

Научная статья

УДК 004.032.26

ББК 32.818.1

А 64

DOI: 10.53598/2410-3225-2023-1-316-80-85

Анализ возможностей нейронной сети на основе языковой модели GPT-3 и способы ее применения на производстве
(Рецензирована)

**Альфира Менлигуловна Кумратова¹, Надежда Витальевна Морозова²,
Андрей Игоревич Василенко³, Ирина Евгеньевна Когай⁴**

^{1, 3, 4} Кубанский государственный аграрный университет, Краснодар, Россия

² Северо-Кавказская государственная академия, Черкесск, Россия, pi@ncsa.ru

¹ kumratova.a@edu.kubsau.ru

³ vasilenkoa0101@gmail.com

⁴ ikogai.me@gmail.com

Аннотация. Целью данной статьи является рассмотрение самого мощного алгоритма обработки естественного языка GPT-3, а также исследование возможностей применения данной нейросети для разработки продуктов в различных сферах производства. GPT-3 (Generative Pre-trained Transformer 3) – это нейронная сеть, способная генерировать тексты, которые могут быть похожи на тексты, написанные человеком. Это одна из самых мощных и продвинутых языковых моделей на сегодняшний день, которая использует машинное обучение для создания текстов на основе предыдущих данных. Она разработана компанией OpenAI и имеет множество вариантов использования в различных отраслях, включая производство.

Ключевые слова: нейронные сети, нейросети, текстовый запрос, GPT-3, обработка естественного языка, обработка текстового запроса, искусственный интеллект, нейролингвистическое программирование

Original Research Paper

Analysis of possibilities of the GPT-3 neural network model and ways of its application in production

**Alfira M. Kumratova¹, Nadezhda V. Morozova²,
Andrey I. Vasilenko³, Irina E. Kogay⁴**

^{1, 3, 4} Кубанский государственный аграрный университет, Krasnodar, Russia

² North Caucasus State Academy, Cherkessk, Russia, pi@ncsa.ru

¹ kumratova.a@edu.kubsau.ru

³ vasilenkoa0101@gmail.com

⁴ ikogai.me@gmail.com

Abstract. The purpose of this article is to consider diffusion neural networks, as well as to study the possibilities of using these neural networks to develop applications that generate images based on text queries and to compare diffusion models that exist today. GPT-3 (Generative Pretrained Transformer 3) is a neural network capable of generating texts that may be similar to human-written texts. This is one of the most powerful and advanced language models to date, which uses machine learning to create texts based on previous data. It was developed by Open AI and has many use cases in various industries, including manufacturing.

Keywords: neural networks, neuronets, text query, GPT-3, natural language processing, text query processing, artificial intelligence, neuro-linguistic programming

Вопросам применения нейронных сетей в различных сферах человеческой жизнедеятельности посвящены труды отечественных ученых: В.А. Бейненсон, Н.Л. Зыховская, А.В. Демин, И.В. Рыбальченко, С.В. Анахов, Т.О. Шаврина, М.В. Флоринская, А.И. Вардинов, Д.Ю. Какутин, А.С. Дмитриев, И.А. Абрамов и других ученых.

GPT-3 (Generative Pre-trained Transformer 3) является алгоритмом обработки естественного языка, разработанным компанией OpenAI и самой крупной языковой моделью в мире. GPT-3 обучен на смеси текстовых данных и кода, опубликованных до конца 2021 г. Данная модель обучена с помощью генеративного предварительного обучения и предсказывает значение будущих токенов на основе предыдущих. В качестве токена авторы подразумевают в данном контексте любой символ, букву или цифры во взаимодействии с нейронной сетью (нейросетью).

Нейросеть умеет находить логические связи между предложениями, словами и частями слов в огромных массивах данных из интернета, включая сотни тысяч страниц Wikipedia, постов из социальных сетей и новостных статей.

Для обучения GPT-3 использованы 175 млрд. параметров, требующих 800 гигабайт для хранения, а также нейросеть держит контекст длиной 2048 токенов.

Далее рассмотрено подробно, в каком процентном соотношении какие наборы данных использованы при обучении (табл. 1).

