТ, факультет інтелектуальних інформаційних систем та автоматики, Вінницький національний технічний університет, Вінниця, e-mail: obisikalo@vntu.edu.ua

Бісікало Лариса Георгіївна – практичний психолог, Агрономіченський ліцей Агрономічної сільської ради, Вінницький район, Вінницька область, e-mail: Igbisikalo@gmail.com

Ihor Moik – student of the Department of Automation and Intelligent Information Technologies, Vinnytsia National Technical University, Vinnytsia, e-mail: cc284692@gmail.com

Oleh Bisikalo – Doctor of Engineering Sciences, Full Professor, Head of Department of Automation & Intelligent Information Technologies, Vinnytsia National Technical University, Vinnytsia, e-mail: obisikalo@vntu.edu.ua

Larysa Bisikalo – practical psychologist, Agronomics' Lyceum of the Agronomic Village Council, Vinnytsia district, Vinnytsia region, e-mail: Igbisikalo@gmail.com