

Characteristics, prospects and limitations of ChatGPT

Jinyue Li

School of Artificial Intelligence, Anhui University of Science and Technology,
Huainan, Anhui, China, 232000

1804628638@qq.com

Abstract. In November 2022, Chat Generative Pre-trained Transformer (ChatGPT) released by OpenAI attracted worldwide attention as soon as it was released. Both artificial intelligence practitioners and the public are scrambling to try this latest chatbot. This paper mainly studies what ChatGPT is, and the remarkable characteristics of ChatGPT compared with other chatbots, as well as the prospect and limitations of ChatGPT. The main purpose of writing this paper is to provide some references for students or practitioners in the direction of artificial intelligence, and to make people who do not know ChatGPT and other large-scale language chat models more systematic understanding. The latest large-scale language model ChatGPT is studied by literature analysis and summary. This paper finds that although ChatGPT has indeed brought shock to the global people in the field of machine learning in the field of artificial intelligence, the hidden dangers and defects behind it (academic integrity, privacy, etc.) are gradually exposed, and relevant practitioners still need to propose measures to improve and develop new uses of chatbots.

Keywords: ChatGPT, artificial intelligence, machine learning.

1. Introduction

Artificial intelligence refers to the concept proposed after 1950 based on the Turing test. When humans cannot judge whether the other party is a machine through dialogue, they pass the Turing test. After 2010, artificial neural networks began to flourish. In 2018, OpenAI proposed a new language learning model-generative pre-trained transformer. From 2018 to 2022, three generations of Generative Pre-trained Transformer (GPT) were born successively. In November 2022, Chat Generative Pre-trained Transformer (ChatGPT) was officially born. Since the advent of ChatGPT, there have been numerous studies on it, currently focusing on the working mechanism of ChatGPT, the impact of ChatGPT on human society, including but not limited to affecting employment rates. However, most of the research on ChatGPT in academic integrity and academic fairness and the direction of future development. There is no detailed investigation and discussion. Knowing ChatGPT is of great significance for people to understand the latest development of artificial intelligence. For most non-artificial intelligence practitioners, ChatGPT's overnight explosion allows them to understand artificial intelligence more concretely, and no longer stick to abstract academic terms. At the same time, ChatGPT has caused countless people's thinking since its inception. In the past, there has never been such an artificial intelligence product that has such a huge public opinion and influence on human society. For such an influential product, it has great research. In this paper, the birth of ChatGPT, the limitations of ChatGPT and the application prospect of ChatGPT are studied. The latest large-scale language model ChatGPT is

studied by literature analysis and summary and review and comparative analysis. This study for artificial intelligence practitioners, can better inventory ChatGPT's various characteristics of the prospects, potential, but also sounded the alarm, ChatGPT in many aspects of the definition and use of restrictions are still vague, uncertain, but at the same time it is full of potential, promote professionals in this area of research. For non-practitioners, this article is a tip to remind them that ChatGPT still has many shortcomings, and for users, privacy and copyright issues cannot be ignored.

2. AI and ChatGPT

The House of Lords in the UK is persuaded that by building on a historically successful research program and recommending five principles, they are: AI should be created for humanity's benefit and for the greater good. AI should work in accordance with fairness and understandability standards. AI must not be applied to undermine the privacy or data rights of people, families, or communities. Everyone has the right to receive an education that will help them develop emotionally, cognitively, and financially alongside AI. AI should never be given the autonomy to harm, destroy, or deceive people [1].

ChatGPT, a generative pre-trained transformer, is now attaching not only AI practitioners but also a human. ChatGPT aims to create a cooperative AI system that can produce language that is helpful, engaging, and contextually relevant [2]. ChatGPT blends supervised fine-tuning with unsupervised pre-training to produce inquiry responses that sound like human speech and topical responses that mimic responses from human experts. The GPT-3 model is the newest update to a language model with 175 billion parameters that was created using a large dataset of naturally occurring text that was compiled from numerous internet sources, including web pages, books, academic publications, and social chats. In contrast to current language models, which often use deep learning with a focus on supervised learning, future evolutionary models may be more based on reinforcement learning [3].

According to the former study, it demonstrated that Generative AI systems are made to use the data they are trained on to make content or output like text, images, audio, simulations, videos, and codes. According to the former study, it showed that the most well-known use of AI is deep learning. To make predictions based on the data, algorithms in this technology are trained on large datasets. Translation of languages, speech recognition, and image recognition might all be included. Natural language processing (NLP) refers to AI that comprehends and uses human language. Examples of this include sentiment analysis, text summarization, and translation. The diverse and innovative AI models used by ChatGPT are built using unsupervised and semi-supervised ML techniques. A few examples include voice recordings, emails, extended text forms, social media posts, images, structured data, and extended text forms. Additional services they might offer include new content, translations, questions and answers, sentiment analysis, summaries, and movies. The potential of ChatGPT to improve people's life in a number of areas, Despite the fact that neither ChatGPT nor Open AI was the first Gen-AI system, it represents a technological leap forward in Generative AI [4].

