A Study of the Usage of Artificial Intelligence in Scientific Research and Education

Ruisen Peng

Shenzhen Middle School, Shenzhen, Guangdong, 518000, China

Abstract. Nowadays, the speed of spearing of artificial intelligence is getting faster. Likely commercialization can make people's lives more convenient, if scientific research gets help from artificial intelligence, the efficiency of scientific research will significantly increase. This paper shows a general situation of the usage of artificial intelligence in scientific research and education. By combining the essential information from recent papers about this field, this paper analyses the crucial problem scientists face and the general solving method. The main problems are the randomness and the nature of Blackbox, and the general solving method is to add a part of the model to let the model "understand" the principles and theorems that are in the field. Also, the paper generates a general situation of the usage of artificial intelligence in education, which is speaking out at an increasing speed. The paper is a conclusion of the situation that the character of artificial intelligence in scientific research and education. This paper expects to offer a clear cognition and deep understanding about this field.

1 Introduction

With the rapid development of artificial intelligence, its application in various fields is becoming increasingly widespread. Also, artificial intelligence plays a significant role in scientific research. It can help scientists analyze data, construct models, and design experiments. What is more, it can accelerate scientific findings and purpose innovation in scientific research. In today's scientific field, facing the increasing amount of data and complex problems, artificial intelligence should be a crucial tool to deal with these challenges [1].

The usage of artificial intelligence keeps developing. Many subjects use artificial intelligence in their research, such as using artificial intelligence to intimate experiments and solve complex equations [2]. There are many papers on this topic. After reading some of these papers and concluding a result, this paper found that the general situation of artificial intelligence in science research is still a crucial problem. For example, the randomness and the nature of BlackBox will interfere with the reliability of the research outcome, and it is unignorable in scientific research [1]. Recent research tries to find a zero-error technique of artificial intelligence, which basis on the principle instead of the direct relationship between problem and answer.

Corresponding author: haomengjiong@ldy.edu.rs

The purpose of this paper is to gain general knowledge of artificial intelligence in science research. The first step is introducing the basic principle of general artificial intelligence, to show the advantages and disadvantages of artificial intelligence in science research. Then, the paper specifies artificial intelligence, which is specialized in scientific research. The part will show the adoption of this kind of artificial intelligence and the difference between specialized artificial intelligence and general artificial intelligence. The third part will how these specialized artificial work in scientific research. What is more, some examples will be shown to illustrate the part. Last but not least, the article will jump out of the limited topic of scientific research and discuss the usage of artificial intelligence in scientific education, which is the basis of the development of the science field.

2 The main problem that prevents artificial intelligence play a role in scientific research

The main types of commercialized artificial intelligence are large language models, natural language processes, and computer vision. In fact, many advantages let it spread out widely. First of all, all commercialized artificial intelligence has an easy operational system, which allows users can use the artificial intelligence easily. What's more another advantage of commercialized artificial intelligence is that the application of civilian AI is very extensive, which means a model can deal with many problems from diverse fields. However, there are several defects in these common artificial intelligence models, and these models cannot deal well with scientific research. To begin with, the data dependency in AI and the following randomness, are unignorable in common artificial intelligence [3]. For example, a large language model needs a large amount of useful data to get a valuable answer. However, the quality of content on the Internet is varied, and this content will affect the answer about the field, especially when the number of kinds of content is relatively small. Content about science also is flown like that, the content in a particular field is very small, and it will make the answer unreliable. What's more the nature of Blackbox also is a critical problem [2]. Nowadays artificial intelligence, user cannot trace how the answer comes, so user cannot directly verify whether the model answer is correct or not. Whatever the usage in complex calculation, solving equations, or the usage in imitation and prediction, accuracy is one of the most important factors scientists need to heed. However, accuracy is hard to detect in scientific research, a field that is full of unknowns, when scientists cannot check if the results are correct via checking the one-by-one step. Even if artificial intelligence solves a problem correctly, it also needs researchers to spend a lot of time to check the outcome. As a result, the nature of Blackbox will make artificial intelligence unable to be used easily in scientific research. Last but not least, the lack of flexibility also prevents the usage of artificial intelligence in scientific research, because the same model's performance is unstable in different tests, and it also will take a lot of time for researchers to modify programs.

3 The improvement to let artificial intelligence "understand" theorem

To prevent the common disadvantage, the specified artificial intelligence has to improve the basic logic of artificial intelligence, just focus on given data. To make the output reliable and easy to verify, it has to import a critical factor- theorem. The only way to make artificial intelligence strict enough is to let artificial intelligence 'think' with the infallible theorem, instead of training by the of the content about the example of these theorem on the internet.

