IJCRT.ORG ISSN : 2320-2882
INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

A Systematic Literature Review: Stock Market Price Prediction Using Reinforcement Learning

Shantanu Dipak Rajmane
Department of Information Technology
Army Institute of Technology
Pune, Maharashtra- 411015

Dr. Rupali Bagate
Department of Information Technology
Army Institute of Technology
Pune, Maharashtra- 411015

Dr Aparna Joshi
Department of Computer Engineering
Pimpri Chinchwad College of Engineering
Pune, Maharashtra 411044

ABSTRACT
This paper presents a systematic literature review of reinforcement learning (RL) techniques, particularly deep reinforcement learning (DRL), applied to stock market price prediction. Through comprehensive analysis of recent research, we find that advanced algorithms such as Q-Learning, Double DQN, and Dueling DQN, especially when combined with sentiment analysis from news and social media, create powerful frameworks that address financial markets' complexity. Our review indicates that DRL approaches significantly outperform traditional methods in the literature, resulting in more adaptive and dynamic solutions for stock market forecasting. This paper synthesizes current research findings and identifies promising directions for future work in this rapidly evolving field.

Keywords
deep reinforcement learning, stock market prediction, Q-learning, Deep Q-Network, financial forecasting, systematic review, sentiment analysis

1. INTRODUCTION

Being a complex, dynamic system, the stock market is influenced by a huge number of factors, among which are economic indicators, political events, and investors' sentiment. Forecasting stock prices has always been a challenge for economists and financial analysts, who have primarily been using statistical models of linear regression and time series analysis. These methods offer a general view of market trends; however, they often discard a lot of the complex, non-linear behaviors that are typical in financial markets. Such challenges to normality with financial data suggest that more advanced research using approaches such as machine learning and artificial intelligence would be needed for better prediction accuracy.

Recently, reinforcement learning—a subfield of machine learning—has garnered a lot of attention because of its promise to enable financial applications.

Where traditional ML models rely on data with labels, RL learns the best sequence of decisions through interaction with a system or environment to maximize cumulative rewards.It is particularly useful in stock market prediction because one has to not only resort on predicting prices but be able to design the most profitable trading strategy under changing conditions of markets. Deep Reinforcement Learning (DRL) which provides the power of Deep Learning with RL also one such tool used now a days in this context. Since models trained with DRL are associated to process massive data and can learn intricate patterns, they seem tailored for the high-dimensional noisy financial dataset. We trained various neural networks in both one-step Q-learning and double DQN schemes for dynamic trading strategy selection depending on the state of financial markets.

Incorporation of sentiment analysis extracted from social media and news sources enhances the view of market dynamics by capturing both historical and real-time data. In this paper, the efficiency of DRL in the prediction and development of stock market strategy is evaluated. This review will include recent developments and methodologies to make a note of the salient points and limitations of DRL approaches, which should identify some research avenues for future work. Our work focuses on how DRL can predict not only the direction of market trends but also how it can provide actionable insights to the investor, which is one of the focal points for furthering financial analytics.

2. Literature Survey

The stock market price forecast is an evolution from the traditional statistical methods of time series analysis and linear regression to more modern methods using machine learning. Although these can provide a basis for understanding, traditional methods often proved inefficient in capturing the non-linear and volatile characteristics of financial markets, as illustrated by Pesaran and Timmermann (1995). One can turn to reinforcement learning, by using Q-learning advancement as a possible approach for modeling the stock markets itself with respect of being able to adapt in new or changing market regimes. Furthermore, sentiment analysis has been employed to capture public opinion via social media and other news sources which not only includes historical data but also enhances predictive efficiency through real-time information (Wang \& Wang, 2020). This complete methodology not only overcomes the drawbacks of prior frameworks but also provides a strong standard way for more informed stock market predictions [1].

There are quite a few ways in which research has been adapted to predict stock prices with reinforcement learning. The most popular were linear regression and time series analysis, but they failed at being used effectively due to some degree of uncertainty and volatility in financial markets. Recent approaches make use of machine learning techniques, particularly Reinforcement Learning, which shows promise in this area. In particular, Deep Reinforcement Learning (DRL) came out as a powerful tool, combining the feature extraction of deep learning with the decision-making power of RL. The philosophy of the hybrid model is to process high-dimensional data and forecast policy results for accurate prediction making. These include using Q-Learning, Double DQN and Dueling DQN on arbitrary trading environments. In addition, models have included sentiment analysis of social media and news feeds to deliver non-stop data that also adds an element of prediction efficiency by helping capture public opinion. It was also intended to serve as a more complete framework than previous models, not just one which allows arbitrage, but an actualy dynamic-system approach for handling the unknown and statistical dynamics at play in financial markets [2].