Таблица 1

Процентное соотношение наборов данных, задействованных в обучении GPT-3

Table 1. Percentage of datasets involved in GPT-3 training

Набор данных	Количество токенов в миллиардах	Процентное соотношение данных в пределах обучения
Common Crawl	410	60%
Web Text2	19	22%
Books1-	12	8%
Books2	55	8%
Wikipedia	3	3%

В связи с тем, что для обучения GPT-3 использовались данные, полученные из интернета, в том числе с различных форумов и социальных сетей, которые содержат токсичные конструкции с нецензурной лексикой, GPT-3 иногда генерирует тексты, содержащие нецензурную лексику.

На сегодняшний день имеется много исследований, которые посвящены использованию GPT-3 в таких целях, как:

- обработка большого объема текстовых данных с целью получения тезисной информации;
- автоматизация написания текстов;
- анализ текста для выявления настроения автора и т.д.

До 2023 года официальный доступ к API имели только несколько компаний, в том числе: Reddit, Quizlet, AIDungeon и др. Второго февраля 2023 года OpenAI предоставили доступ к API GPT-3.5 Turbo всем желающим разработчикам, в связи с чем все больше и больше компаний будут использовать данную нейросеть на своих производствах.

ChatGPT (продукт, построенный вокруг GPT-3) стал одним из самых быстрорастущих ИТ-продуктов в истории, набрав 1 млн. пользователей за 5 дней после выпуска пред релизной версии и 100 млн. пользователей в январе 2023 г.

GPT-3 можно успешно применять для автоматизации ручных процессов и рутинных задач, в связи с чем нейросеть уже сейчас может снять часть нагрузки с производственных работников. Ниже рассмотрены возможные способы применения данного алгоритма на производстве.

1. Создание контента. GPT-3 отлично справляется с такими задачами, как написание блог постов, сценариев, художественных произведений и даже написание самого контент-плана (рис. 1).



Привет! Напиши, пожалуйста, контент план для блога о психотерапии



Конечно, я могу помочь с контент-планом для блога о психотерапии! Вот несколько идей для статей:

1. Что такое психотерапия и как она работает?
2. Типы психотерапии: краткое описание каждого вида и его применение.
3. Как выбрать психотерапевта и что искать в нем?
4. Основные причины, по которым люди обращаются к психотерапевту.
5. Терапевтические подходы для лечения депрессии и тревожности.
6. Как психотерапия может помочь при проблемах в отношениях?
7. Психотерапия для детей: как это работает и когда ее нужно применять?
8. Психотерапия для пар: что это и как это работает?
9. Медитация и майндфулнесс как дополнительные методы поддержки в психотерапии.
10. Как продолжать работу над собой после завершения курса психотерапии.

Надеюсь, эти идеи помогут вам начать создание контент-плана для вашего блога о психотерапии. Желаю вам успехов в написании!

Рис. 1. Ответ ChatGPT, работающего на основе GPT-3.5, на просьбу о создании контента

Fig. 1. Response of ChatGPT, working on the basis of GPT-3.5 to content creation request

2. Составление документов. Например, данную нейросеть можно использовать для составления исковых заявлений в юридических компаниях или сопроводительных писем для отправки резюме в компаниях, помогающих людям с трудоустройством. Для того чтобы получить необходимый документ, в запросе нужно указать как можно больше подробностей. На примере, представленном ниже, авторами задействована нейросеть для составления искового заявления об истребовании долговых обязательств. На что алгоритм не просто написал заявление, но также сообщил, какие обязательные данные должны присутствовать в документе, а в конце посоветовал обратиться к профессиональным юристам (рис. 2).

3. Составление плана обучения. Нейросеть не только может составить план для развития специалистов, но и поможет разобраться со сложностями, возникающими при обучении.

4. Чаты поддержки. Очень часто у больших компаний есть целые отделы поддержки продукта, на поддержание которых уходят большие бюджеты. На рисунке 3 показано, насколько хорошо нейросеть справляется с такого рода задачами.

