International Journal of Research Studies in Computer Science and Engineering (IJRSCSE)
Volume 10, Issue 1, 2024, PP 1-7
ISSN 2349-4840 (Print) & ISSN 2349-4859 (Online)
https://doi.org/10.20431/2349-4859.1001001
www.arcjournals.org

ChatGPT Related Principles and Future Development

Hongyan Sun, Bin Li

South China Normal University Guangzhou

Corresponding Author:** ***Hongyan Sun, *South China Normal University Guangzhou*

Abstract: *As GPT technology continues to advance and become more popular, its impact on various fields will gradually become apparent. The emergence of large language models has promoted the development of the field of natural language processing. This article will discuss the relevant principles of ChatGPT (Generative Pre-trained Transformer) and its model architecture, then introduce the new paradigm brought by GPT, the development process of each stage of GPT and the data sets used, and then introduce the core technology of ChatGPT. It includes prompt learning and instruction fine-tuning, thought chain and human feedback reinforcement learning. Finally, we combines the problems existing in reality and introduces the limitations and future development of large language models.*

Keywords: *GPT; large language model; natural language processing; ChatGPT; prompt learning, new paradigm*

1. Introduction

Natural language processing is an important direction in the field of artificial intelligence. Its research is to achieve effective communication between humans and computers using natural language. The emergence of ChatGPT based on GPT (Generative Pre-trained Transformer) has made it a major research hotspot at present. Early natural language processing was based on rules to build vocabulary, syntax and semantic analysis, question and answer, chat and machine translation systems; later, statistical-based methods became popular, and the main idea at that time was to build machine learning systems based on artificially defined features and use data After learning, the parameters of the machine learning system were determined; subsequently, the emergence of neural network natural language processing, and the proposal of deep learning, researchers modeled deep learning and conducted end-to-end training. Currently, they have achieved results in different fields. During this period, the language model based on the MLP (multi-layer perceptron) used position embedding representation of words and used it to fit probability functions, which improved generalization and efficiency, but it was difficult for MLP to capture local information. The emergence of CNN (convolutional neural network), which uses one or more convolution kernels to sequentially convolve local input sequences, can obtain better local features, but it is difficult for CNN to capture long-range features. Then the language model based on RNN (Recurrent Neural Network) was proposed. RNN is a type of neural network used to process sequence data. Its essence has memory, and the output depends on the current input and memory. It is specially used to deal with timing problems. and sequence-to-sequence problems, characterized by the ability to capture long-term dependencies in sequences. Since then, many variants of RNN have been extended, including long short-term memory network (LSTM), gated recurrent unit (GRU), bidirectional RNN (BiRNN), etc. Until 2017,Google proposed the transformer model based on the self-attention mechanism that can solve the sequence-to-sequence problem by relying only on the attention mechanism, creating a number of new records. Even though the Transformer training time is short, it still has Good performance. Therefore, researchers started various studies based on Transformer. The emergence of GPT and BERT (Bidirectional Encoder Representation from Transformers) once again pushed the research focus of Transformer to the top. The overall architecture of the transformer model is shown in Figure 1. As a typical application of large language model GPT, ChatGPT has greatly promoted the field of natural language processing. Through large-scale pre-training and context understanding, it has the ability to generate natural language text and can conduct conversations and answer questions. and tasks such as providing information.

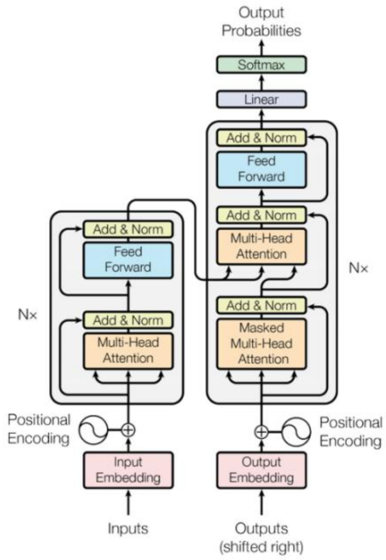

Figure1. *Transformer model architecture*

2. GPT DEVELOPMENT HISTORY AND NEW PARADIGM

GPT is developing and extending its application at a rapid speed. GPT-1 adopts the pre-training model of the Transformer decoder and adopts the paradigm of pre-training and fine-tuning. As the scale of large language models expands, researchers gradually discover the challenges brought by "pre-training + fine-tuning". Slowly, new paradigms are quietly arriving, and ChatGPT is one of the products of this scenario.

