

ChatGPT: limitations, challenges and potential applications

Matheus Vinicius Abadia Ventura^{1,2} & Antonio Carlos Pereira de Menezes Filho²

¹ UniBRAS University Center of Southwest Goiano, UniBRAS, Rio Verde, Goiás state, Brazil

² Goiano Federal Institute, IF Goiano, Rio Verde, Goiás state, Brazil

Correspondence: Matheus Vinicius Abadia Ventura, UniBRAS University Center of Southwest Goiano, UniBRAS, Rio Verde, Goiás, Brazil. E-mail: matheusvinicius10@hotmail.com

Received: June 03, 2023

DOI: 10.14295/bjs.v3i1.427

Accepted: July 14, 2023

URL: <https://doi.org/10.14295/bjs.v3i1.427>

Abstract

This article introduces ChatGPT, a language model based on text generation developed by OpenAI. Built on top of the GPT-3.5 architecture, ChatGPT is capable of engaging in interactive conversations with users, answering questions, providing information, and even simulating natural conversations. The model is trained on large amounts of text data and uses an attention mechanism to learn language structure and generate coherent and relevant responses. The purpose of this article is to analyze and discuss the training process, main characteristics, challenges and applications of ChatGPT. ChatGPT represents an important step towards more sophisticated and interactive dialog systems, driving the advancement of artificial intelligence in the field of conversation.

Keywords: Artificial Intelligence, OpenAI, AI.

ChatGPT: limitações, desafios e aplicações potenciais

Resumo

Este artigo apresenta o ChatGPT, um modelo de linguagem baseado na geração de texto desenvolvido pela OpenAI. Construído sobre a arquitetura GPT-3.5, o ChatGPT é capaz de se envolver em conversas interativas com os usuários, respondendo a perguntas, fornecendo informações e até simulando conversas naturais. O modelo é treinado em grandes quantidades de dados de texto e usa um mecanismo de atenção para aprender a estrutura da linguagem e gerar respostas coerentes e relevantes. O objetivo deste artigo é analisar e discutir o processo de treinamento, principais características, desafios e aplicações do ChatGPT. O ChatGPT representa um passo importante para sistemas de diálogo mais sofisticados e interativos, impulsionando o avanço da inteligência artificial no campo da conversação.

Palavras-chave: Inteligência Artificial, OpenAI, IA.

1. Short Communication

1.1 Introduction

In recent years, driven by the development of machine learning-based language models such as GPT (Generative Pretrained Transformer), the field of language generation has made great progress. GPT, a language model based on text generation developed by OpenAI (OpenAI, 2023), has proven to be very effective in tasks such as text generation, machine translation and document summarization (Figure 1). However, applying these models to interactive dialogues remains a challenge due to the need to generate contextually coherent and relevant responses (Deng; Lin, 2022).

