ITM Web of Conferences **84**, 02009 (2026) https://doi.org/10.1051/itmconf/20268402009
ATCIDS 2026

Stock Price Prediction Based on Large Language Models and Reinforcement Learning

Xuanyu Li*

Math Department, University of Washington, Seattle, Washington, United States

Abstract. Stock prediction has always been a key and difficult point of research in the financial field. For a long time, prediction models mostly relied on structured data such as historical market data and technical indicators, making it difficult to fully utilize the rich textual information contained in news, financial reports, and social media, and also difficult to adapt to the dynamic changes of the market. In recent years, the rapid development of artificial intelligence technology, especially the rise of large language models (LLMs) and reinforcement learning (RL), has provided new ideas for stock prediction. This paper systematically reviews the application status of large language models and reinforcement learning in stock prediction and explores the potential of their combination. The article first introduces the basic concepts of large language models and reinforcement learning respectively, then analyzes the advantages of large language models in semantic understanding, context modeling, interpretable prediction, and factor mining, then summarizes the advantages of reinforcement learning in sequential decision-making, return optimization, and risk control, and finally summarizes typical frameworks and synergistic effects of their integration. The purpose of this work is to offer theoretical references and useful advice for developing financial analysis systems that are more intelligent, comprehensible, and self-adaptive.

1 Introduction

As the foundation of the contemporary financial system, the stock market is subject to price changes caused by a variety of factors including market sentiment, corporate fundamentals, macroeconomics, and unforeseen events. Traditional quantitative models mainly rely on historical prices and structured indicators for prediction. Although they have achieved certain results, they have obvious deficiencies in processing massive unstructured text information and simulating the dynamic decision-making behaviors of market participants. With the continuous development of artificial intelligence, large language models and reinforcement learning have provided new ideas and methods for stock prediction. The large language model excels at understanding and analyzing textual information, capturing valuable information from various types of text and generating explanatory predictions; Reinforcement learning focuses more on long-term performance by simulating the process of continuous decision-making, while also taking into account risk factors. At present, there

* Corresponding author's email: xuanyl2@uw.edu

is still a lack of systematic sorting and comparative research on these two types of technologies. Therefore, this article will review the latest developments of large language models and reinforcement learning in stock prediction from three aspects: method principles, practical applications, and fusion methods. It is hoped that their respective advantages, current problems, and future development directions can be clarified, providing some references for promoting the deeper application of artificial intelligence in the financial field.

2 Basic Information

2.1 Large language models

Large Language Models are AI models based on Transformer as the core architecture and obtained through large-scale Internet text training. This type of model can deeply understand the inherent laws of human language and generate smooth and natural text. Its key ability lies in learning the complex relationships between vocabulary, grammar, and semantics from training data. In recent years, large language models have not only performed outstandingly in general language tasks, but have also gradually been applied in multiple professional fields such as finance, healthcare, and law. In the field of finance, typical applications include tasks such as financial news analysis, market sentiment judgment, causal inference of events, and trend prediction. Their advantage lies in their ability to process unstructured textual data and output logical and interpretable content. Specifically, large language models mainly support stock prediction in the following ways: First, they can achieve deep text understanding and context modeling. For example, some studies use models like Mistral-7B combined with historical news information to predict market impact [1]; some large language models can also generate interpretable predictions and support self-reflection, such as the SEP framework which uses models like GPT-3.5 or Vicuna for information summarization, reasoning prediction, and learning from mistakes to improve [2]; furthermore, large language models can directly extract factors affecting stock prices from text, for example, by using prompt engineering to guide models like GPT-3.5-turbo and GPT-4 to identify key factors and company relationships [3]. With the growing demand, specialized large models for the financial field continue to emerge, such as BloombergGPT [4] and FinBERT, focused on sentiment analysis. As large language models are trained on more high-quality financial data, their grasp of professional terminology and industry logic becomes more accurate. All this indicates that, with their powerful understanding and reasoning capabilities, large language models have become an indispensable part of building the next generation of intelligent financial analysis systems.

2.2 Reinforcement learning

Reinforcement learning is a category of machine learning methods centered on interactive learning. The agent continuously interacts with the environment, selects actions based on the current state, and the environment changes accordingly and provides reward feedback. Its goal is to learn a policy that maximizes long-term cumulative reward. In stock prediction and trading, this process is usually modeled as a Markov Decision Process: the state includes information such as stock prices, technical indicators, and position status; actions are buy, sell, or hold; rewards are associated with objectives such as trading profit and risk control. At the specific algorithm level, various reinforcement learning methods have been successfully applied in the field of stock prediction. For example, value-based algorithms like Deep Q-Network, by training a Q-function to evaluate the long-term value of states, were used for initial learning of trading strategies in the DAIS system [5]. Policy optimization algorithms like Proximal Policy Optimization ensure training stability by constraining policy update step sizes and have been used as the core framework of the CGPPO method [6] and applied as one of the key algorithms in IFF-DRL [7]. Furthermore, there are distributed algorithms, such as the quantum-enhanced version of Asynchronous Advantage Actor-Critic, which have been verified to effectively improve trading performance in the foreign exchange market [8]. In addition, other algorithms such as Advantage Actor Critic and Double DQN have also been incorporated into frameworks such as IFF-DRL as important components, enriching the selection of methods for strategy learning.