The paper "Stock Trading Strategies Based on Deep Reinforcement Learning" by Yawei Li, Peipei Liu, and Ze Wang describes a deep reinforcement learning (DRL) model that is specifically designed to learn and adapt in real time to cope with the market dynamics of stock trading. The study underscores the inadequacies of static trading strategies and challenges related to noisy financial data. For better market state analysis, the authors integrate multiple sources of data, including stock data, technical indicators, and candlestick charts. This is achieved using LSTM networks and CNN-BiLSTM for feature extraction. Sharpe Ratio (SR) and profit rate (PR) were used to balance risk and returns in the new reward function. Experiments on the Chinese stock market datasets, S\&P 500 datasets showed that our model significantly (with a p-value of <0.01) beat other models in profit rate and annualized return as well as Sharpe Ratio with its higher efficiency and generalization ability. This work improves algorithmic trading culture by enabling the development of dynamic and adaptive strategies, which are able to optimally manage risks whilst maintaining profitability [3].

Xing Wu et al. in the paper "Adaptive Stock Trading Strategies with Deep Reinforcement Learning Methods" address the challenges of designing robust trading strategies in complex and dynamic stock markets. It proposes two adaptive methodologies: GDQN (Gated Deep Q-learning Network) and GDPG (Gated Deterministic Policy Gradient), used to extract informative features from raw financial data and technical indicators, improving the accuracy and robustness of stock market representation. To make predictions on such an ever-changing stock market, the previous work attempts to leverage deep reinforcement learning in applications of non-stationary domains (non-metrics spaces) that learn strategies keeping as main goal return maximization. Experimental results on U.S., U.K. and Chinese stocks demonstrate that compared to the traditional Turtle trading strategy, both GDQN and GDPG are able to significantly outperform state-of-the-art DRL strategies with more stable returns in trending markets as well volatile ones after transaction costs retained. Our results show that, in the actor-critic method with concrete optimization algorithm GDQN, GDPG can perform more stably and efficiently than critic-only (GDQN) however this suggests a future direction — Deep reinforcement learning grandient recurrent unit is both powerful for adaptive quantitative trading strategies[4].

A paper titled, A Survey of Forex and Stock Price Prediction Using Deep Learning by Zexin Hu, Yiqi Zhao, Matloob Khushi on how forex prices have been predicted using deep learning techniques It categorizes various deep learning techniques like Convolutional Neural Networks (CNN), Long Short-Term Memory(LSTM), Deep Neural Networks(DNN), Recurrent Neural networks(RNN) and Reinforcement Learning, including hybrid models such as Hybrid Attention Network(HAN) or Wavenet. The survey has arranged them in terms of dataset, variable and models used along with the performance matrix such as RMSE, MAPE,MAE MSE Accuracy Shapre ratio Return rate etc. The recent models LSTM with techniques like DNN have shown a lot of promise. Similarly, other advanced deep learning methods like reinforcement learning have also been developed showing high returns and performance benefits. The paper evidence that deep learning is gaining momentum in financial modeling, as it excels at capturing complex and nonlinear relationships between input features for a given historical dataset, which do create better results (accuracy) when comes to the prediction of stock market price [5].

In the research of Feizi-Derakhshi, Lotfimanesh and Amani (2024), they introduce a pricing prediction model for Iranian gold market which is based on Double Deep Q-Network (DDQN) algorithm using CNN \& LSTM layers. This algorithm takes advantage of the feature extraction ability of CNN and long-term dependency learning capacity in LSTM, thus overcomes drawbacks exposed by price prediction problems with abrupt oscillations. The experiment found that the model had better performance in projecting price fluctuations with high accuracy and reliability compared to other methods when six different types of gold from 2009–2020 are evaluated. Results of the study Results showed that DDQN -LSTM-CNN proposed method reduced prediction errors overarcking their baselines, further confirming it as a robust tool for investros and market analysts to rely on in challenging markets. This is the first work to combine coupled LSTM-CNN neural network models in DDQN for financial forecasts. The results demonstrate its superior performance compared to other commonly used models [6].