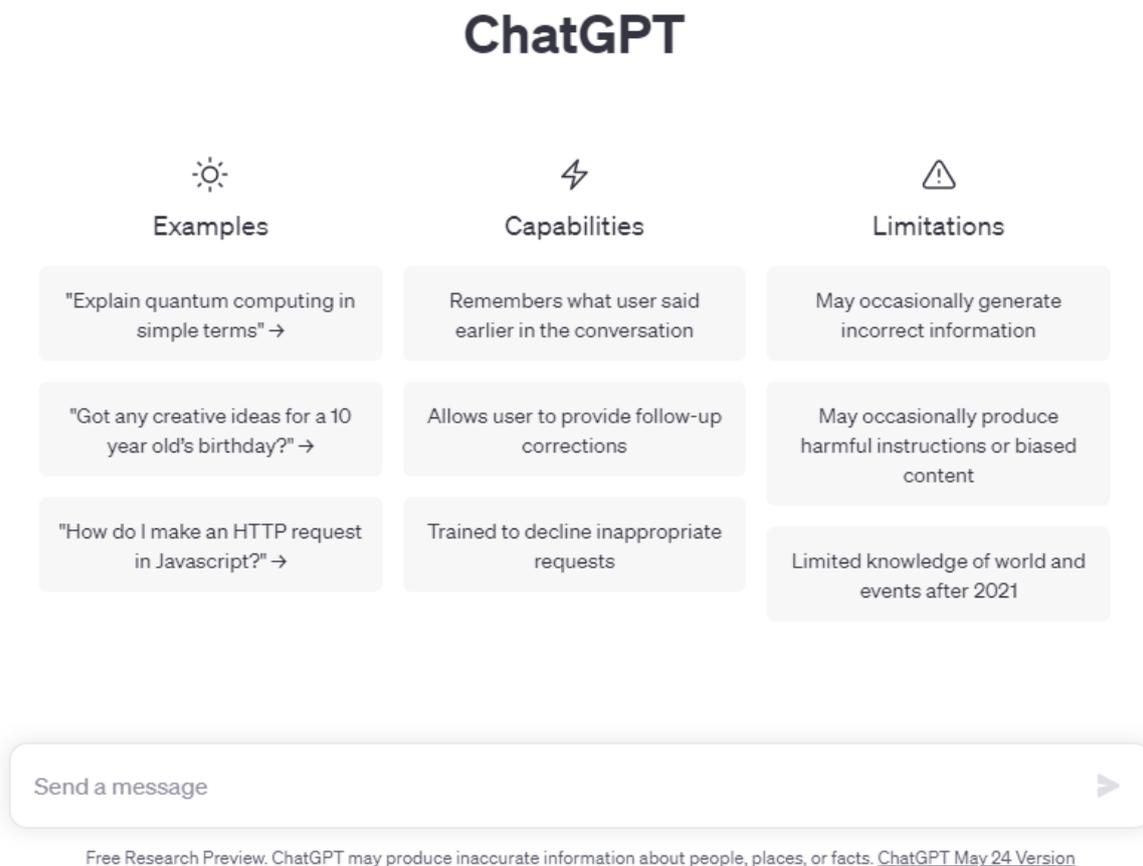


Figure 1. Screen view of the user logged into ChatGPT. Source: OpenAI, 2023.

The ability to have a natural conversation with a machine has been a long-standing goal in the field of artificial intelligence. ChatGPT is a recent and promising solution for this task. It is trained on large amounts of textual data to understand and generate coherent responses, making interaction with the model more enjoyable and realistic.

When interacting with ChatGPT, you can ask questions, ask for advice, request explanations or just chat. However, it's important to remember that ChatGPT is a machine with no real-world knowledge or experience beyond what it learned from the training data. Therefore, your responses are based on patterns found in the data and may not always be accurate or up to date.

ChatGPT is trained on a "text completion" task, where it takes an incomplete piece of text and generates a plausible continuation. It can understand and generate text in a variety of styles and domains, from casual conversation to technical writing, factual information, descriptions, answers to questions and much more.

Regarding applications, ChatGPT was used in several areas such as health and medicine, business and finance, law and legal services, creative writing and content generation, education and training, programming and code debugging, media and entertainment, sales and marketing, banking, scientific research and data processing and analysis (Ray, 2023). In this article, we explore the main features, limitations, challenges and potential applications of ChatGPT.

1.2. ChatGPT Methodology

ChatGPT is trained on a dataset of human conversations, which includes questions and answers from various sources such as forums, chat, and customer service. The model adopts the GPT-3.5 architecture and consists of a multi-layer Transformer-type encoder and decoder. These layers allow the model to learn the structure of the language and capture the contextual dependencies needed to generate coherent responses.

ChatGPT training involves maximizing the conditional probability of each response given a sequence of previous conversations. The model is fed a sequence of input tokens and trained to predict the next correct token. This process is repeated many times using techniques such as random sampling and reinforcement training to improve the quality of the responses generated.

ChatGPT has several features that make it an effective template for interactive conversations. He is able to deal with questions from various fields and provide relevant and useful answers. The template is also capable of asking for clarification when a user's question is unclear or lacks sufficient information. Additionally, ChatGPT can incorporate contextual information throughout the conversation to generate more accurate and personalized responses.