3 Stock Prediction Based on Large Language Models

Compared to traditional quantitative models that rely on historical prices and structured indicators, stock prediction methods based on large language models have outstanding advantages in deep semantic understanding and logical reasoning ability for unstructured text. The financial market is, in fact, information-intensive, and financial market news, financial reports and social media views, which are text formats, play a decisive role in influencing stock prices. The large language model can handle this raw text directly and extract semantic information and create event logic and assess market sentiment using it, giving a better understanding of how market changing behaviour works, increasing their predictive foresight.

In terms of information understanding and processing depth, large language models have achieved a leap from "keyword matching" to "semantic understanding". Traditional methods such as keyword extraction or sentiment analysis, often can only provide rough and context-lacking information summaries. Models trained with high-quality financial corpora, such as BloombergGPT, can accurately understand professional terms, distinguish complex relationships among companies, such as competition and supply chain cooperation, and infer the differentiated impact of events on different market entities [4]. This deep understanding is an important foundation for generating high-quality predictions and explainable reasons.

The value of financial information often depends on the historical background. However, the Prefix Summary Context method proposed by Koval et al. effectively solves the problem of long context modeling [1], giving the model unique advantages in context modeling and information fusion. This approach greatly improves the model's capacity to assess the "novelty" and "importance" of information by using a small model to condense and summarize historical news and using it as prefix information to guide the large model in understanding current news. Experiments demonstrate that this approach outperforms baseline models that solely concentrate on current data in anticipating news market impact, with a Sharpe ratio of 1.14 in portfolio simulation as historical context length increases [1].

Prediction interpretability has significantly improved due to the SEP framework created by Koa et al. The "self-reflection" mechanism and the "Summarize-Explain-Predict" approach allow the model to organize data, make predictions, and reason like an analyst. It may self-analyze the reasons why the forecast is incorrect and create plans for improvement [2]. In addition to producing better natural language explanations (as assessed by GPT-4), this method steadily increases prediction accuracy by learning from errors, exceeding several conventional models with an accuracy rate of 51.38% on the test set [2].

Additionally, a new factor mining approach has been encouraged by large language models. The LLMFactor framework proposed by Wang et al. is no longer limited to predefined factor libraries. Instead, it uses carefully designed prompts to guide large language models to dynamically extract "conceptual factors" affecting stock prices directly from news texts [3]. Ablation experiments show that these factors generated by large language models contribute significantly to performance improvement, accounting for about a 46% improvement in the MCC metric, demonstrating their potential in building transparent, interpretable AI investment systems [3].

Of course, prediction methods based on large language models still face challenges such as unstable output, hallucination problems, and reliance on high-quality textual data [2, 3, 4]. However, this technology can not only deeply analyze complex financial information but also clearly explain its own reasoning process. At the same time, it is good at handling multi-step complex tasks, so it is expected to become the core of the next generation of intelligent financial analysis tools.

4 Stock Prediction Based on Reinforcement Learning

When traditional methods analyze complex markets, they often oversimplify the problem - they treat consecutive trading decisions as a single issue of predicting the next stock price. In this way, models often place too much emphasis on the accuracy of predicted prices, while neglecting that the transaction itself is a continuous and multi-step decision-making process. Moreover, such core investment objectives as risk management and return optimization can also hardly be put into the construction of traditional models. The inherent distinction of reinforcement learning and classical thought is that the former believes transactions as a sequence of choices. This framework maximizes long-term returns not only, but also naturally incorporates a number of market constraints.

Reinforcement learning breaks away from the limitations of traditional predictive models when setting learning objectives. Traditional models usually only aim for the smallest possible error in predicting figures, but accurate predictions do not necessarily mean making money in actual transactions. In contrast, reinforcement learning skips this step. It does not demand the accuracy of the intermediate prediction stage but directly takes the final performance of the investment portfolio (such as total return and risk-adjusted return) as the optimization direction. This point has been confirmed by many studies. For instance, the DAIS system developed by Park et al. significantly increased the Sharpe ratio of the model from -3.07 in the traditional method to 1.09 by integrating reinforcement learning and supervised learning, with a very obvious improvement in performance [5]. Another IFF-DRL framework proposed by Zhou et al. achieved a cumulative return of up to 1,462.42% and a Sharpe ratio of 4.59 on the Hang Seng Index [7]. These empirical results further confirm that optimization with direct benefits as the objective is a feasible path.