Ding, Chen, Jiang, Lin, and Lin (2024) develop a reinforcement learning framework to advance stock portfolio management using information on financial news sentiment associated with price trends. More generally, while extracting sentiment scores from financial news using a large language model, TrendTrader creates heterogeneous contexts that could benefit the stability of multimodal feature processing. Such an approach helps to reduce the problems of feature disparity and fragile convergence that are common with RL methods. A spatial-temporal backbone network combining supervised and reinforcement learning with incremental training to achieve better convergence of policy networks is also used in TrendTrader. Comprehensive experimental validation over both US and China markets shows the superiority of TrendTrader compared with a series of state-of-the-art methods, presenting better annualized returns, lower volatility, and improved risk management. The key role of embedding news sentiment and trend heuristics in RL frameworks for robust and profitable portfolio management is highlighted [7].

This paper by Li, Zhu, Chen, and Hu (2024) studies a stock trading strategy that uses deep reinforcement learning (DRL) to use capital position data from Hong Kong to enhance trading performance in the volatile Chinese A-share market. The study formulates a trading simulation environment along with the PPO algorithm to train the trading agent in handling some key problems like environmental unpredictability and data heterogeneity. The agent incorporates technical indicators and the factors extracted from Hong Kong capital flows, such as position changes, capital flow, profit-loss ratio, and excess inflows/outflows. The backtesting results show the DRL-based agent can outperform the baseline methods to achieve higher annualized returns, which can also manage risk better, especially when the market is volatile and other critical events such as COVID-19 hit it. This research put an emphasis on reducing market noise by taking in capital data from Hong Kong for the robustness of trading strategies and indicated the possibility of future research by incorporating sources that are much more diverse in financial information [8].

In the DRL optimization portfolio literature, the effectiveness of directly outputting investment actions rather than requiring some unrealistic prediction about future asset prices shows. The thesis aimed to make use of DRL, in particular the PPO algorithm, and deal with the limitations through the introduction of variants like GTrXL and imitation learning techniques. These were model stability and data efficiency approaches, essential because of the relative scarcity of financial data. Particularly, the GTrXL model could manage sequential information across long horizons to better embed past prices and indicators. Finally, imitation learning was done using methods such as Behavioral Cloning and Offline Reinforcement Learning for the constraint of data with pretrained model on expert actions. However, the study also found that decision stability was an issue and technical indicators alone were not sufficient to make stable decisions. This branch of work emphasizes that while DRL shows promise for financial applications, additional pursuit is required to create consistent and accompany reliable investment strategies [9].

In this regard, the "NIFTY Financial News Headlines" dataset has been released with the aim of enabling and promoting research in the field of financial market forecasting using large language models. The data includes two mutually independent forms: NIFTY-LM for supervised fine-tuning using an auto-regressive causal language modeling objective for LLMs and NIFTY-RL, which is highly adapted to be used for reinforcement learning alignment methods. The dataset used contains deduplicated financial news headlines of high quality, coupled with comprehensive metadata and market indices. Each headline has been systematically filtered and ranked for relevance. This structured dataset is aimed at overcoming the challenges of financial forecasting, i.e., market volatility and scarcity of data; thus, it will help to avail a rich corpus that will boost the predictive power through LLMs. Experiments on the dataset have achieved promising results for tasks like stock price movement forecasting, which indicates the effectiveness of large-scale language models in financial forecasting applications [10].

Additionally, greater quality of data analysis in finance is described by LLM embeddings. The research demonstrated that models that are larger in size are more informative in producing embeddings that add up to more accurate information acquisition and better data clustering related to the market.This result further highlights the fact that "the rich get richer" in LLMs – bigger model will simply be able to pick up more subtle aspects of the data. The paper also suggests a way forward involves building on similar reinforcement learning techniques such as RL from human feedback (RLHF) and then extending it to large language models in order for them to better generalize market movement prediction efficiently. As a publicly hosted dataset on the Hugging Face repository, this corpus not only serves as foundation for financial market forecasting applications but also enables further computational research via systematic data truncation \& handling facilities and will benefit to financial data science community [10].

3. Research Findings and Analysis

Based on our systematic review of the literature, several key observations emerge regarding reinforcement learning applications in stock market prediction:

3.1 Advantages of DRL in Financial Forecasting

Our analysis reveals that DRL methodologies consistently demonstrate superior capabilities in processing complex financial data compared to traditional approaches. The reviewed studies show that DRL frameworks can effectively:

- Process multi-dimensional market data including price histories, technical indicators, and sentiment signals
- Adapt to changing market conditions in ways static models cannot
- Generate actionable trading strategies rather than just price predictions
- Balance risk and reward considerations when properly configured

3.2 Prominent Algorithmic Approaches

The literature shows particular success with several DRL variants:

- Q-Learning and its extensions (Double DQN, Dueling DQN)
- Policy Gradient methods (especially PPO)
- Hybrid approaches incorporating LSTM and CNN architectures for feature extraction