ChatGPT training is a complex process involving the use of supervised and reinforcement learning techniques. Initially, the model is pre-trained on a large and diverse text corpus, which allows the model to learn general language representations. It is then tuned using specific data from the dialog task. During training, an "autoregressive" technique is applied, in which the model predicts the next word in a sequence based on the previous context.

1.3. Limitations, Challenges and Potential Applications

Despite representing a significant advance in the generation of conversational language, ChatGPT still has some limitations. Although the model is capable of generating apparently correct answers, it is possible that these answers are factually inaccurate. Furthermore, ChatGPT can be sensitive to malicious input, resulting in undesirable or inappropriate responses in certain situations (Deng & Lin, 2022).

Dealing with these limitations is an ongoing challenge with the aim of improving the accuracy and ethics of the model. Bias mitigation and ethical responsibility are also important concerns, as the model can reproduce or amplify biases present in the training data. While ChatGPT has achieved impressive results, it is essential to face and address the challenges and limitations involved. According to Zhou et al. (2023), ethical concerns: bias, privacy and security, transparency, abuse and authorship.

ChatGPT has the ability to generate answers that seem plausible, but occasionally it may give incorrect information or vague answers. Also, the model may be sensitive to slightly differently rephrased questions. Improving consistency, handling multiple shifts, and understanding more complex commands are all areas of active research to improve ChatGPT performance.

AI-based chat technology like ChatGPT requires a careful approach to ethics and responsibility (Khlaif, 2023). It is important to mitigate bias, ensure model transparency and implement control mechanisms to prevent abuse. In addition, future research should focus on improving contextual understanding, improving response generation, and exploring more advanced training techniques.

ChatGPT presents a wide range of potential applications (Raj, 2023). It can be used as a virtual assistant on websites, chatbots, customer service systems and educational support. The model has the potential to improve the user experience, providing fast and accurate responses, in addition to enabling real-time interaction in various domains.

2. Final considerations

ChatGPT represents a significant advance in the ability to generate coherent and relevant responses in interactive conversations. Its architecture based on GPT-3.5 and training with large datasets allow the model to produce more natural and useful dialogs. While there are still challenges to be overcome, ChatGPT demonstrates enormous potential in a variety of applications, from customer service to education and virtual assistants. With continued improvements, ChatGPT is expected to contribute to enhancing human-machine interaction and driving innovation in artificial intelligence.

Despite the promising potential of ChatGPT to enhance various applications and improve the user experience in conversational interactions, there are still challenges to be faced. Issues related to response accuracy, mitigation of bias, and ethical accountability must be addressed to ensure responsible and ethical use of AI-based

conversational systems. ChatGPT represents an important step towards more sophisticated and interactive dialog systems, driving the advancement of artificial intelligence in the field of conversation.

3. Authors' Contributions

Matheus Vinicius Abadia Ventura: writing, grammatical and scientific corrections in the manuscript and publication. *Antonio Carlos Pereira de Menezes Filho*: writing, grammatical and scientific corrections in the manuscript and publication.

4. Conflicts of Interest

No conflicts of interest.

5. Ethics Approval

Not applicable.

6. References

- Deng, J., & Lin, Y. (2022). The benefits and challenges of ChatGPT: An overview. *Frontiers in Computing and Intelligent Systems*, 2(2), 81-83. <https://doi.org/10.54097/fcis.v2i2.4465>
- Khlaif, Z. N. (2023). Ethical Concerns about Using AI-Generated Text in Scientific Research. Available at SSRN 4387984. <https://doi.org/10.2139/ssrn.4387984>
- OpenAI (2023). ChatGPT. Available: <https://chat.openai.com/>
- Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, 3, 121-154. <https://doi.org/10.1016/j.iotcps.2023.04.003>
- Zhou, J., Müller, H., Holzinger, A., & Chen, F. (2023). Ethical ChatGPT: Concerns, challenges, and commandments. *arXiv preprint arXiv:2305.10646*. <https://doi.org/10.48550/arXiv.2305.10646>

Funding

Not applicable.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).