However, it also solicits two necessary problems: how to let the artificial intelligence model understand and think with the theorem, and how to make the artificial intelligence

model understand the question with a known theorem. First of all, to let artificial intelligence deal well with such a great amount of theorem, some model such as decision trees and random forests are abandoned, because their performance is relatively worse in situations involving a large volume of data [4,5]. On the other hand, a neural network's performance is better in a larger volume of data, and it also can use different neural to represent different theorems, so it widely be used in scientific research [6]. To let artificial intelligence "understand" theorems, scientists use varied methods, but the purpose is the same using an additional part to deal with theorem and avoid randomness.

Here are several examples. The artificial intelligence model of Google- Alphageometric, using symbolic deduction to make artificial intelligence "understand" theorem [7]. To be specific, they used code-like statements to represent theorem and mathematical information. And the team radon chose nearly one billion theorems to train the model. Every theorem stands for a node, if the two theorems can prove each other, these nodes will be linked. Given the question and condition, Alphageometrics can use a nature language model to change it to a symbolic statement and use the nature language process model to change it to code-like statements. After that, it will try to find a way to link the condition and problem randomly, and it also will create some new nodes to try to solve the problem.

4 Advantages of models and application analysis

Scientific research model has several significant advantages. First, since each process relies on mathematical theorems, the model excels in terms of accuracy. Secondly, this process can combine mathematical symbols with natural language and transform them into clear mathematical steps, greatly enhancing the readability of the model. Although the model was primarily trained to solve geometric problems, it can equally be trusted to handle other types of problems in the field of mathematics, such as arithmetic, number theory, and combinatorics.

However, this model also has some shortcomings. First, because the solutions obtained by the model may be very complex and the path to solve the problem has a certain degree of randomness, this makes it difficult for human readers to understand these solutions. Second, due to the limited number of nodes, the model cannot handle all geometric problems. In addition, in order to expand the application scope of the model to solve more types of problems, including calculus and linear algebra, the model needs to be trained with a large number of theorems, and the huge number of nodes will lead to a sharp increase in the time and computing power required for calculations.

Despite these shortcomings, the model successfully solves the two main problems faced by common artificial intelligence - randomness and black-boxness, and can therefore be effectively used in applications in the field of mathematics. Another model in "Machine Learning Rigor from Field Theory to Poincare Conjecture" takes a different approach, namely reinforcement learning. In this approach, the model is rewarded for completing correct proofs and is punished for errors. After training itself over a period of time, the model will be able to use theorems to correctly solve problems through other algorithms

5 Analysis of example

Artificial intelligence is becoming increasingly prevalent across various domains of scientific research, with a particularly significant application in the field of mathematics. One of the major challenges in mathematics involves solving complex partial differential equations (PDEs), which are often resistant to traditional analytical methods. The inherent complexity of these equations, combined with the limitations of conventional programming approaches, creates a demand for innovative solutions. Theoretically, by employing models based on

neural networks, scientists can approximate solutions to virtually any partial differential equation. Furthermore, AI has shown promise in addressing a range of optimization problems, thereby expanding its utility beyond purely mathematical inquiries.

In the realm of materials science, the quest for new and useful materials is of paramount importance. However, the vast array of potential combinations of different materials complicates this endeavor significantly. Artificial intelligence can play a crucial role in this context by analyzing a multitude of factors associated with various materials and their combinations. Through sophisticated algorithms and data modeling, AI can predict the properties of unknown materials, facilitating the discovery of innovative substances that could have valuable applications across industries.

A noteworthy example of AI's capabilities in scientific research is the development of AlphaFold, an advanced artificial intelligence model specifically designed to predict protein structures [8]. Proteins, as organic polymers, possess a complex arrangement of composition factors, making it exceedingly challenging for researchers to accurately predict the stability of new proteins. AlphaFold has not only addressed this significant hurdle but has also demonstrated superior performance compared to other contemporary AI models in the field. By integrating knowledge from both physics and biology into its framework, scientists have enabled AlphaFold to generate protein structures responsibly and with high reliability [9].

Moreover, AlphaFold is adept at analyzing extensive datasets of known protein structures, allowing it to uncover patterns that indicate the stability of various proteins. The model employs additional methodologies, including simulating the evolutionary processes of proteins and applying physical principles to create three-dimensional representations of these biomolecules. After a rigorous training period, AlphaFold has achieved remarkable accuracy in its predictions, with its simulation capabilities extending to the atomic level. The team behind this innovative model not only utilized it to predict new protein structures but also made the code open-source, allowing researchers and developers worldwide to leverage and enhance its functionality. This collaborative approach has led to the paper detailing AlphaFold being cited over 26,000 times, underscoring its widespread impact and significance within the bioinformatics community.