2.1. GPT Development History

Google Brain, Google Research and other teams jointly published the article "Attentions Is All You Need" in 2017, replacing the previous Seq2seq model implemented with RNN as the encoder/decoder. Seq2seq with attention added has different degrees of performance in various tasks. promote. In the paper, Google proposed a Transformer model based on the attention mechanism that can process sequence data by relying only on the attention mechanism. The Transformer model has achieved great success in natural language processing. Researchers have started various studies based on the Transformer, and it is now widely used. Applied to various fields of NLP. There are currently many language models based on the Transformer model in NLP, and GPT is OpenAI based on the Transformer decoder. It learns language models from a large amount of unlabeled data by "predicting the future", focusing on language generation and instruction. It is generated by fine-tuning to solve downstream natural language processing tasks.

In June 2018, OpenAI published an article on GPT-1. Before GPT-1, traditional NLP models often used a large amount of data to perform task-related model training on supervised models. There are two main disadvantages: 1. It requires a large amount of annotated data; 2. A model trained on one task is difficult to generalize to other tasks. GPT-1 mainly uses Transformer's decoder as a model. By performing production pre-training on language modeling on different unlabeled text corpora, and then discriminatively fine-tuning each task, it can achieve huge gains in these tasks.

In February 2019, OpenAI released GPT-2, which aims to train a word vector model with stronger generalization capabilities. It did not carry out too many structural innovations and designs on the GPT-1 network, but only used More network parameters and larger data sets. The learning goal of GPT-2 is to use unsupervised pre-training models to do supervised tasks. Based on the above ideas, when the capacity of a language model is large enough, it is enough to cover all supervised tasks, that is to say, all Supervised learning is a subset of unsupervised language models. The GPT-2 data set is taken from highly praised articles on Reddit and is named Web Text. The data set contains a total of about 8 million articles, with a cumulative volume of about 40G, and articles on the design of Wikipedia have been removed. Although there is a 40G data set, the performance on some tasks is no better than random. In practical applications, GPT-2's zero-shot is far from being usable, and its unsupervised learning capabilities still have a lot of room for improvement.

In May 2020, OpenAI released GPT-3, which demonstrated strong performance on many NLP tasks and benchmarks, matching even the state-of-the-art fine-tuning systems in some performances only in zero-shot, one-shot, and few-shot settings. Of course, this is related to GPT-3's 175 billion parameters and 45TB data set. Not only that, but a new paradigm which is Prompt-learning was also created at this time. ChatGPT is further fine-tuned on the basis of GPT-3.5 using the Instruct GPT training method. GPT-3.5 is further trained on the basis of GPT-3 by adding ways such as thought chains, codes and multiple rounds of conversations. The code and thought chains are Training makes the model have stronger logical reasoning capabilities. This optimization and improvement gives ChatGPT thinking understanding, multi-round dialogue and general capabilities.

In March 2023, OpenAI released GPT-4. Compared with GPT-3.5, GPT-4 has context learning capabilities. In context learning, the model will learn and predict based on the context information of the current input data. Unlike supervised learning or fine-tuning, context learning does not require updating the model parameters. Instead, analogy learning and task prediction are directly based on the pre-trained language model, which improves the efficiency of model application. GPT-4 is also a multi-modal upgrade to ChatGPT. GPT-4 also has powerful image understanding capabilities. It can directly conduct visual question and answer on images in a natural language, and its performance on many tasks even exceeds human level.

The impact of the GPT series is obvious to all, and its development process is shown in Figure 2. In the future, GPT and other large models will develop towards a higher level of intelligence and more versatility, not only in dialogue generation, but also in various natural language processing tasks.

Table1. *The development history of GPT*

Model	release time	Parameter quantity	Amount of pre-training data	data set
GPT	June 2018	117 million	About 5GB	Books Corpus, Wikidia
GPT-2	February 2019	1.5 billion	40GB	WebText
GPT-3	May 2020	175 billion	45TB	Common Crawl etc.
GPT-4	March 2023	unknown	unknown	unknown

2.2. GPT New Paradigm

Long before GPT appeared, there were several paradigms: the fully supervised learning paradigm in the non-neural network era, the fully supervised learning paradigm based on neural networks, and the "pre-training + fine-tuning" paradigm. We mainly introduce the "pre-training + fine-tuning" paradigm and the new GPT paradigm which is Prompt-learning.