A large language model (LLM) called ChatGPT employs deep learning to produce writings that resemble human speech in response to commands. It was unveiled on November 30, 2022, as the most recent version of Open AI's massive language models, which enable 'intelligent' discussions. The California-based company's Generative Pre-trained Transformer (GPT) models include this [4].

ChatGPT, however, differs significantly from the earlier models in a number of areas. The main distinction between it and GPT-3 is that the latter is intended to carry out a variety of natural language processing (NLP) tasks, including text summarization, question answering, and the creation of creative writing (such as poetry or fiction) as well as high-quality long or short form copy (such as blog posts) [4]. ChatGPT can respond to follow-up questions, reject improper solicitations, challenge wrong premises, and concede its errors. According to the former study, it demonstrated that contains a lot of emergent skills, like high-quality conversation, complex reasoning, chain-of-thought (CoT), few-shot learning (in-context learning), generalizing across tasks, and understanding code [5].

3. Characteristics and advantages of ChatGPT

ChatGPT re-developed the search engine field. As a key partner of OpenAI, Microsoft has integrated ChatGPT into its search engine product Bing for the first time. The new Bing can be used as a dialog system to respond to user queries, retrieve context through conversations with users, refine searches based on dialog context, narrow searches to improve search efficiency, and add references in responses, including retrieved web pages. In this approach, ChatGPT fulfills the function of information extraction and lessens the strain of visiting pointless web pages, resulting in a more natural connection between users and search engines. We believe that the annoyance of this conversation search will gradually replace the inefficient mode of keyword search [5]. ChatGPT also has the following advantages:

Generalization: ChatGPT creates replies with many turns that correspond to the user's purpose. The user experience in conversational interaction mode is substantially improved by ChatGPT's ability to record past conversations in order to provide context for some hypothetical inquiries. ChatGPT is improved using RLHF and instruction tweaking so that it may learn to generalize tasks and respond to input from people [5].

Correction: According to the former study, it showed that ChatGPT is willing to own up to its errors. If users point out its errors, the model adjusts the response in light of their suggestions. Additionally, ChatGPT has the ability to debate wrong questions and then offer a logical guess [5].

Creativity: ChatGPT excels in creative writing tasks such as copywriting, planning, and even refining the answers to users' requests. It has been utilized to create chatbots and other conversational AI systems that may be employed in applications for support and customer care. Generative AI may replace several professions that can produce different and unique text, audio and visual materials based on human needs, the most popular concept being its use as a tool like an "assistant" to make certain professions more accessible and acceptable to everyone [5]. The majority of participants' responses indicate that ChatGPT is effective in boosting students' prospects of academic success by giving users (teachers and students) a foundational understanding of a range of subjects. The participants also acknowledged ChatGPT's effectiveness in conveying a thorough comprehension of various issues in a clear manner. In light of this, it is possible to argue that ChatGPT will bring about a paradigm shift in current methods of instruction delivery and propel learning reform in a future full of digital possibilities [6].

Dialogue continuity: during the dialogue process, the dialogue information of the previous user will be remembered to realize continuous dialogue and improve the user experience in the dialogue interaction mode.

4. Limitations of ChatGPT

The majority of users are positive about the integration of ChatGPT and other AI-powered chatbots into educational systems. ChatGPT cannot effectively handle precise logical problems. ChatGPT often provides incorrect answers to math or logic problems that have explicit rather than probabilistic solutions. ChatGPT doesn't know if it's answering questions correctly or not, and it's more likely to answer like a human being than to answer correctly.

GPT is now unable to update itself in order to learn fresh facts and the most recent knowledge. Furthermore, it is challenging to encompass the information in the model. The model stores knowledge learned from large-scale databases in a distributed representation, where the process is a black box that is difficult to manipulate or interpret. artificial intelligence requires innovativeness, compassion, and other interestingly human capacities. Because they are only intended to carry out tasks that have been given to them, artificial intelligence systems are unable to creatively conceive of or comprehend the subtle human emotions. ChatGPT makes use of only plain text without any references or links. Also, Google is fostering its own huge language model and is utilizing artificial intelligence vigorously in its web search tool as ChatGPT develops. ChatGPT is a huge language model that is constantly trained to make answers more accurate [2].

