Analysis of the Application of Large Language Models in the Financial Field

Hongyi Qian

SWUFE-UD Institute of Data Science at SWUFE, Southwestern University of Finance and Economics, Chengdu, 611130, China
qianhongyi13@outlook.com

Abstract. In recent years, the Large Language Models (LLMs) have developed continuously, providing new methods for the application of artificial intelligence (AI) in the financial field. This paper introduces the applications and limitations of LLMs and, based on this, proposes prospects for the application of LLMs in the financial field. This article primarily discusses the advantages of LLMs in investment advice, risk management, and processing unstructured data in the financial field. It also mentions that the "black box" nature of LLMs in dealing with regulations still needs further resolution in the future. Furthermore, it proposes that LLMs need to be more professional, transparent, and automated in the future to better adapt to the financial field and promote their deeper application in finance. This article aims to provide a summary of the latest developments in LLMs in the financial field for professionals in related industries, and to offer a reference for them.

Keywords: Large Language Models, Finance, Artificial Intelligence.

1 Introduction

With the continuous development and deepening of artificial intelligence (AI) technology in recent years, the development of Large Language Models (LLMs) has received much attention, especially in the financial field. In the context of globalization, datafication and informatization of the financial market, the application of LLMs in the financial field has received particular attention. Furthermore, in the financial sector, traditional methods are gradually showing their limitations in terms of timeliness and foresight, making it difficult for practitioners to achieve significant progress when using them. However, with the continuous development of LLMs, models like ChatGPT, such as GPT-3, can understand and generate natural language [1]. LLMs can also quickly understand and capture market reactions to government policies and public opinion, and make advanced predictions on stock prices, etc. It can also help practitioners train models at a lower cost and with fewer samples, enabling them to build better models for fields such as trading and risk control. The financial sector has high requirements for real-time performance, accuracy, interpretability, and compliance. LLMs can meet these requirements to a certain extent and are increasingly outperforming traditional methods in some aspects. However, LLMs still have certain

© The Author(s) 2026
L. C. Yoong (ed.), *Proceedings of the 2026 2nd International Conference on Financial Innovation and Marketing Management (FIMM 2026)*, Advances in Economics, Business and Management Research 408,
https://doi.org/10.2991/978-94-6239-801-6_18

shortcomings in terms of reliability and ethics and need to be controlled at the review and regulatory levels.

To date, studies have shown that LLMs have been applied in the five core financial fields of data analysis, investment research, trading, investment management, and risk management, demonstrating the breadth of application of LLMs in the financial field [2]. Meanwhile, professional financial LLMs such as FinGPT and XuanYuan are constantly being developed [3], providing a research foundation for exploring professional financial LLMs. However, LLMs also have problems such as privacy leakage and "black box". Existing articles point out that LLMs can remember personal privacy or private and institutional data, which leads to the leakage of personal or institutional privacy [4]. However, the interpretability required by regulations may further increase the risk of privacy leakage, resulting in the inability to solve privacy issues and interpretability at the same time. Moreover, LLMs may also produce illusions when generating content, leading to incorrect investment advice, etc [5].

Based on the advantages and disadvantages of LLMs mentioned above, this review summarizes and analyzes the applications of existing LLMs in the financial field. This paper summarize the advantages, disadvantage and future development of LLMs. In the part of the advantage of LLMs, it highlights the strengths of LLMs in core application areas such as investment trading and risk management. In the part of the disadvantage of LLMs, it identifies their shortcomings in areas including data privacy, model interpretability, and regulatory compliance. And in the part of the future development of LLMs, it proposes potential research directions and pathways for LLMs within the financial sector, providing a comprehensive reference for practitioners in the intersection of AI and finance to help with subsequent research.