6 Artificial intelligence in education

Beyond its applications in scientific research, artificial intelligence also plays a critical role in the preparatory aspects of scientific inquiry, particularly in the field of education. The importance of AI in educational settings is increasingly recognized, as it encompasses three primary areas: appreciation, guidance, and learning. In terms of educational management, AI can significantly alleviate the administrative burdens faced by teachers. For instance, it can assist in evaluating students' homework, providing timely feedback, and analyzing performance data to give educators insights into their students' interests and overall progress. Additionally, AI can serve as a tool to detect instances of academic dishonesty, such as cheating during assignments and examinations.

When it comes to teaching methodologies, AI's capability to process vast amounts of data allows for the development of tailored teaching plans for specific classes or even individual students. This customization enables educators to identify students who may benefit from additional extracurricular learning opportunities, as well as those who might be struggling to keep up with the curriculum. Such personalized approaches can enhance the learning experience for students from diverse backgrounds, ensuring that each learner receives the support they need.

AI also offers substantial benefits directly to students, providing daily feedback on their performance to help them track their progress and identify areas for improvement. More importantly, AI can recommend course selections and formulate long-term learning plans

based on each student's unique needs and aspirations. The implementation of these functions can be accomplished without relying on overly complicated software or substantial computational resources, significantly simplifying the educational process. However, the challenge remains for educators to define the boundaries between AI assistance and instances of academic dishonesty that may arise in conjunction with its use.

A prominent example of AI's integration into education is the Khan Academy, which has recently introduced an innovative AI assistant known as Khanmigo AI Assistant [10]. This AI assistant enhances the educational experience in several key ways, including real-time tutoring, personalized learning suggestions, and problem explanations. The real-time tutoring feature empowers students to seek assistance from AI whenever they encounter difficulties, thus improving educational efficiency and alleviating some of the pressures on teachers. Additionally, personalized learning suggestions facilitate the self-learning process by helping students locate appropriate learning materials. By monitoring individual learners' performance, the system can automatically adjust the difficulty level of the materials, further enhancing learning outcomes.

Finally, the problem explanation feature allows students to reflect on their learning challenges. The AI model is capable of reviewing students' mistakes, enabling them to easily identify their weaknesses and areas for improvement. This comprehensive approach to education, driven by artificial intelligence, holds great potential for transforming the learning experience, making it more accessible, personalized, and effective for students worldwide.

7 Conclusion

To gain a deeper understanding of the current application status of artificial intelligence in scientific research and education, this article conducts a systematic analysis of a large number of relevant papers published recently. By summarizing the common points of these papers, this article draws a general conclusion. First, the study discovered and pointed out the main flaws that hinder the role of general artificial intelligence in scientific research. Specifically, these flaws are reflected in black box characteristics and randomness. To overcome these two problems, scientific AI models incorporate domain-specific principles and theorems, allowing AI to better understand the context and purpose of its application.

In fact, there are many examples of the application of artificial intelligence in scientific research, covering many fields. In addition to its wide application in scientific research, the application of artificial intelligence in the field of education has become increasingly prominent, which is mainly reflected in three aspects: appreciation, guidance and learning. Khan Academy is a successful example of how artificial intelligence can improve educational outcomes.

Although scientific artificial intelligence models can effectively deal with the above two problems, the diversity of solutions makes their application scope different. If a universally applicable model could be developed, it would mark another important breakthrough. With the rapid development of artificial intelligence technology, it is expected to cover more aspects of the education field in the future, providing new opportunities for educational innovation and improvement.

Reference

1. W.J. Eschenbach, Transparency and the Black Box Problem: Why We Do Not Trust AI. Philos. Technol. 34, 1607–1622 (2021).
2. Y. Xu et al., Artificial Intelligence: A Powerful Paradigm for Scientific Research. The Innovation 2, 4, 100179 (2021).

3. D.E. Maydan, J.L. Hennessy, M.S. Lam, Efficient and Exact Data Dependence Analysis, in Proceedings of the ACM SIGPLAN '91 Conference on Programming Language Design and Implementation (pp. 1-12), Toronto, Ontario, Canada: Association for Computing Machinery (1991).
4. L. Breiman, Random Forests. Machine Learning 45, 5-32 (2001).
5. J.R. Quinlan, C4.5: Programs for Machine Learning. Morgan Kaufmann, San Francisco, CA (1993).
6. A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet Classification with Deep Convolutional Neural Networks. (2012).
7. A. Davies, P. Veličković, L. Buesing et al., Advancing Mathematics by Guiding Human Intuition with AI. Nature 600, 70–74 (2021).
8. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, D. Hassabis et al., Highly Accurate Protein Structure Prediction with AlphaFold. Nature 596, 583-589 (2021).
9. L. Chen, P. Chen, Z. Lin, Artificial Intelligence in Education: A Review, in IEEE Access 8, 75264-75278 (2020).
10. Khan Academy. Available online: www.khanacademy.org (accessed on 2024.9.2).