THE IMPACT OF LARGE-SCALE LANGUAGE MODELS ON THE FUTURE DEVELOPMENT OF ACCOUNTING

MingFong Wong
The University of Hong Kong, HongKong 999077, China.
Corresponding email: mingfongwong@126.com

Abstract: With the rapid development of Large-scale language models (LLMs), attempts to apply artificial intelligence (AI) to various fields have become increasingly common, and accounting-related professions have been identified as among those vulnerable to replacement by AI. This study aims to provide a concise overview of the development of AI and the technological framework of LLMs, and to explore in depth the implications of their development for the accounting. This study argues that the power of LLMs lies in their vast amount of data and large number of parameters; however, the interpretability of the models is challenging and the application of general-purpose AI to the accounting profession requires further development. Although it is difficult for general-purpose LLMs to have a direct impact on accounting, the development of AI will change the accounting environment and specific objects, which will reshape the content of accounting work, the competence framework of accountants and the market for accounting services.

Keywords: Large-scale language model; Accounting

1 INTRODUCTION

Large-scale language models (LLMs), a neural network-based natural language processing technique, have gradually become a research focus in the field of natural language processing since 2018. With the continuous development of LLMs, researchers have begun to explore their applications in various fields, including but not limited to natural language generation, machine translation, dialogue systems, text classification, and information retrieval. Among them, the Generative Pre-trained Transformer (GPT) family of models, as one of the most popular LLMs, has achieved remarkable results in various natural language processing tasks[1]. However, with the expansion of the model size, a number of challenges and problems such as computational complexity, data bias, and model interpretability have emerged, making a comprehensive and deep understanding of these models particularly important.

In the field of accounting, LLMs have a wide range of potential applications. LLMs can automatically classify, capture and report financial data, improving the accuracy and efficiency of accounting and reducing accounting costs. In addition, LLMs can help accountants process large amounts of financial data more quickly, for example by automatically identifying and correcting erroneous financial data and reducing manual intervention. These applications will drive the accounting industry towards digital transformation and promote the development of the accounting industry. Overall, LLMs will drive the accounting industry towards digital transformation, improve accounting accuracy and efficiency, reduce accounting costs, and promote the development of the accounting industry. However, it also faces some challenges and problems, such as privacy and security issues, model interpretability issues, and model reliability and stability issues. Therefore, it is particularly important to conduct in-depth studies and research on these issues.

2 FUNDATION OF LLMS

2.1 Concept and History of LLMs

LLMs are large-scale machine learning models capable of processing and generating natural language[2]. These models are generated by deep learning techniques that use large amounts of training data and powerful computational capabilities to produce models capable of understanding and generating human language. The heart of LLMs is the Transformer model, a neural network model based on a self-attention mechanism. The self-attention mechanism allows the model to assign different attentional weights based on the relevance of different positions in the sequence when processing the input sequence, and this mechanism allows the model to better capture long-range dependencies, thus improving the ability to understand and generate language.

As LLMs have been developed, the size of the model and the number of parameters have increased. This increase in size has allowed the models to handle more complex linguistic tasks and to demonstrate strong capabilities in several domains. This development trend is of great significance for natural language processing research and provides new opportunities and challenges for future research.

The history of LLMs can be traced back to around 2010, when researchers began experimenting with using neural networks to train language models. These early models could usually only handle relatively small datasets, but with continuous improvements in hardware and algorithms, the performance of LLMs has improved significantly. In 2018, researchers proposed the Transformer model, a neural network model based on the self-attention mechanism, which significantly improves the performance of natural language processing tasks. The emergence of the Transformer model laid the foundation for the development of LLMs. In the same year, OpenAI released the first generation of GPT

(Generative Pre-trained Transformer) model, which is a large-scale language model with 117 million parameters[3].The GPT model is trained by pre-training and fine-tuning, and is capable of generating high-quality text, and has achieved significant results in several natural language processing tasks. In 2019, OpenAI released an improved GPT-2 model with more parameters and a deeper transformer architecture with 1.5 billion parameters.GPT-2 performs well on tasks such as text generation, Q&A, translation and summarisation.In 2020, OpenAI further released the GPT-3 model with 17.5 billion parameters[4]. The GPT-3 model has demonstrated strong capabilities in tasks such as natural language processing, machine translation, Q&A systems, and even complex tasks such as text generation and code writing; in early 2023, OpenAI further released the GPT-4 model, which is one of the most advanced language models available, with higher reliability, creativity, and collaboration GPT-4 has 1.8 trillion parameters, making it one of the largest known language models. At the same time, the performance of GPT-4 has been greatly improved to handle more complex linguistic tasks and has demonstrated its power in a number of domains.

