

УДК 004.942

DOI: <https://doi.org/10.53920/ITS-2022-2-2>

Олександр Іванович ГОЛУБЕНКО,

канд. техн. наук, доцент
ЗВО «Міжнародний науково-технічний університет
імені академіка Юрія Бугая»
ORCID ID: 0000-0002-1776-5160

Олександр Олександрович ПІДМОГИЛЬНИЙ,

аспірант
Державний університет телекомунікацій
ORCID ID: 0000-0001-8689-2086

GENERATIVE PRE-TRAINED TRANSFORMER 3

GPT (Generative Pre-training Transformer) — це тип штучного інтелекту (AI), який використовує алгоритми машинного навчання для створення тексту природною мовою. Перша версія GPT, випущена в 2018 році, стала революційним досягненням у сфері ШІ та обробки природної мови (NLP). Однак він також мав деякі обмеження та проблеми, які були розглянуті в наступних версіях моделі.

Однією з головних проблем першої версії GPT була відсутність контролю над контентом, який вона генерувала. Модель було навчено на великому наборі даних тексту, створеного людиною, і вона змогла створити зв'язний і, здавалося б, людиноподібний текст на широкий спектр тем. Однак він часто створював текст, який був упередженим, образливим або іншим чином недоречним, оскільки він не міг повністю зрозуміти контекст або значення використаних слів.

Іншою проблемою першої версії GPT була її нездатність виконувати складніші завдання NLP, такі як переклад або конспектування. Хоча він міг створити зв'язний текст, він не міг зрозуміти значення чи структуру тексту так, як це може зробити людина.

Подальші версії GPT, такі як GPT-2 і GPT-3, вирішували ці проблеми та додавали нові можливості, такі як здатність виконувати складніші завдання NLP і генерувати більш зв'язний і відповідний контексту текст. Однак вони все ще мають обмеження і можуть давати необ'єктивні або невідповідні результати, якщо не використовувати їх відповідально.

Ключові слова: штучний інтелект (AI), машинне навчання, обробка природної мови (NLP), генеративний передтрениувальний трансформатор (GPT), генерація тексту, глибоке навчання, нейронна мережа.

Oleksandr GOLUBENKO

Candidate of technical sciences, associate professor
IHE «Academician Yuri Bugay
international science and technical university»
ORCID ID: 0000-0002-1776-5160

Oleksandr PIDMOGYLNYI

Postgraduate
State University of Telecommunications
ORCID ID: 0000-0001-8689-2086

GENERATIVE PRE-TRAINED TRANSFORMER 3

GPT (Generative Pre-training Transformer) is a type of artificial intelligence (AI) that uses machine learning algorithms to generate text in natural language. The first version of GPT, released in 2018, was a revolutionary breakthrough in AI and natural language processing (NLP). However, it also had some limitations and issues that were addressed in subsequent versions of the model.

One of the main problems with the first version of GPT was the lack of control over the content it generated. The model was trained on a large dataset of human-generated text and was able to generate coherent and seemingly human-like text on a wide range of topics. However, he often produced text that was biased, offensive, or otherwise inappropriate because he could not fully understand the context or meaning of the words used.

Another problem with the first version of GPT was its inability to handle more complex NLP tasks such as translation or annotation. Although he could produce coherent text, he could not understand the meaning or structure of the text as a human could.

Later versions of GPT, such as GPT-2 and GPT-3, addressed these issues and added new capabilities, such as the ability to perform more complex NLP tasks and generate more coherent and context-appropriate text. However, they still have limitations and can produce biased or inconsistent results if not used responsibly.

Keywords: artificial intelligence (AI), machine learning, natural language processing (NLP), generative pretraining transformer (GPT), text generation, deep learning, neural network.

Постановка проблеми. Generative Pre-trained Transformer 3 (GPT-3) — це модель штучного інтелекту (AI) для генерації мови,

розроблена OpenAI, яка привернула значну увагу як ЗМІ, так і технічної спільноти. Маючи 175 мільярдів параметрів, GPT-3 наразі є найбільшою та найпотужнішою мовною моделлю з існуючих, а її можливості виходять далеко за рамки простого генерування тексту, що потребує детального аналізу та досліджень.

Аналіз останніх досліджень і публікацій. GPT-3 — це тип мовної моделі на основі Transformer це означає, що він використовує архітектуру transformer для обробки та генерації тексту. Архітектура трансформатора базується на ідеї самоуважності, що дозволяє моделі обробляти вхідні послідовності паралельно, а не послідовно. Це дає змогу моделі фіксувати довгострокові залежності та зв'язки між словами в реченні, що важливо для створення зв'язного тексту.