3.3 Integration of Alternative Data Sources

A significant trend in recent research is the integration of non-traditional data:

- Sentiment analysis from financial news and social media
- Cross-market information flows (e.g., Hong Kong capital position data for Chinese markets)
- Technical indicators combined with fundamental data

3.4 Current Limitations and Challenges

Despite promising results, several challenges persist:

- Model stability across different market regimes
- Model stability across different market regimes
- Interpretability of DRL-based trading decisions
- Data efficiency given the relative scarcity of financial data
- Potential for overfitting to historical patterns

4. Conclusion

Deep Reinforcement Learning: DRL is instrumental in forecasting the stock market which went far away from conventional methods, indeed it brings a more advanced and dynamic view on how markets work. This work highlights that DRL methods are able to digest a wide range of complex financial data like historical prices, technical indicators and sentiment signals obtained from social media or news. To that extent: they are able to make faster and better decisions during volatile markets, this is achieved by having more comprehensive systems which cover multiple data types.

The research underscores the potential for using DRL to make effective and tactical predictions of financial markets. Using the powerful ADQN algorithm (utilizing techniques like Q-Learning, Double DQN and Dueling with a hint of RNN to always overcome different trading environments) — outperforms traditional models in all scenarios. Integrating advanced neural networks, such as Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN), enable DRL systems to model temporal dependencies more effectively and improve feature extraction across big data sets.

Nevertheless, difficulties persist on having the model stability and the interpretability of DRL strategies. Finally, we still need more work to redesign models and acquire data from a variety of sources in order for us to fully make sense on how DRL systems decide. In addition, we need to investigate the moral questions raised by something as advanced and powerful as AI being used in financial markets (for example regarding market manipulation etc.)

Although the results indicate that DRL is a powerful tool for stock market prediction and trading, further exploration remains to be done. These innovations offer an era of opportunity on financial analysis transformation, aiding investors and analysts to master today's complex markets in the expanding landscape with new AI and ML based technologies iteratively developed.

5. Acknowledgements

The authors are thankful to the faculty and staff of Department of Information Technology, Army Institute of technology, Pune for their valuable advices. The authors would like to thank the faculty members who participated in this work also for anonymous reviewers whose commentary truly helped us improve the quality of our manuscript. This work was funded by the Army Institute of Technology. We would also like to thank the Army Institute of Technology for allowing us to use their computational resources which made our survey extremely effective.

6. REFERENCES

[1] M. H. Pesaran and A. Timmermann, "Predictability of stock returns: Robustness and economic significance," The Journal of Finance, vol. 50, no. 4, pp. 1201-1228, 1995.

[2] Y. Li, P. Ni, and V. Chang, "Application of deep reinforcement learning in stock trading strategies and stock forecasting," Computing, vol. 102, no. 7, pp. 1305–1322, 2020.

[3] Y. Li, P. Liu, and Z. Wang, "Stock trading strategies based on deep reinforcement learning," Scientific Programming, vol. 2022, Art. no. 4698656, pp. 1–15, 2022.

[4] X. Wu, H. Chen, J. Wang, L. Troiano, V. Loia, and H. Fujita, "Adaptive stock trading strategies with deep reinforcement learning methods," Information Sciences, vol. 538, pp. 142–158, 2020.

[5] Z. Hu, Y. Zhao, and M. Khushi, "A survey of forex and stock price prediction using deep learning," Applied System Innovation, vol. 4, no. 1, p. 9, 2021.

[6] M.-R. Feizi-Derakhshi, B. Lotfimanesh, and O. Amani, "Accurate price prediction by double deep q-network," Inteligencia Artificial, vol. 27, no. 74, pp. 12-21, 2024.

[7] W. Ding, Z. Chen, H. Jiang, Y. Lin, and F. Lin, "Trend-heuristic reinforcement learning framework for news-oriented stock portfolio management," ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5120-5124, 2024.

[8] J. Li, C. Zhu, X. Chen, and X. Hu, "A stock trading strategy based on deep reinforcement learning and hong-kong capital position," International Journal of Innovative Computing, Information and Control, vol. 20, no. 2, pp. 643-652, 2024.

[9] G. Scaletta, "Deep Reinforcement Learning for Portfolio Optimization," Master's Thesis, Politecnico di Torino, 2024.

[10] R. Saqur, K. Kato, N. Vinden, and F. Rudzicz, "NIFTY Financial News Headlines Dataset: A Comprehensive Resource for Financial Market Forecasting," 2024.