As LLMs continue to evolve, researchers are exploring even larger models and more training data. Training these models requires significant computational resources and datasets, but they have nevertheless made remarkable progress. The development of LLMs has attracted considerable attention not only in academia but also in industry. However, LLMs also face some challenges and controversies. One of the challenges is the huge amount of computational resources required for the models and the relatively high environmental and energy consumption. In addition, LLMs can suffer from bias and inaccuracy in text generation, which requires further research and improvement. Overall, the concept and development history of LLMs demonstrate the great potential of artificial intelligence in the field of natural language processing. With the continuous advancement of technology and in-depth research, we can foresee that LLMs will play an increasingly important role in various fields to provide people with better language communication and understanding. However, to achieve this goal, we need to overcome the above challenges to ensure the sustainable development and widespread application of LLMs. This will undoubtedly be an important direction for future research.

2.2 Application Scenarios and Advantages of LLMs

LLMs are widely used in various fields due to their powerful language understanding and generation capabilities. The following are some of the main application scenarios and advantages of LLMs:

1) Text generation: LLMs can generate high-quality text, which is of great significance in the fields of automatic writing, content creation, literary creation, etc. It helps writers and creators to provide inspiration and generate a large amount of textual content that can be tailored to specific topics or styles.

2) Machine Translation: LLMs excel in machine translation tasks. It can translate from one language to another while maintaining sentence fluency and accuracy. This is important for cross-language communication, multinational companies and cross-cultural communication. The translation ability of the large language model can help people to better understand and communicate between different languages.

3) Question and answer system: LLMs can be used to build an intelligent question and answer system that can answer questions posed by the user. It can understand the intent of the question and generate accurate answers based on pre-trained knowledge and language models. This is very useful in areas such as online customer service, intelligent assistants and knowledge base retrieval.

4) Sentiment analysis: LLMs can analyse sentiment and emotion in text. It can identify positive, negative or neutral sentiments in text and help companies and organisations understand users' attitudes and emotional tendencies towards products, services or events. This is important in areas such as market research, opinion monitoring and sentiment analysis.

5) Text summary: LLMs can automatically generate a summary or overview of a text. It can extract key information from long articles or documents and produce concise and accurate summaries. This is very useful in areas such as information retrieval, news reporting and knowledge management.

The advantages of LLMs are mainly reflected in the following aspects

1) Contextual understanding: LLMs are able to understand the contextual relationships in a text, leading to a better understanding of semantics and context. It can capture long-range dependencies, providing more accurate language understanding and generation capabilities.

2) Multi-domain adaptability: The LLMs use a large amount of multi-domain data during pre-training and fine-tuning, which makes the model highly adaptable. It can be applied to tasks in different domains and achieve good performance on various language processing tasks.

3) Flexibility and adaptability: The LLMs can be fine-tuned and optimised for specific tasks to improve performance and accuracy. It can be customised to meet different needs and adapt to different application scenarios.

4) Large-scale data processing: TheLLMs can process large amounts of textual data, enabling it to learn richer and more accurate language patterns. It can extract knowledge and information from massive data and apply it to various practical scenarios.

5) Continuous improvement and iteration: The development of LLMs is a process of continuous iteration and improvement. Researchers and developers can improve the performance and capability of the models through continuous training and optimisation to make them better meet practical needs.

Overall, LLMs have a wide range of applications in text generation, machine translation, question and answer systems, sentiment analysis and text summarisation. Their contextual understanding, multi-domain adaptability, flexibility and adaptability, and ability to handle large amounts of data make them important tools and techniques for natural language

processing. As technology advances and research intensifies, LLMs will continue to evolve and provide even more powerful language processing capabilities.

2.3 Prospects of LLMs in the Accounting

Accounting is an information-intensive field that has to deal with a large amount of textual data, such as financial statements, tax documents, audit reports, etc. LLMs can help accountants to analyse this textual data, identify potential problems and risks, and provide appropriate suggestions and solutions. Here are some possible application scenarios:

1) Financial statement analysis: LLMs can help accountants analyse financial statements, identify potential problems and risks, and provide appropriate suggestions and solutions. For example, LLMs can automatically identify anomalous data and trends in financial statements to help accountants better understand the financial position and business operations.

