Proceedings of CONF-CDS 2026 Symposium: Computer Vision-Based Multimodal Cognitive Load Estimation for Adaptive Media Communication
DOI: 10.54254/2755-2721/2026.GL35291

Machine Learning for Stock Market Prediction: A Review of Sentiment, Technical, Macroeconomic, and Fundamental Factors

Tianyuan Shi

Department of Computer Science, Stonybrook University, Stonybrook, USA
Tianyuan.Shi@stonybrook.edu

Abstract. Predicting the stock market has become an important research topic in recent years, because accurate result can support investment decision and reduce financial risks.Traditional statistical models are hard to analyze in the nonlinear market, leading researchers to develop machine learning models. This paper reviews recent studies about stock prediction based on machine-learning methods from four perspectives: traditional machine learning models, sentiment analysis, technical indicators, macroeconomic and fundamental factors. The reviewed literature shows that machine learning methods can have better result than traditional models in handling complex financial data. Furthermore, sentiment information, technical indicators, and macroeconomic or fundamental factors can significantly improve prediction performance. However, each approach has limitations. The review finds that using multiple factors may relate to accurate predictions. Therefore, combining different factors through hybrid models is considered the most effective strategy. Finally, this review reveals a clear transition from single-factor prediction models to multi-factor prediction systems and provides a comprehensive understanding of current developments in stock prediction and highlights potential directions for further research.

Keywords: Machine Learning, Stock Prediction, Sentiment Analysis, Technical Indicators, Macroeconomic and Fundamental Factors

1. Introduction

The stock market affected modern financial systems significantly. People can invest and companies to finance by stock market directly. Accurate stock prediction is valuable and important for investors, financial institutions, and policymakers because it can improve investment decision-making, reduce financial risks, and support market stability. However, predicting stock market movements is challenging. The stock market is dynamic and nonlinear. Moreover, the stock market is associated with many factors. Some main factors are historical market prices, investors' moods, company financial conditions, and macroeconomic environments.

Traditional stock prediction used mathematical methods such as linear regression and ARIMA models. Because of the nonlinear and complicated characteristics of the stock market, their

DOI: 10.54254/2755-2721/2026.GL35291

performances of them are inaccurate. In order to have a better prediction, many researchers put their minds to machine learning models.

Different from traditional methods, machine-learning algorithms could automatically learn complex patterns from large datasets. Early studies have tried several simple machine learning models. The most common and well-known ones are Artificial Neural Networks(ANN) and Support Vector Machines(SVM).

After trying some more machine learning algorithms, researchers have found that some other factors can also affect the stock prices. Later studies then added specific information sources such as investor sentiment, technical indicators, macroeconomic conditions, and company fundamental data. These factors may help models better understand market behavior.

This paper reviews machine-learning-based stock prediction research from four aspects: traditional machine learning models, sentiment-based prediction approaches, technical-indicators methods, and models with macroeconomic and fundamental factors. The aim of this paper is to provide a comprehensive overview of current research developments and identify important directions for future studies.

2. Machine learning-based stock prediction

In the past, many stock prediction studies used statistical models such as linear regression and ARIMA models. The result is inaccurate because the market can not be predicted in a simple, linear and mathematical model. Because of the complexity of the stock market, the development of new models is complicated but necessary. Recently, researchers have turned their attention to machine learning. Then it started becoming popular in financial forecasting because it could adapt to a strange and complicated environment automatically.

There are many machine learning models for stock prediction. The earliest and classic models are ANN and SVM. They have different advantages. ANN models perform better at learning nonlinear relationships from large datasets. SVM models are good at working with classification problems and high-dimensional data. Instead of depending on a fixed mathematical model, machine learning models can find unique features from historical data and fix the problems themselves.

Because of the different characteristics of ANN and SVM, their performance on the same data is different. In order to develop a better model for the stock market, research was conducted by Yasin Kara et al. [1]. After predicting the direction of movement in the daily Istanbul Stock Exchange (ISE) National 100 Index by ANN and SVM, they find the average performance of ANN models was 75.74%, and SVM models' was 71.52% [1]. The research showed that ANN models were better than SVM models when they predicted the stock market with the same data.

In order to understand and predict the stock market better. Later studies explored other machine learning methods. Ballings compared multiple machine-learning classifiers for stock price direction prediction. Their results showed that "Random Forest achieved the best overall performance among several other machine learning models". Their study suggested that ensemble learning methods should be considered more when stock markets are predicted [2]. The stock market is unstable and noisy. Using multiple models can reduce errors and overfitting. Different models have different advantages. They could provide more stable prediction performance when changing conditions of the market.

Processing data before using any machine-learning models also helps improve the performance of machine-learning models. According to Patel's research, "the performance of all the prediction models improves when these technical parameters are represented as trend deterministic data" [3]. In

DOI: 10.54254/2755-2721/2026.GL35291

other words, focusing on data pre-processing, like changing the exact numbers of data to trend, can improve the accuracy of the prediction.

However, although many authors have tried several models and methods, basic machine learning-based stock prediction still has limitations. The three authors relied on historical data, which means their models mainly focused on historical prices, volume and index movement in the market; they didn't take outside factors such as personal sentiments and technical indicators into consideration. To make improvements, later studies started considering machine learning with different factors to understand the stock market better.

3. Machine learning with sentiment factors

Besides the historical data of the stock market, emotions of investors also strongly affect the stock market. Emotions are usually considered as sentiment factors. Investors will be affected by news, including financial types and political events. According to Bollen's study, public sentiment can be used to predict the stock market. Their results show that when certain mood dimensions are included, the prediction accuracy of stock market is improved [4].