Structural benefits of reinforcement learning in the incorporation of risk control mechanisms are also present. Reinforcement learning can naturally enable the practical constraints to be built into the learning model, like transaction costs and position limits, whereas traditional models can find it hard to introduce these key elements into the learning procedure in a systematic way. To explain, the CGPPO algorithm developed by Chung et al. takes into account simultaneously profitability, volatility, and transaction costs in the context of this scoring system, in effect balancing prospective returns and risks [6]. The last model is QLSTM-QA3C, created by Chen et al., which regulates a maximum drawdown to only 0.92% by a structure that consists of a composite reward system in which time penalties and position penalties are created, entirely showing the superior risk management functions of this model [8].

With the time-dependent nature of market data, the more realistic method of analysis is provided through reinforcement learning. The correlation between stock data itself is quite high, whereas traditional models tend to assume that the data are unrelated to one another, which is not always the case with the real market. Reinforcement learning with deep learning offers a good solution to the problem of acquiring such a temporal dependency. As an example, Zhou et al. integrated a dynamic mechanism of learning into their IFF-DRL model, allowing the model to adapt itself dynamically according to new data and constantly adapt to the new developments in the market, thus effectively solving the issue of the old models that are likely to forget the old but learn the new [7]. The LSTM and graph neural networks employed by Chung et al. prove to have a better environmental adaptability, as the CGPPO approach makes use of both LSTM and graph neural networks in capturing the transient and geographical characteristics of the market on the fly [6].

Even though reinforcement learning is a highly promising tool in stock prediction, the available literature also revealed certain limitations, which exactly indicate the directions that can be further investigated in the future. To begin with, the efficiency and speed of the model computation and response are a practical problem. Some studies have mentioned that training a model for a single stock takes about 10 minutes [5], which, to some extent, limits its application in real-time trading with high-frequency updates. Another study also pointed out that when complex structures such as graph neural networks are introduced, the computational burden will significantly increase, leading to a slower training process [6]. Second, the process of creating models still involves a great deal of subjectivity and dependence. [5] admits that the 22 technical indicators used were chosen subjectively and were not optimized for selection, whereas the studies in [6] and [8] rely on correlations or trends in historical data. When market structure undergoes sudden changes, its assumptions may no longer hold, leading to performance degradation. Finally, the generalization ability of models is generally insufficient, which seriously affects the breadth of their practical application. The models in [5] and [6] were only validated in the Korean stock market, and [7] also only used six major stock indices, without testing in broader asset classes (such as bonds, cryptocurrencies) or cross-market environments; [8] only targeted the USD/TWD single currency pair and did not consider transaction costs, all of which leave the model's effectiveness to be further proven.

5 Stock Prediction Based on the Combination of Large Language Models and Reinforcement Learning

Large language models and reinforcement learning have different methodological focuses, but they show obvious complementarity in stock prediction tasks. Large language models are good at extracting semantic information from unstructured text, performing logical reasoning, and generating interpretable outputs; reinforcement learning is good at sequential decision-making in dynamic environments, optimizing long-term returns, and controlling risks. Combining the two can build an intelligent investment system with perception, reasoning, and decision-making capabilities, which has become a hot research direction.

Existing research demonstrates two typical integration paths. One is the deep integration mode represented by the SCRL-LG framework proposed by Ding et al. [9]. This framework uses large language models as information perceivers to extract global semantic features from news and uses PPO-based reinforcement learning as an optimizer to dynamically learn how to optimally fuse these semantic features with traditional numerical stock features. The other is the division-of-labor collaboration mode adopted by the adaptive margin trading framework proposed by Gu et al. [10]. In this mode, the large language model (such as GPT-4o, Claude-3.5, etc.) first analyzes macro and micro information, outputting interpretable judgments on market trends, and then reinforcement learning (A2C) executes specific fund allocation and trading decisions based on this judgment.

Despite the different specific architectures, these combination methods all demonstrate significant advantages. They have performed a more flexible task of information fusion: large language models can comprehend the multifaceted content of news and financial reports, and reinforcement learning is the one that has to combine all these in-depth comprehensions with the dynamics of the current market, trading restrictions and other constraints, which has achieved more than the traditional models that merely assembled information. In addition, the hybrid system makes the overall decision-making process more transparent and easier to justify. Huge language models are capable of offering a calculative reasoning that humans can internalize, whereas reinforcement learning is constantly adaptive and optimizing, by making changes using feedback information. The two complement each other, making the decision-making thinking of the model no longer like a "black box", and making it easier for people to understand and trust.