While ChatGPT is powerful at producing safe and harmless responses, the system can still be vulnerable to attacks, or someone with ulterior motives can exploit ChatGPT for attacks, including directive attacks (making models follow the wrong instructions to do illegal or unethical things) and hint

injection [4]. Also, while ChatGPT does a good job in English and respecting United States culture, for other languages and cultures, it still requires a lot of data and training to optimize. According to Open AI, ChatGPT may sometimes react to instructions that are hurtful or discriminatory in their responses, occasionally making responses that sound reasonable but are incorrect or meaningless. These technologies have reason to be exciting, not least because they help reduce the factors that reduce quality of life, such as dyslexia. On the other hand, some technologies can also reduce these risks. The choice of initial data used to train these algorithms is critical to preventing the addition of hazardous materials [5].

ChatGPT has data security risks, and personal privacy (such as age, gender, address, and hobbies, and even personal privacy such as fund accounts) may be exposed during interaction with ChatGPT. Users' unintentional communication can reveal a significant amount of this personal information, and we believe that relevant laws and regulations should be issued to limit data collection and data processing methods to legal limits.

For experts, it's important for them to have faith in their surroundings and that the system's input won't be changed in the future. Additionally, usability is a key component of successfully integrating specialists into AI systems, necessitating the comprehension of the system's fundamentals by interface designers. It's vital to emphasize that usability and security are sometimes seen as fundamentally opposed concepts, which highlights the pressing need for research in the field of "usability security."

ChatGPT cannot be listed as an author, and ChatGPT's answers cannot be copyrighted. Many research papers list ChatGPT as an author. However, neither Nature nor Science have made it clear that LLMs are not eligible to be published as authors in their respective journals. Numerous concerns have arisen as a result of the absence of current user guidance on how to verify ChatGPT usage. The majority of the participants' responses revealed a statement that ChatGPT is effective in giving users (e.g. teachers and students) the opportunity to successfully understand the responses by providing them with basic knowledge of various aspects. In addition, participants in this experiment believe that ChatGPT is more efficient in its comprehensive understanding of various complex topics after users provide easy-to-understand language. From this point of view, we can say that ChatGPT will lead to an increase in the possibility of a paradigm shift in traditional teaching methods, and in the future promote the reform of learning methods with artificial intelligence machine learning [4].

Different users often give different answers when using ChatGPT, and even the answer form is different, which may be a text narrative or a table form. We believe this leads to unfair and uneven efficiency in learning, for example, in horizontal comparison problems, tables are often better for questioners to compare and understand than text statements [3].

The operation of ChatGPT requires a lot of computing power support. A few years ago, the number of parameters of gpt3 launched by openAI reached 175 billion, which is about 117 times that of cpt2. ChatGPT observes human feedback at every step, hoping to better understand human intentions [4].

ChatGPT's smooth conversations impress most users. Interactions with ChatGPT are often considered exciting and fun. However, it is believed that it is still a long way from being fully human because it is currently constrained by the text interface and is unable to comprehend the meaning of the utterance or the user's language cues.

5. Conclusion

This paper discusses the characteristics, advantages, limitations of ChatGPT, and the development direction and prospect of ChatGPT. ChatGPT can help people complete some simple tasks, but there are also problems such as academic integrity and privacy leakage. People need to formulate rules to constrain artificial intelligence and maximize the use of artificial intelligence within the feasible scope. The data accuracy and data sources of this paper are not extensive enough, future research on ChatGPT may focus on how to avoid the use of ChatGPT defects, how to maximize the use of ChatGPT and how to define the copyright of ChatGPT output content and how to prevent cheating with ChatGPT and search engines combined with ChatGPT application research.

References

- [1] Andreas Holzinger^{1,2}. (2018) Current Advances, Trends and Challenges of Machine Learning and Knowledge Extraction: From Machine Learning to Explainable AI. c IFIP International Federation for Information Processing 2018 Published by Springer Nature Switzerland AG. pp. 1–8, 2018. https://doi.org/10.1007/978-3-319-99740-7_1.
- [2] Abid Haleem a, Mohd Javaid , Ravi Pratap Singh. (2022) An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges. BenchCouncil Transactions on Benchmarks, Standards and Evaluations.
- [3] Yogesh K. Dwivedi a,b, (2023) “So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. International Journal of Information Management .
- [4] Damian Okaibedi Eke.(2023) ChatGPT and the rise of generative AI: Threat to academic integrity? Journal of Responsible Technology .
- [5] Jie ZHOU^{1,3},Pei KE²,Xipeng QIU^{1,3}, Minlie HUANG², Junping ZHANG^{1,3}. (2023) ChatGPT:potential, prospects, and limitations. Frontiers of Information Technology & Electronic Engineering.
- [6] Ahmed Tlili¹ , Boulus Shehata¹ , Michael Agyemang Adarkwah^{1*}, Aras Bozkurt² , Daniel T. Hickey³ , Ronghuai Huang¹ and Brighter Agyemang⁴. (2023) What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. Tlili et al. Smart Learning Environments. <https://doi.org/10.1186/s40561-023-00237-x>