2.2.1. Pre-training + Fine-tuning

After the introduction of neural networks, most deep learning methods still require a significant amount of manually labeled data, which limits their applicability in many domains with scarce labeling resources. In such case, models that utilize linguistic information in unlabeled data are used to collect more annotations that provide a valuable alternative, and collecting more annotations is a time-consuming and expensive work. In addition, learning good representations unsupervised can also significantly improve the performance in the presence of heavy supervision. With the emergence of pre-trained language models like GPT and BERT, "Pre-training + fine-tuning" has become one of the foundational paradigms in the domain of natural processing. This paradigm explores a semi-supervised approach that uses a combination of unsupervised pre-training and supervised fine-tuning to accomplish language understanding tasks, with the goal of learning a universal representation that can be transferred to in a wide range of tasks.

The dataset Books Corpus was used for unsupervised pre-training in GPT-1, and supervised fine-tuning was performed in various tasks such as natural language reasoning, question answering and common-sense reasoning, semantic similarity, and text classification. The resulting model acquired significant world knowledge and the ability to handle long-range dependencies. It successfully transferred to discriminative tasks, demonstrating that achieving substantial performance gains is indeed possible through the use of informative pre-training. The data sets Books Corpus and Wikipedia are used for pre-training in BERT, and a masked language model is used to achieve pre-trained deep bidirectional representation, which illustrates the importance of bidirectional pre-training for language representation. Therefore, the "pre-training + fine-tuning" paradigm does not require a large amount of supervised downstream task data. The model is mainly trained on large-scale unsupervised data and only requires a small amount of downstream task data to fine-tune a small number of network layers.

The "Pre-training + Fine-tuning" paradigm requires saving a copy of the entire pre-trained model for each downstream task, and inference must be performed in a separate batch. As the scale of language models continues to grow, this introduces several challenges. There are two main challenges:

1. Semantic Drift: To adapt the model to new downstream tasks, it is necessary to introduce new training parameters, resulting in a significant gap between the two-stage objectives. This can lead to a decrease in semantic coherence between the pre-training and fine-tuning stages;
2. Overfitting: Additional parameters are introduced during both the pre-training and fine-tuning stages to accommodate specific task requirements. This can lead to overfitting, particularly when dealing with limited samples, and can reduce the model's generalization capabilities.

2.2.2. Pre-training + Prompt + Prediction

Led by GPT-3 and PET, a new fine-tuning paradigm based on pre-trained language models, prompt-Tuning, is proposed. Its purpose is to avoid the introduction of additional parameters by adding templates, so that the language model can be used in small samples (Few-shot) or zero-shot (Zero-shot) scenario to achieve the ideal effect, which is converting the fine-tuned downstream task target into a pre-trained task.

Prompt-based learning consists of three steps: template construction, label-word mapping, and training.

1. Template Construction: A template containing a [MASK] token related to the given sentence is generated using methods such as manual definition, automatic search, or text generation. This template is then concatenated with the original text to create the input for Prompt-Tuning. By feeding it into a BERT model and using a pre-trained MLM classifier, you can obtain the predicted token probability distribution;
2. Label-Word Mapping: Because in the [MASK] part we are only interested in some words, we need to establish a mapping relationship. For example, if the predicted word belongs to the "positive" category, it is considered to the great category.
3. Training: According to the labels produced by Verbalizer, the predicted probability distribution of the specified keywords can be obtained, and cross information entropy is used for training. At this point, because only the pre-trained MLM head is fine-tuned, the over-fitting problem is avoided. Continuous Template Construction:

There are many engineering choices for prompt-based methods, and template construction methods are one of them. Templates are additional short texts marked with [mask]. Typically, each sample has only one template, and there are discrete and continuous template construction methods.

1. Discrete Template Construction: This method directly concatenates discrete characters with the original text and keeps the word embeddings of discrete characters fixed during training.
2. Continuous Template Construction: The continuous template construction method aims to allow the model to continuously adjust template parameters based on specific contextual semantics and task objectives during training. The prompt information becomes a continuous vector unconstrained by word embeddings, rather than existing words in the vocabulary.