In terms of enhancing model performance, the combination of large language models and reinforcement learning naturally unifies the two matters of long-term benefits and risk control. Reinforcement learning, through its reward mechanism, can directly optimize core investment goals such as the Sharpe ratio and cumulative returns, and at the same time, it can naturally take into account practical constraints such as transaction costs and drawdown control. Large language models, with their reasoning capabilities, capture the medium - and long-term, continuous driving factors in the market. Experimental data fully support this advantage. Whether it is the 28.8% annualized return and 1.24 Sharpe ratio achieved by the SCRL-LG framework in the A-share market [9], or the 243.852% cumulative return reached by the adaptive framework in margin trading [10], they are all significantly superior to A single method or traditional benchmarks.

In conclusion, the combination of large language models and reinforcement learning not only technically achieves a closed loop of "perception - reasoning - decision-making", enhancing the system's information processing and adaptive capabilities, but also improves the profitability, robustness and interpretability of investment strategies in practical applications. It has opened up new ideas for breaking through the bottlenecks of traditional quantitative models in the in-depth utilization of information and the flexibility of decision-making, and demonstrated a brand-new direction for the application of artificial intelligence in the financial field.

6 Conclusion

This article reviews the application progress of large language models and reinforcement learning in stock prediction, with a focus on discussing the advantages of their combination. It has been shown that large language models have exceptional semantic understanding and logical reasoning skills to extract information about large quantities of unstructured text, and produce interpretable analytical outputs. This helps not only to make the process of prediction more transparent, but at the same time leads to an increased efficiency of the information use. On the other hand, reinforcement learning, through the framework of sequential decision-making, transforms the simple problem of stock price prediction into the strategic optimization of long-term returns, thereby naturally introducing the ability to control risks and adapt to market changes. Combining the two can build a complete system from "perceiving" information to "reasoning" analysis and ultimately to "decision-making". Multiple experiments have shown that such systems generally perform better than traditional models or single methods. Of course, this field still faces some challenges at present, such as large language models may generate inaccurate content, the training process of reinforcement learning is not stable enough, and there is a strong reliance on high-quality data, etc. Future research can focus on the lightweight deployment of models, the efficient integration of multimodal information, and enhancing the generalization ability of models in different markets, thereby promoting the practical application of this technology in real-world scenarios.

References

1. R. Koval, N. Andrews, X. Yan, Context-aware language models for forecasting market impact from sequences of financial news. arXiv:2509.12519 (2025)
2. K.J. Koa, Y. Ma, R. Ng, T.-S. Chua, Learning to generate explainable stock predictions using self-reflective large language models. in Proceedings of the ACM Web Conference 2024 (WWW '24), 4304–4315 (2024, May)
3. M. Wang, K. Izumi, H. Sakaji, LLMFactor: Extracting profitable factors through prompts for explainable stock movement prediction. in Proceedings of the 2024 Conf.

on Empirical Methods in Natural Language Processing (EMNLP 2024), Miami, USA, 1–12 (2024)

4. S. Wu, O. Irsoy, S. Lu, V. Dabravolski, M. Dredze, S. Gehrmann, ... G. Mann, BloombergGPT: A large language model for finance. arXiv:2303.17564 (2023)
5. M. Park, J. Kim, D. Enke, A novel trading system for the stock market using Deep Q-Network action and instance selection. Expert Syst. Appl. 257, 125043 (2024)
6. J. Chung, M. Kim, S. Min, H. Choi, S. Park, J. Kim, Correlation-assisted spatio-temporal reinforcement learning for stock revenue maximization. Expert Syst. Appl. 289, 128361 (2025)
7. C. Zhou, Y. Huang, Y. Kong, X. Lu, Enhancing trading strategies by combining incremental reinforcement learning and self-supervised prediction. Expert Syst. Appl. 289, 128297 (2025)
8. J.-H. Chen, Y.-C. Huang, Y.-C. Tsai, S.Y.-C. Chen, Quantum-enhanced forecasting for deep reinforcement learning in algorithmic trading. (2025)
9. Y. Ding, S. Jia, T. Ma, B. Mao, X. Zhou, L. Li, D. Han, Integrating stock features and global information via large language models for enhanced stock return prediction. (2023)
10. J. Gu, J. Ye, G. Wang, W. Yin, Adaptive and explainable margin trading via large language models on portfolio management. in Proceedings of the 5th ACM Int. Conf. on AI in Finance, 248–256 (2024, November)