2 Basics of LLMs

LLMs are statistical models that can be trained on large amounts of text and predict using probability distributions [6]. For a word sequence W consisting of $w_1, w_2, \ldots, w_n$ (where wi represents the i-th word in the sequence), its probability $P(W)$ can be calculated using the following probability formula [7].

$$P(W)=P(w_1,w_2,\ldots,w_n)=P(w_1)P(w_2|w_1)P(w_3|w_1,w_2)\cdots P(w_n|w_1,w_2,\ldots,w_{n-1}) \quad (1)$$

LLMs are trained on a large amount of text and have the ability to understand and generate natural language. They are pre-trained neural network models and typically have a parameter scale of hundreds of millions to hundreds of billions [8]. In 2017, after the transformer architecture was introduced, LLMs used the self-attention mechanism to model the associations between words in parallel, thus enabling efficient training on ultra-large-scale datasets and leading to the development of many LLMs [9].

The GPT series is a representative of this, and GPT-3 is a milestone in the development of LLMs[1]. As shown in Table 1, the GPT series of models has many sub-models and a wide coverage. They can demonstrate great capabilities (such as chain thinking, contextual learning, etc.) by expanding the parameter scale and the amount of training data. The evolution of LLMs is mainly due to the improvement of computing power, the availability of large-scale datasets, and the development of new

neural network architectures. These models significantly enhance language understanding and generation capabilities, enabling their widespread application across various industries and fields.

Table 1. GPT-3 series of large language models (GLLMs) from GPT-3 to the latest GPT-4

GPT-3	Ada Babbage Curie Davinci text-ada text-babage text-curie text-davinci
GPT-3.5	code-davinci-002 text-davinci-002 text-davinci-003
ChatGPT	gpt-3.5-turbo gpt-3.5-turbo-16k
GPT-4	gpt-4 gpt-4-32k

3 Application of LLMs In the Financial Field

LLMs have powerful learning and reasoning capabilities, and they have significant advantages in processing unstructured data such as news and media sentiment and cross-market correlation analysis[3]. Meanwhile, LLMs can integrate text, quantitative data and other methods to achieve excellent performance in risk control and fraud detection, which is better than traditional methods[10]. Moreover, LLMs can automate complex tasks such as text (e.g., financial reports, news), charts, predictions, and question answering, extract actionable trading signals from multimodal heterogeneous data, reduce human workload, and provide assistance to financial analysts and other professionals [11,12].

From a data perspective, as shown in Figure 1, the LLM achieves a profit direction prediction accuracy of 64.35% when reasoning based solely on purely numerical financial statements, which is higher than the 58.37% accuracy of human analysts. Moreover, the complex model ranks highly in all three data points, which indicates that it has achieved a balance in the three data points and is more suitable for handling tasks[13].

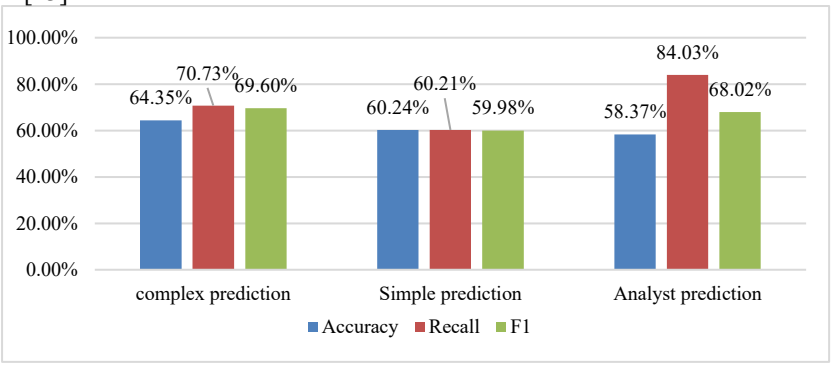


Fig. 1. Accuracy, recall, and F1 score of large language models (complex and simple models) compared to human analysts' predictions.

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (2)$$

From the perspective of different LLMs, according to the data in Table 2 and Table 3, after comparing GPT-4, Claude, and Gemini, it was found that Claude scored the highest in terms of analytical depth and reasoning, while Gemini performed well in terms of information coverage breadth, citing 2.3 times more unique sources than other models [14]. Horizontal comparisons can provide a reference for selecting different LLMs to complete different tasks in the financial field.