5. Помощь в написании кода программистами. Ввиду того, что нейросеть обучена на большом объеме данных, в материалах для обучения в том числе использованы и примеры написания кода из публичных проектов. Стоит отметить, что на данный момент алгоритм не способен полноценно заменить программистов, но может помогать им решать рутинные задачи.

Например, продукт Copilot (<https://github.com/features/copilot>) может встраивать в среду разработки и помогать программистам в написании продуктов.



Меня зовут Ирина, гражданин Иванов Е.Е. взял у меня в долг 50 000 рублей и не возвращает. Помогите, пожалуйста, написать исковое заявление в суд



Добрый день, Ирина. Я готов помочь вам с написанием искового заявления.



Ниже я предоставляю вам пример, который вы можете использовать для своего дела.

Обратите внимание, что конкретные детали вашего дела, такие как точная дата, место подписания договора и т.д., должны быть указаны в документе. Также убедитесь, что вы включили все необходимые доказательства, подтверждающие ваше требование.

Исковое заявление:

"К Иванову Евгению Евгеньевичу,
адрес: _____

Исковое заявление

Я, _____ (ФИО), действующая от своего имени, зарегистрированная по адресу _____, заявляю иск к Иванову Евгению Евгеньевичу, зарегистрированному по адресу _____, о возврате денежных средств в размере 50 000 рублей, взятых в долг в период с _____ по _____ годы.

Regenerate response

Рис. 2. Исковое заявление, написанное ChatGPT, работающего на основе GPT-3.5

Fig. 2. Complaint written by ChatGPT, working on the basis of GPT-3.5



Представь, что ты работник технической поддержки мебельного магазина и должен ответить на вопрос покупателя:

У меня сломался диван, я купила его 2 недели назад в вашем магазине, что мне делать?



Здравствуйте! Я понимаю, что это может быть очень неудобно для вас. Наш магазин готов помочь вам решить эту проблему.



В соответствии с нашей политикой гарантии, вы можете обратиться к нам в течение двух недель после покупки, если возникли какие-либо проблемы с товаром. В данном случае, поскольку вы приобрели диван две недели назад, вы можете обратиться к нам для получения замены или ремонта.

Чтобы начать процесс обращения по гарантии, вам нужно связаться с нашим отделом обслуживания клиентов и предоставить информацию о поломке дивана. Мы также можем попросить вас предоставить фотографии поломки, чтобы помочь нам лучше понять проблему и предоставить наилучший вариант решения.

Спасибо, что выбрали наш магазин. Мы готовы помочь вам и решить эту проблему как можно быстрее.

Рис. 3. Ответ, сгенерированный нейросетью GPT-3.5 в роли сотрудника технической поддержки

Fig. 3. Response generated by the GPT-3.5 neural network as a technical support officer

GPT-3 может быть использована в производстве для автоматизации рутинных процессов. Она может быть использована для создания инструкций и руководств, например, для обучения новых сотрудников и улучшения процессов производства. Также она может помочь в создании более эффективных систем управления запасами и планирования производства.

Нейросеть GPT-3 также может быть использована в исследованиях для создания новых идей и концепций. Она применяется в получении более точных и надежных прогнозов, основанных на предыдущих данных. Также она может быть использована для создания новых моделей и алгоритмов, которые могут помочь улучшить процессы производства. В маркетинге GPT-3 может быть использована для создания более эффективных, более привлекательных объявлений для потребителей. Также ее используют в получении более точных прогнозов спроса на продукты.

GPT-3 – это мощный инструмент, который может помочь улучшить производственные процессы и повысить производительность. Его использование в различных отраслях, включая производство, исследования и маркетинг, становится все более популярным, и это свидетельствует о его эффективности и потенциале в будущем.

В данной статье рассмотрено несколько способов использования нейросети GPT-3 на производстве. Нет сомнений, что открытый доступ к API алгоритма является значимым моментом для всех компаний, где присутствует много ручных процессов, и может помочь сократить бюджеты на найм людей, выполняющих рутинные процессы. Также данная нейросеть может помочь небольшим производствам развиваться быстрее без дополнительных затрат на наем таких специалистов, как маркетолог, технический писатель или специалист технической поддержки в чате.