2) Tax planning: Large language models can assist accountants in tax planning by identifying tax incentives and reductions and providing appropriate suggestions and solutions. For example, big language models can automatically identify key information and provisions in tax documents to help accountants better understand tax policies and regulations.

3) Audit and compliance: Large language models can help accountants with audit and compliance, identifying potential violations and risks, and providing suggestions and solutions. For example, LLMs can automatically identify unusual data and trends in audit reports to help accountants better understand audit results and business operations.

4) Intelligent accounting: LLMs can help accountants perform intelligent accounting, automatically identify and classify financial data, and generate relevant accounting documents and reports. For example, LLMs can automatically identify key information and data in financial documents to help accountants better manage and process financial data.

The application of LLMs in accounting is very promising and can help accountants improve efficiency and accuracy, and reduce costs and risks. With the continuous advancement of technology and in-depth research, LLMs will play an increasingly important role in the accounting field, providing accountants with better tools and support.

3 IMPACT OF LLMS ON THE ACCOUNTING

Regardless of management activity theory or information systems theory, the basic work of accounting is the generation of accounting information[5].LLMs such as ChatGPT, as a representative product of generative AI, are not accounting-specific tools and are not designed to be used directly for accounting information processing. However, on the one hand, the development of generative AI applications will bring about changes in economic and social activities, business models and organisational forms, just like the Internet and digital intelligence, which will be deeply embedded in social and market operations, and at the same time will reshape the operation mode of enterprises as well as the business model, etc., which will pose the adjustment requirements for accounting information processing, On the other hand, generative AI will also intervene in the accounting information processing process and even combine with cloud computing and accounting robots, which will redefine the logic of accounting information processing. The following are some typical implications of the big language model for accounting.

1) In accounting information processing, accounting information processing is the core part of accounting work, including data entry, classification, reconciliation and other tasks. Traditional accounting information processing usually requires a lot of manpower and time, and is prone to errors and delays[6]. However, the advent of large language models has changed this situation. By training the model, automated accounting information processing can be achieved, greatly improving efficiency and accuracy. Models can identify and extract key information, automatically classify and reconcile data, reducing manual errors and time costs.

2) In terms of financial reporting analysis, financial reporting is the main way for companies to communicate financial information to the outside world, which is of great importance to investors, creditors and other stakeholders in making decisions. However, financial reports typically contain a large amount of data and complex information that requires in-depth analysis and interpretation. Big Language models can provide more comprehensive and accurate insights and predictions by analysing financial reporting data. Models can automatically interpret and parse financial reports, identify potential risks and opportunities, and generate detailed analysis. This provides accountants and financial analysts with powerful tools to better understand and evaluate financial reports and provide more accurate decision support.

3) In terms of risk management, risk management is an important part of corporate governance and is essential for protecting business interests and sustainable development. Big language models can help companies develop risk management strategies by analysing large amounts of financial and non-financial data, identifying unusual patterns and trends, and providing early warnings and recommendations. Models can identify potential risk factors, predict the likelihood and impact of risks, and provide a basis for business decisions. This objective and accurate method of risk analysis can help entities to better manage potential risks and reduce losses and uncertainty.

4) In accounting decision making, accounting decision making is the core part of enterprise management, involving investment, financing, cost control and other aspects. Big Language models can use historical data and market trends for financial forecasting and decision support. The model can analyse market changes, competition and economic trends to provide companies with accurate forecasting results and decision-making advice. This provides business decision-

makers with more comprehensive and accurate information to help them make informed decisions and improve the competitiveness and profitability of their organisations.

However, the application of LLMs in accounting faces a number of challenges and limitations. First, data security and privacy is an important consideration. Big language models require a large amount of data for training and application, but they also need to protect data security and privacy. When applying big language models, it is necessary to comply with relevant laws and regulations and to take appropriate data protection measures. Second, the accuracy and interpretability of the model is also a key issue. Despite the advantages of big language models in dealing with large data and complex problems, the accuracy and interpretability of their results still need further research and improvement. In addition, the application of big language models needs to be integrated with the ethical and moral principles of AI to ensure that its application meets social and moral requirements.