In general, discrete methods do not require the introduction of any additional parameters, while the continuous methods require the introduction of a small number of parameters and allow the model to update parameters during training introduced human feedback reinforcement learning (RLHF) methods.

3. The Core Technology of ChatGPT

OpenAI has further introduced the Instruct GPT model. The model uses supervised instruction fine-tuning and human feedback reinforcement learning (RLHF) methods. It utilizes a proximal policy optimization approach to enable self-improvement and updates of the model, ensuring better alignment with user intent.

3.1. Prompt Learning and Instruction Fine-tuning

Prompt learning involves designing a template that embeds the original sentence, including a [MASK] token related to the given sentence, to create input for Prompt-Tuning. The created input should be feed into the BERT model, and by using a pre-training MLM classifier, the probability distribution of the predicted token could be obtained. Prompt learning aims to stimulate the completion ability of the language model.

Instruction fine-tuning is an enhanced version of prompt learning, using manually annotated instruction data for supervised training to stimulate the model's understanding ability. Instruction fine-tuning involves further training LLMs (Language Model Models) on a dataset composed of (instruction, output) pairs. In this context, instructions represent human commands to the model, and outputs represent the expected results following those instructions. This approach helps bridge the gap between the LLM's next-word prediction objective and the goal of making LLMs follow human instructions. It enhances the model's generalization capabilities, even in cases where no explicit examples are provided for following instructions.

The ability of ChatGPT to respond to human instructions is a direct outcome of instruction fine-tuning. For new tasks that have not been trained by the model, we only need to design a language description of the task and provide task instances to the model, and the model can learn new tasks from the given scenario and provide appropriate responses. Chain of Thought

3.2. Chain of Thought

Strong logical reasoning ability is one of the core abilities derived from the "emerge with intelligence" of large language models, and the key to reasoning ability is the chain of thinking. Google researchers, including Jason Wei, introduced the concept of Chain of Thought, which involves inserting a series of intermediate reasoning steps into small-sample prompt learning to effectively enhance a language model's reasoning capabilities. In other words, the Chain of Thought prompts is to decompose a multi-step reasoning problem into many intermediate steps, allocate more calculations, generate more tokens, and allow the model to accumulate correct inferences throughout the reasoning process. These accumulated tokens are then concatenated to find the solution, resulting in a significant improvement in the model's accuracy when dealing with complex reasoning tasks.

3.2.1. Zero-shot-CoTChain of Thought

The Zero Shot Chain of Thought (Zero-shot-CoT) is a follow-up study of CoT prompting, which introduces a very simple zero-shot prompt. By appending the words "Let's think step by step" to the end of the question, the large language model was able to generate a Chain of Thought that answered the question. From this Chain of Thought, they can extract more accurate answers.

3.3. Human Feedback Reinforcement Learning

Reinforcement Learning from Human Feedback RLHF is an important key technology for ChatGPT to understand human instructions and align human behavior. This algorithm consists of two stages in the context of reinforcement learning:

1. Reward Model Training: This stage aims to obtain a reward model that fits human preferences. A reward model is a model used in reinforcement learning to describe and calculate reward values for actions. For the given input information and the two results generated by the model, the annotator needs to choose the generated result that is more suitable for human preferences. Everyone scores answers to questions differently, as a result it is impossible to use the same numerical value to score every answer. The reward model interacts with human experts to gather feedback on the quality of generated responses, further improving the generation capabilities and naturalness of the large language model. Unlike the supervised model, the reward model makes the generated text more natural and realistic through scoring, which further improves the generation ability of large language models. The reward model is trained by calculating the probability of preference based on the difference in reward values between the two responses in each sample.

2. Proximal Policy Optimization: As an advanced reinforcement learning algorithm, proximal policy optimization aims to enable intelligent systems to learn optimal action strategies to maximize a specific reward signal by interacting with the environment. Proximal Policy

Optimization (PPO) is a (model-free) policy optimization gradient-based algorithm that aims to learn a policy that maximizes the cumulative reward obtained based on testing during training. Given a trained reward model, the parameters of ChatGPT will be regarded as a policy, trained under the framework of reinforcement learning, so that it generates results with higher reward scores.

4. ChatGPT's Real Challenges and Future Development

Although ChatGPT has made significant progress in dialogue generation and the generalization ability of GPT has been significantly improved, ChatGPT still faces some limitations and challenges, which are representative of many large language models.