Table 2. Comparison of the performance of different large language models in terms of information diversity, output length, and information density

Model Name	Information Diversity Score	Avg. Output Character Count	Information Density
GPT 4.1	37.20	5106.30	7.29
Claude Sonnet 4	79.25	4628.90	17.12
Gemini 2.5 Pro Preview	366.08	6318.10	57.94

Table 3. Comparison of scores of different models on multi-dimensional evaluation indicators

Model Name	Relevance	Accuracy	Reasoning	Problem Solving	Industry Relevance	Human Opinion
GPT 4.1	6.53	6.66	6.59	6.44	6.75	6.75
Claude Sonnet 4	7.44	8.25	8.06	8.13	8.19	8.06
Gemini 2.5 Pro Preview	6.00	6.31	6.25	4.88	6.00	5.44

4 The Shortcomings of LLMs In the Financial Field

The current problem of the LLM is that the path of decision-making is not transparent. When the LLM gives investment advice or risk judgment, it is often difficult to give a complete reasoning chain, but it only gives human practitioners an answer. This will often lead to obvious obstacles in subsequent requirements, such as the determination of responsibility, audit, and review of compliance[15]. LLMs may also appear when facing highly specialized financial issues, such as actively fabricating non-existent data, terms, or fictitious sources. This may lead human practitioners to mistakenly believe in

the LLM after using it for assistance, resulting in incorrect investment advice, reasonable violations, or customer losses, leading to serious problems like this[5].

After training LLMs using a general corpus, they may encounter problems such as insufficient understanding of financial terminology, regulatory rules, and regulations, which leads to a decline in performance[16]. Meanwhile, when using sensitive data containing private or institutional information during the training of LLMs, the model may inadvertently leak training content, leading to risks such as privacy leaks [4]. Based on the above two points, a new problem arises: high-quality financial corpora cannot be used to train LLMs due to privacy issues, while using ordinary corpora will lead to a decline in the capabilities of LLMs. This results in a scarcity of high-quality financial corpora when training LLMs. In addition, publicly available data is often insufficient to cover industry diversity and has historical biases, which will amplify historical errors after training, thus affecting the assessment of the reliability of LLMs [15,16].

5 The Future Development of LLMs In the Financial Field

Based on the above-mentioned application and shortcomings of LLMs in the financial field, this paper believes that LLMs need to promote the development of professional financial LLMs in the future. Dedicated LLMs need to be designed specifically for financial models from the underlying architecture, and financial basic models (FFMs) should be built [17]. Professional terms, relevant regulations and other content should be deeply trained in the model so that LLMs can play a better role in the financial field. Meanwhile, it is necessary to strengthen the reliability and interpretability of the model, and also to develop tools that can conduct compliance audits on LLMs. In this way, the model can not only generate results, but also produce a compliant, complete and reliable reasoning chain, so that regulatory authorities and the company's internal risk assessment team can conduct compliance checks [17]. With the continuous development of technology, LLMs need to promote further automated integration, such as expanding the application of financial models in real-world scenarios and further deepening the automation of model business, making the help of LLMs more readily available to relevant practitioners. Related reports indicate that industry research and practical predictions suggest that the financial industry will accelerate the integration of LLMs into core business processes within the next two years[18].

6 Conclusion

This paper researches the application of LLMs in the financial field. It introduces the concept of LLMs, their advantages and disadvantages in the financial field, and their future development. This paper reviews the clear advantages of LLMs in processing unstructured data, and points out LLMs need imporve in areas such as opaque pathways and difficulty in regulation. It also discusses the future development direction based on the advantages and disadvantages of LLMs.

In summary, the application of LLMs in the financial field breaks through the inherent pattern of traditional models, providing a new and efficient method for financial practitioners. Through its excellent corpus understanding and multimodal data processing capabilities, LLMs demonstrate the potential to exceed the capabilities of human analysts. However, the "black box" nature of the decision-making path within LLMs makes them difficult to regulate, and the results they provide may not reliable. Therefore, it is important to shift from general to specialized development, and to improve the LLMs by building underlying logic specific to the financial field, making it more suitable for the needs of the financial field.

This paper believes that LLMs will become an necessary tool in the financial field in the future, driving the financial field to further progress and achieve better development.