На додаток до самоуважності, мовні моделі на основі трансформаторів також використовують інші методи, такі як моделювання замаскованої мови та динамічне керування. Моделювання замаскованої мови передбачає маскування частини вхідного тексту та передбачення відсутніх слів на основі контексту, наданого незамаскованими словами. Це допомагає моделі навчитися розуміти зв'язки між словами та створювати текст, який є зв'язним і має сенс у контексті.

Динамічний контроль, з іншого боку, дозволяє моделі регулювати рівень деталізації та конкретності вихідних даних на основі вхідних даних. Це дозволяє моделі створювати текст, який підходить для даного контексту та завдання.

Мета статті — дослідження технічних особливостей GPT-3, аналіз потенційного застосування та впливу, а також розгляд деяких етичних проблем, пов'язаних з його використанням.

Виклад основного матеріалу. Архітектура трансформатора була представлена в оригінальній моделі GPT, але GPT-3 робить крок далі, використовуючи попередньо навчену версію моделі. Попереднє навчання передбачає навчання моделі на великому наборі даних і подальше її тонке налаштування для конкретного завдання, наприклад перекладу мови або відповідей на запитання.

Однією з ключових особливостей GPT-3 є його здатність генерувати зв'язний і схожий на людину текст. Це досягається за рахунок використання механізмів уваги, які дозволяють моделі враховувати контекст і зв'язки між словами в реченні. GPT-3 також має можливість виконувати широкий спектр мовних завдань,

включаючи переклад, узагальнення та відповіді на запитання, і навіть може виконувати завдання програмування [1].

Архітектура трансформатора, що використовується в GPT-3, базується на ідеї самоуважності, яка дозволяє моделі обробляти вхідні послідовності паралельно, а не послідовно. Це дозволяє моделі фіксувати довгострокові залежності та зв'язки між словами в реченні, що важливо для створення зв'язного тексту.

GPT-3 також використовує техніку під назвою масковане моделювання мови, яка передбачає маскування частини вхідного тексту та передбачення відсутніх слів на основі контексту, наданого незамаскованими словами. Щоб реалізувати масковане моделювання мови, частина вхідного тексту вибирається випадковим чином і замінюється спеціальним маркером, таким як «[MASK]». Потім модель навчається передбачати пропущене слово або слова на основі контексту, наданого незамаскованими словами. Наприклад, враховуючи вхідні дані «[МАСКА] сиділа на килимку», модель буде навчена передбачати пропущене слово «кіт» на основі контексту, наданого іншими словами в реченні. Це допомагає моделі навчитися розуміти зв'язки між словами та створювати текст, який є зв'язним і має сенс у контексті [2].

На додаток до цих методів GPT-3 також використовує техніку, яка називається динамічним керуванням, яка дозволяє моделі регулювати рівень деталізації та конкретності вихідних даних на основі вхідних даних. Модуль динамічного керування може бути реалізований за допомогою різноманітних методів, таких як Transformer encoder або окрема нейронна мережа прямого зв'язку [5]. Transformer encoder в модулі керування обробляє вхідні дані та генерує керуючий сигнал, який використовується для налаштування рівня деталізації та конкретності на виході.

Динамічний контроль дозволяє GPT-3 генерувати текст, який відповідає заданому контексту та завданню. Наприклад, якщо введенням є запит на детальний опис особи, модуль керування генеруватиме керуючий сигнал, який повідомляє моделі створити більш детальний і конкретний опис. Якщо введенням є підказка із запитом про загальний огляд теми, модуль керування згенерує керуючий сигнал, який скаже моделі створити більш загальний і високорівневий опис. Це дозволяє GPT-3 генерувати текст, який підходить для заданого контексту та завдання.

GPT-3 також може включати зовнішні знання у свій вихід, що дозволяє генерувати текст, який є фактично точним та інформативним. Це досягається завдяки використанню трансформатора знань, який є окремим компонентом моделі, яка навчається на великому наборі даних із багатим на знання текстом.

GPT-3 був навчений на великому наборі текстових даних під назвою Internet Archive Books dataset. Цей набір даних складається з понад 8 мільйонів книг та інших текстів, які були оцифровані Інтернет-архівом, некомерційною організацією, яка працює над збереженням і наданням доступу до творів культури. Набір даних включає широкий спектр текстів, включаючи книги, статті та веб-сайти, і охоплює широкий діапазон мов і тем.