4.1. Inaccurate Semantic Understanding

ChatGPT's semantic understanding ability is limited and it struggles to grasp the deep meanings of language. This will cause ChatGPT's answers to be inaccurate and even cause misunderstandings.

4.2. Feasibility of Reasoning

Feasibility is one of the crucial limitations of current large language models. Firstly, the models may generate solutions to problems that do not exist and that is illusions. Secondly, the model may generate text that conflicts with common sense or professional knowledge; Additionally, when faced with uncertain or unknown facts, they may attempt to build arguments without a basis, leading to the generation of inaccurate information.

4.3. Security

ChatGPT has security issues, and it could be used for malicious purposes, such as disclosing users' personal information for online fraud or personal attacks.

A large amount of training data for large language models comes from unannotated text on the Internet, which inevitably introduces harmful falsehoods or discriminatory content. Constructing appropriate security evaluation standards and corresponding training data is a significant challenge for the practical application of large language models.

4.4. Evaluation and Future Development of Large Language Models

With the rapid development of large language models, the evaluation of large language models also brings challenges. In the future, large language models will gradually enter all aspects of human life. It is important to obtain a reasonable and reliable evaluation method. As these models continue to increase in scale and enhance their cross-lingual abilities while integrating multiple modalities, it also raises concerns about computational resource support. There is no doubt that large language models such as ChatGPT are driving the innovation of emerging applications and the development of human-computer interaction.

5. Conclusion

This article has explored the importance of ChatGPT in the field of natural language processing and its related technical principles, and discussed the development process of the GPT model and the proposal of a new paradigm (hint learning) and core technology. Furthermore, it has analyzed the ubiquity of ChatGPT and large language models limitations and challenges, including inaccurate semantic understanding, feasibility of reasoning, security, and evaluation of large language models. In the future, with the rapid development of large language models, large language models are poised to transform the work and lifestyles of millions of people. Therefore, addressing the existing challenges requires collaborative efforts from various sectors of society.

References

[1] 张熙,杨小汕,徐常胜.ChatGPT及生成式人工智能现状及未来发展方向[J/OL].中国科学基金:1-8[2023-11-01].https://doi.org/10.16262/j.cnki.1000-8217.20231026.002.

[2] 秦涛,杜尚恒,常元元等.ChatGPT工作原理、关键技术及未来发展趋势[J/OL].西安交通大学学报,2024(01):1-11[2023-11-01].http://kns.cnki.net/kcms/detail/61.1069.T.20231016.1403.012.html.

[3] 冯宇涛,强继朋,李云等.基于prompt的文本可读性评估[J].扬州大学学报(自然科学版),2023,26(03):56-60.DOI:10.19411/j.1007-824x.2023.03.010.

[4] 陶建华,聂帅,车飞虎.语言大模型的演进与启示[J/OL].中国科学基金:1-9[2023-11-01].https://doi.org/10.16262/j.cnki.1000-8217.20231026.005.

[5] 舒文韬,李睿潇,孙天祥等.大型语言模型：原理、实现与发展[J/OL].计算机研究与发展:1-12[2023-11-01].http://kns.cnki.net/kcms/detail/11.1777.TP.20231008.1142.004.html.

[6] Radford A, Narasimhan K, Salimans T, et al. Improving language understanding by generative pre-training[J]. 2018.

[7] Radford A, Wu J, Child R, et al. Language models are unsupervised multitask learners[J]. OpenAI blog, 2019, 1(8): 9.

[8] Brown T, Mann B, Ryder N, et al. Language models are few-shot learners[J]. Advances in neural information processing systems, 2020, 33: 1877-1901.

[9] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.

[10] Devlin J, Chang M W, Lee K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding[J]. arXiv preprint arXiv:1810.04805, 2018.

[11] Ouyang L, Wu J, Jiang X, et al. Training language models to follow instructions with human feedback[J]. Advances in Neural Information Processing Systems, 2022, 35: 27730-27744.

Citation: *Hongyan Sun, Bin Li. "ChatGPT Related Principles and Future Development" International Journal of Research Studies in Computer Science and Engineering (IJRSCSE), vol 10, no. 1, 2024, pp. 1-7. DOI: https://doi.org/10.20431/2349-4859.1001001.*