References

1. Kalyan, K.S.: A survey of GPT-3 family large language models including ChatGPT and GPT-4. arXiv:2310.12321 (2023)
2. Dong, Y., Wu, F., Zhang, K., Dai, Y., Zhang, S., Ye, W., Chen, S., Cheng, Z.Q.: Large language model agents in finance: a survey bridging research, practice, and real-world deployment. In: Findings of the Association for Computational Linguistics: EMNLP 2025, pp. 17889-17907. Association for Computational Linguistics, Suzhou (2025)
3. Wu, R., Liu, H.: A survey on the application and research progress of large language models in financial forecasting. AIP Advances 15(6), 060704 (2025).
4. Dhaini, M., Meisenbacher, S., Erdogan, E., Matthes, F., Kasneci, G.: When explainability meets privacy: an investigation at the intersection of post-hoc explainability and differential privacy in the context of natural language processing. In: Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, vol. 8, pp. 731-744 (2025).
5. Mahmud, M.A., Alam, M.A., Alam, M.K.: Exploring the transformative potential of generative AI and large language models (LLMs) in financial applications: opportunities, risks and strategic implications. Journal of Computer Science and Technology Studies 7(8), 1069-1088 (2025).
6. Bengio, Y., Ducharme, R., Vincent, P.: A neural probabilistic language model. Advances in Neural Information Processing Systems 13 (pp. 932–938) (2000)
7. Li, Y., Wang, S., Ding, H., Chen, H.: Large language models in finance: a survey. arXiv:2311.10723 (2023)
8. Minaee, S., Mikolov, T., Nikzad, N., Chenaghlu, M., Socher, R., Amatriain, X., Gao, J.: Large language models: a survey. arXiv:2402.06196 (2024)
9. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. In Advances in Neural Information Processing Systems 30 (pp. 6000-6010). Curran Associates, Inc.
10. Trachova, D., Trachov, V., Demchuk, O., Karpenko, A., Prykhodko, I.: Applications of large language models in the innovative analysis of economic and accounting documentation: a systematic review of recent advances and practical implications. In: Bahaaeddin, A., Hamdan, A. (eds.) The Digital Blueprint: Crafting Business Models for the Tech Future. ICBT 2025. Lecture Notes in Networks and Systems, vol. 1576, pp. 1-15. Springer, Cham (2026).

11. Khan, A.T., Li, S., Cao, X.: Bridging finance and AI: a comprehensive survey of large language models in financial system. Digital Finance 7, 679-701 (2025).
12. Jadhav, A., Mirza, V.: Large language models in equity markets: applications, techniques, and insights. Frontiers in Artificial Intelligence 8, 1608365 (2025).
13. Yin, M., Guo, M.: Complex forecasting and investment strategy optimization via chain-of-thought of large language models. Expert Systems with Applications 298(Part D), 129913 (2026).
14. Bermudez, A., Boeye, M.: Evaluation of modern large language models as tools for financial due diligence: Gemini 2.5 Pro Preview, GPT 4.1, & Claude Sonnet 4. ToltIQ (2025-07-21). https://www.toltiq.com/insights/evaluation-of-modern-large-language-models-as-tools-for-financial-due-diligence-gemini-2-5-pro-preview-gpt-4-1-claude-sonnet-4
15. Bender, E.M., Gebru, T., McMillan-Major, A., Shmitchell, S.: On the dangers of stochastic parrots: can language models be too big? In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21), pp. 610-623 (2021).
16. Guo, Z., et al.: FinEval: a Chinese financial domain knowledge evaluation benchmark for large language models. arXiv:2308.09975 (2023)
17. Chen, L., Liu, S., Yan, J., Wang, X., Liu, H., Li, C., Jiao, K., Ying, J., Liu, Y., Yang, Q., Li, X.: Advancing financial engineering with foundation models: progress, applications, and challenges. Engineering (2025).
18. Dennehy, F.: Large language models could revolutionise finance sector within two years. The Alan Turing Institute (2024-03-27). https://www.turing.ac.uk/news/large-language-models-could-revolutionise-finance-sector-within-two-years

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