Окрім набору даних Internet Archive Books, GPT-3 також навчався на інших наборах даних, включаючи англійську Вікіпедію та набір даних WebText, який складається з тексту, взятого з Інтернету [4]. Комбінація цих наборів даних дозволяє GPT-3 навчатися на різноманітних і репрезентативних зразках тексту, що допомагає генерувати високоякісний і зв'язний результат.

Потенційне застосування та вплив

GPT-3 має потенціал для революції в обробці природної мови та штучному інтелекті, і він уже використовується для широкого спектру програм. Деякі з найбільш перспективних потенційних застосувань GPT-3 включають:

- Чат-боти: GPT-3 можна використовувати для створення чат-ботів, які можуть вести природні розмови з користувачами. Це може мати широкий спектр застосувань, наприклад, обслуговування клієнтів, освіта та розваги.
- Мовний переклад: GPT-3 може виконувати мовний переклад з високою точністю, що може мати значні наслідки для глобального спілкування та співпраці.
- Створення вмісту: GPT-3 має можливість генерувати текст, схожий на людину, який можна використовувати для автоматизації створення вмісту для веб-сайтів, соціальних мереж та інших платформ.
- Композиція музики: GPT-3 навіть використовувався для створення музики, демонструючи його здатність розуміти та генерувати складні моделі.

Одним із ключових застосувань GPT-3 є створення контенту [3]. Модель може генерувати високоякісні статті, публікації в блогах і со-

ціальних мережах, які можуть бути корисними для підприємств і організацій, які прагнуть швидко й ефективно створювати вміст. GPT-3 також можна використовувати для більш творчих додатків, таких як створення музики та написання історій [6].

Окрім можливостей створення мови, GPT-3 також має здатність виконувати такі завдання, як очищення та форматування даних, що може бути корисним для підприємств і організацій, які обробляють великі набори даних.

Незважаючи на свої вражаючі можливості, GPT-3 не позбавлений обмежень. Одним із потенційних обмежень GPT-3 є його залежність від даних, на яких він навчався [7]. Якщо навчальні дані містять упереджений або невідповідний вміст, модель може створити упереджений або невідповідний вихід. Важливо ретельно розглянути різноманітність і якість даних, які використовуються для навчання GPT-3, і переконатися, що вони є репрезентативними для населення або завдання, для якого вони використовуються.

Іншим потенційним обмеженням GPT-3 є його вартість і вимоги до ресурсів. Модель вимагає великої кількості обчислювальних ресурсів і даних для ефективного навчання, що може зробити її дорогим і ресурсомістким у використанні [8]. Це може обмежити його доступність для деяких організацій.

Висновки та пропозиції. Загалом GPT-3 — це потужна та універсальна мовна модель, яка здатна виконувати широкий спектр мовних завдань і генерувати високоякісний текст. Однак важливо ретельно розглянути потенційні обмеження та етичні міркування моделі, використовуючи її для конкретних цілей.

Є кілька ключових особливостей і характеристик GPT-3, які роблять його унікальним і потужним інструментом для досліджень НЛП і ШІ. Ось деякі з основних ключів до розуміння GPT-3:

- Великий масштаб: GPT-3 є однією з найбільших мовних моделей, коли-небудь розроблених, із 175 мільярдами параметрів. Це дозволяє генерувати високоякісний текст і виконувати складні завдання NLP з рівнем продуктивності, якого важко досягти з меншими моделями.
- Трансформаторна архітектура: GPT-3 використовує трансформаторну архітектуру, яка є типом нейронної мережі, яка широко використовується в завданнях НЛП. Архітектура трансформатора дозволяє GPT-3 обробляти введе-

ний текст більш ефективним і результативним способом, дозволяючи генерувати текст, який є більш зв'язним і відповідним контексту.

- Моделювання замаскованої мови (MLM): GPT-3 використовує техніку під назвою моделювання замаскованої мови (MLM), яка передбачає маскуванню частини вхідного тексту та прогнозування відсутніх слів на основі навколишнього контексту. Це допомагає GPT-3 дізнатися про зв'язки між словами та структурою мови.
- Динамічне керування: GPT-3 представляє нову техніку, яка називається динамічним керуванням, яка дозволяє моделі адаптувати свої результати на основі контексту чи поточного завдання. Це дозволяє GPT-3 генерувати текст, який є більш актуальним і підходить для різних ситуацій.