In summary, the application of big language models in accounting has important implications for accounting practice and decision making. By automating the processing of accounting information, providing accurate analysis of financial reports, and supporting risk management and decision making, big language models provide new ideas and tools for accounting. However, there are still some challenges and limitations in the application of big language models that require further research and practice. In the future, with the continuous development of technology and deeper application, the prospect of applying big language models in the field of accounting will be broader.

4 CHALLENGES AND COPING STRATEGIES OF LLMS IN ACCOUNTING

In the accounting domain, LLMs face a number of challenges, including data quality and privacy protection, domain expertise, interpretability, model bias and error, and regulatory and supervisory requirements.

1) Data quality and privacy protection. Accounting data often contains sensitive information, so ensuring data quality and privacy protection is critical. To address this challenge, data desensitisation techniques such as data encryption and anonymisation can be adopted to protect the privacy of data. In addition, data cleansing and validation are necessary to ensure the accuracy and integrity of data.

2) Domain expertise. The accounting domain has complex rules and standards that require in-depth expertise to properly understand and process accounting data. LLMs may not be able to accurately understand and apply this expertise, resulting in inaccurate processing results or non-compliance with accounting standards. To address this issue, the knowledge of accounting experts can be combined and their expertise can be incorporated into the training and validation process of the model to ensure that the model can accurately understand and apply accounting rules and standards. In addition, domain knowledge transfer learning using existing accounting domain knowledge and the model can also improve the effectiveness and accuracy of the model's application in the accounting domain.

3) Interpretability. LLMs are usually black box models and it is difficult to explain their decision making process and results. However, in the accounting domain, interpretability is very important because decisions need to be clearly justified and explained. To address this problem, interpretable machine learning methods such as decision trees and rule extraction can be used to provide explanations and understanding of model decisions. In addition, model interpretability can be enhanced by using model interpretation tools to visualise and explain the model's decision process and results.

4) Model bias and error. Model bias refers to a model's erroneous preference or tendency to misrepresent data, while error refers to uncertainties and errors in the model's predictions or analyses. To address this challenge, models can be evaluated and validated by testing them against real accounting data and comparing them with professional judgement to ensure the accuracy and reliability of the model. In addition, continuous optimisation and updating of the model, taking into account feedback and practical applications, and correction of deviations and errors in the model can also improve the accuracy and stability of the model.

5) Regulatory and supervisory requirements. The accounting field is subject to strict regulations and supervisory requirements, and the large language model must comply with these requirements in the application process. To address this challenge, regulatory and supervisory compliance frameworks and processes can be established to ensure that the application of models meets the relevant regulatory and supervisory requirements. In addition, the establishment of a model audit and monitoring mechanism to monitor and audit the model application process can also ensure model compliance and stability.

5 CONCLUSION

In summary, LLMs have a wide range of applications in the accounting field. It can bring new ideas and tools to accounting by automating the processing of accounting information, providing accurate analyses of financial reports, and supporting risk management and decision making. However, the application of big language modelling in accounting also faces challenges such as data privacy and security, model interpretability, model accuracy and stability. To overcome these challenges, data protection measures, integration of domain expertise, enhancement of model interpretability, correction of model biases and errors, and compliance with relevant regulations and regulatory requirements are required. With the continuous development of technology and in-depth research, the application of LLMs in accounting will have a broader future, providing better tools and support for accountants and driving the digital transformation and development of the accounting industry.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Gao Lei. Explainable Legal Judgment Prediction Driven by Data and Knowledge. Southwest University of Finance and Economics.

[2] Tang Dalu. The Operational Logic and Media Response of Artificial Intelligence Large scale Language Models. China Media Technology. 2023, (02): 19-24.

[3] Zhang Bo, Dong Ruihai. How Natural Language Processing Technology Empowers the AIED:The Perspective of AI Scientist. Journal of East China Normal University (Education Science Edition). 2022, 40 (09): 19-31.

[4] The Era of Artificial Intelligence under Deep Learning. Big Data Era. 2021, (06): 56-76.

[5] Zhou Weihua. The Evolution and Development of the Impact of Information Technology on Accounting Theory and Practice. Friends of Accounting. 2019, (05): 120-124.

[6] Xie Shuangshuang. Analysis of the Impact of Accounting Computerization on Accounting Methods. China Storage and Transportation. 2020, (12): 107-108.