Примечания

1. Бейнсен В.А. Применение генеративных нейросетей в журналистике: проблемы и перспективы // Динамика медиасистем. 2023. Т. 3, № 1. С. 352–359.
2. Зыховская Н.Л. Цифровизация творчества: возможности Искусственного Интеллекта в литературе, журналистике и Искусстве // Медиасреда. 2022. № 2. С. 107–110.
3. Демин А.В., Рыбальченко И.В. Управленческие приемы использования нейронных сетей как новейшей парадигмы действий современного руководителя // Муниципальная академия. 2022. № 4. С. 21–30.
4. Анахов С.В. Достижения и тренды цифровой трансформации в научно-образовательной и технической сферах // Новые информационные технологии в образовании и науке. 2022. № 5. С. 5–16.
5. Шаврина Т.О. О методах компьютерной лингвистики в оценке систем искусственного интеллекта // Вопросы языкознания. 2021. № 6. С. 117–138.
6. Флоринская М.В., Вардинов А.И. Искусственный интеллект GPT-3 // Вестник Ессентукского института управления, бизнеса и права. 2021. № 18. С. 150–156.
7. Какутин Д.Ю., Дмитриев А.С., Абрамов И.А. Обзор вспомогательных инструментов на основе машинного обучения для написания исходного кода программ // Инженерный вестник Дона. 2022. № 5 (89). С. 27–33.
8. Wikipedia. GPT-3. URL: <https://en.wikipedia.org/wiki/GPT-3> (дата обращения: 10.03.2023).
9. Ben Lutkevich, Ronald Schmelzer. GPT-3. URL: <https://lilianweng.github.io/posts/2021-07-11-diffusion-models/> (дата обращения: 10.12.2022).
10. Михалевич Ю. О возможности использования больших языковых моделей в разработке ПО. URL: <https://www.itweek.ru/ai/article/detail.php?ID=225994> (дата обращения: 10.03.2023).

References

1. Beinenson V A. Application of generative neural networks in journalism: problems and perspectives // Dynamics of Media Systems. 2023. Vol. 3, No. 1. P. 352–359.

2. Zyhovskaya N.L. Digitalization of creativity: the possibilities of Artificial Intelligence in literature, journalism and art // Media Environment. 2022. No. 2. P. 107–110.
3. Demin A.V., Rybalchenko I.V. Managerial techniques of using neural networks as the newest paradigm of actions of a modern leader // Municipal Academy. 2022. No. 4. P. 21–30.
4. Anakhov S.V. Achievements and trends of digital transformation in scientific, educational and technological spheres // New Information Technologies in Education and Science. 2022. No. 5. P. 5–16.
5. Shavrina T.O. On methods of computer linguistics in the evaluation of artificial intelligence systems // Questions of Linguistics. 2021. No. 6. P. 117–138.
6. Florinskaya M.V., Vardikov A.I. Artificial Intelligence GPT-3 // Bulletin of the Essentuki Institute of Management, Business and Law. 2021. No. 18. P. 150–156.
7. Kakutin D.Yu., Dmitriev A.S., Abramov I.A. Overview of auxiliary tools based on machine learning for writing the source code of programs // Engineering Bulletin of the Don. 2022. No. 5 (89). P. 27–33.
8. Wikipedia. GPT-3. URL: <https://en.wikipedia.org/wiki/GPT-3> (access date: 10/03/2023).
9. Ben Lutkevich, Ronald Schmelzer. GPT-3. URL: <https://lilianweng.github.io/posts/2021-07-11-diffusion-models/> (access date: 10/12/2022).
10. Mikhalevich Yu. About the possibility of using large language models in software development. URL: <https://www.itweek.ru/ai/article/detail.php?ID=225994> (access date: 10/03/2023).

Авторы заявляют об отсутствии конфликта интересов.

The authors declare no conflicts of interests.

Статья поступила в редакцию 21.02.2023; одобрена после рецензирования 10.03.2023; принята к публикации 11.03.2023.

The article was submitted 21.02.2023; approved after reviewing 10.03.2023; accepted for publication 11.03.2023.

© А.М. Кумратова, Н.В. Морозова, А.И. Василенко, И.Е. Когай, 2023