GPT-3 є потужним і універсальним інструментом, який має потенціал зробити революцію в дослідженнях НЛП та ШІ. Його великий масштаб і розширені можливості роблять його цінним ресурсом для дослідників і розробників, які працюють над широким спектром програм і проєктів.

© Голубенко О.І., Підмогильний О.О., 2022

ЛІТЕРАТУРА

1. Language models are better than humans at next-token prediction Authors: Buck Shlegeris, Fabien Roger, Lawrence Chan, Euan McLean. arXiv:2212.10560.

2. Self-Instruct: Aligning Language Model with Self Generated Instructions. Authors: Yizhong Wang, Yeganeh Kordi, Swaroop Mishra, Alisa Liu, Noah A. Smith, Daniel Khashabi, Hannaneh Hajishirzi. arXiv:2212.10509.

3. Interleaving Retrieval with Chain-of-Thought Reasoning for Knowledge-Intensive Multi-Step Questions Authors: Harsh Trivedi, Niranjana Balasubramanian, Tushar Khot, Ashish Sabharwal. arXiv:2212.08072.

4. Foresight -- Deep Generative Modelling of Patient Timelines using Electronic Health Records Authors: Zeljko Kraljevic, Dan Bean, Anthony Shek, Rebecca Bendayan, Joshua Au Yeung, Alexander Deng, Alfie Baston, Jack Ross, Esther Idowu, James T Teo, Richard J Dobson. arXiv:2212.04037.

5. Demystifying Prompts in Language Models via Perplexity Estimation Authors: Hila Gonen, Srini Iyer, Terra Blevins, Noah A. Smith, Luke Zettlemoyer. arXiv:2211.09267.

6. Reflect, Not Reflex: Inference-Based Common Ground Improves Dialogue Response Quality.

Authors: Pei Zhou, Hyundong Cho, Pegah Jandaghi, Dong-Ho Lee, Bill Yuchen Lin, Jay Pujara, Xiang Ren. arXiv:2211.07615.

7. UGIF: UI Grounded Instruction Following. Authors: Sagar Gubbi Venkatesh, Partha Talukdar, Srini Narayanan. arXiv:2210.17497.

8. Leveraging Pre-trained Models for Failure Analysis Triplets Generation. Authors: Kenneth Ezukwoke, Anis Hoayek, Mireille Batton-Hubert, Xavier Boucher, Pascal Gounet, Jerome Adrian. arXiv:2210.17238.

REFERENCES

1. Language models are better than humans at next-token prediction Authors: Buck Shlegeris, Fabien Roger, Lawrence Chan, Euan McLean. arXiv:2212.10560

2. Self-Instruct: Aligning Language Model with Self Generated Instructions. Authors: Yizhong Wang, Yeganeh Kordi, Swaroop Mishra, Alisa Liu, Noah A. Smith, Daniel Khashabi, Hannaneh Hajishirzi. arXiv:2212.10509

3. Interleaving Retrieval with Chain-of-Thought Reasoning for Knowledge-Intensive Multi-Step Questions Authors: Harsh Trivedi, Niranjana Balasubramanian, Tushar Khot, Ashish Sabharwal. arXiv:2212.08072

4. Foresight -- Deep Generative Modelling of Patient Timelines using Electronic Health Records Authors: Zeljko Kraljevic, Dan Bean, Anthony Shek, Rebecca Bendayan, Joshua Au Yeung, Alexander Deng, Alfie Baston, Jack Ross, Esther Idowu, James T Teo, Richard J Dobson. arXiv:2212.04037

5. Demystifying Prompts in Language Models via Perplexity Estimation Authors: Hila Gonen, Srini Iyer, Terra Blevins, Noah A. Smith, Luke Zettlemoyer. arXiv:2211.09267

6. Reflect, Not Reflex: Inference-Based Common Ground Improves Dialogue Response Quality Authors: Pei Zhou, Hyundong Cho, Pegah Jandaghi, Dong-Ho Lee, Bill Yuchen Lin, Jay Pujara, Xiang Ren. arXiv:2211.07615

7. UGIF: UI Grounded Instruction Following. Authors: Sagar Gubbi Venkatesh, Partha Talukdar, Srinu Narayanan. arXiv:2210.17497

8. Leveraging Pre-trained Models for Failure Analysis Triplets Generation. Authors: Kenneth Ezukwoke, Anis Hoayek, Mireille Batton-Hubert, Xavier Boucher, Pascal Gounet, Jerome Adrian. arXiv:2210.17238

СТАТТЯ НАДІЙШЛА ДО РЕДАКЦІЇ 23.11.2022