After noticing the importance of sentiment, some researchers began to evaluate public emotions and news as data and put them into machine learning models. From Gu's research, the team tried three models and methods to predict financial stock market prices. They compared Fin-BERT Embedded LSTM Architecture, LSTM Architecture, and DNN Architecture [5]. The author developed the Fin-BERT Embedding LSTM Architecture, which analyzed news and combined them with historical data. The other two architectures only use historical data. As a result, the Embedding LSTM Architecture method performed the best with the least computation time [5]. Gu's research showed that adding sentiment analysis helps machine learning models come closer to reality.

Moreover, some attention should be paid. Davidovic and McCleary pointed out the relationship between the stock market and sentiment. Sentiment would be influenced by macroeconomic news and change gradually over a period. News did not affect the market instantaneously but unfolded in phases. The most important warning was that sentiment alone had weak power. In other words, "sentiment-only regression models confirm virtually no predictive power of any sentiment metrics" [5]. Other factors, such as market volatility and liquidity, still played important roles in prediction.

Some more limitations exist when using sentiment-based models. Online news is not reliable all the time. Fake news and misleading information can not reflect the real market actions. Processing and evaluating emotional data requires a huge amount of resources.

Although there are many limitations, sentiment factors are regarded as an important direction and source for stock market prediction. To figure out these limitations, some researchers tried to combine other technical indicators, such as volatility and liquidity, with machine learning models to improve performance.

4. Machine learning with technical indicators

Technical indicators come from historical data. They are new features mathematically calculated from original data. Machine learning models can use them as patterns that can be learned. Some common technical indicators are moving averages, RSI (Relative Strength Index), and MACD (Moving Average Convergence Divergence). Different technical indicators represent different features of market behavior, such as trend, momentum and volatility. Compared with original historical prices, using indicators can provide more financial patterns for machine-learning models to learn the stock market [6].

DOI: 10.54254/2755-2721/2026.GL35291

In Mostafavi's and Hooman's study, they focused on identifying the most useful technical indicators for stock market prediction. Their research analyzed 88 technical indicators using several machine learning models, including Random Forest, XGBoost, SVR, and LSTM. Their results show the most effective technical indicators are trend, momentum, volatility, and volume categories [7].

Surprisingly, technical indicators can be designed and adjusted automatically. According to Chong's research, without using previous predictors, the author and his team used a deep learning network that can learn features from original data [8]. However, the study also warned that deep learning models required lots of data and resources.

On the other side, Wang studied stock return prediction by using neural network models with multiple financial measures. The author emphasized that different types of indicators contributed to different indicators. Moreover, the author suggested that combining different, distinct features and variables can improve the performance of prediction [9].

Although technical indicators and deep learning methods improved stock prediction performance, several limitations still existed. Also in Wang's research, financial features may affect each other. For example, similar trading signals of a company may have different meanings in different periods. In addition, different market conditions may affect the correctness of technical indicators [9].

Even with these limitations, using technical indicators is still one of the most important tools in machine-learning-based stock prediction models. Most of the limitations contain an important concept: market conditions or macroeconomic factors. Therefore, many researchers tried to discover and analyze them to get better results of prediction.

5. Machine learning with macroeconomic and fundamental factors

Macroeconomics is one of the most important concepts in normal markets. In stock predictions, macroeconomics has a high status too. Many researchers pointed out that stock market behavior is affected under different economic conditions.

In Wang's research, machine learning models had poor performances under recessions [10]. The author found that ML model performances have a close connection with market volatility [10]. The more stable a stock market is, the better a machine learning model can predict. The author recommends that other ML practitioners should evaluate their ML models during both recession and non-recession periods because of the different conditions [10].

Later studies focused more directly on macroeconomic variables. Macroeconomic factors refer to economic conditions that affect the overall financial market, such as inflation rates, interest rates and economic uncertainty. According to Zeng's research, the author found that macroeconomics had a great impact on machine learning performances. Then Zeng "developed a hybrid model that combines both random forest and the LASSO" [11]. To a certain extent, the new model can reduce the impact caused by market volatility.

Besides macroeconomic conditions, some researchers also noticed fundamental factors. Fundamental factors focused on individual companies usually include company earnings, revenues and debts. A research study by Huang studied stock prediction based on fundamental analysis using machine learning models. Their research suggested that company financial information and fundamental variables could improve prediction performance and support investment decision-making. Compared with prediction systems that only used historical prices, combining machine learning models with company financial data helped explain long-term stock market behavior more effectively [12].

However, macroeconomic and fundamental-factor-based prediction systems still have several limitations. Macroeconomic data is usually updated more slowly than stock market data, which

DOI: 10.54254/2755-2721/2026.GL35291

makes short-term prediction more difficult. Moreover, in Huang's research, fundamental analysis also requires large amounts of financial data, and some company information may not fully reflect future market conditions [12].

Even with these limitations, macroeconomic and fundamental factors have become increasingly important in machine-learning-based stock prediction research. Many recent studies have suggested that historical market data, sentiment factors, technical indicators, and macroeconomic information should be used together. In this way, prediction performance can be improved and help machine learning systems understand complicated financial stock markets to the maximum extent.

6. Analysis and discussion

The studies mentioned above show the development of stock pridiction by machine-learning models. Early researches mainly focused on improving prediction accuracy through traditional machine learning models such as ANN and SVM. These models showed the differences between traditional mathematical methods and machine learning models. Machine learning models perform better because they can learn nonlinear relationships as patterns from data. ANN performs better than SVM.

Later, researchers recognized that historical market data alone could not fully let the model understand the stock market. Therefore, additional information sources began to be added to machine learning models. Sentiment analysis considered investor emotions and market psychology into prediction systems; Technical indicators provided mathematical features representing market momentum, trends and volatility; Macroeconomic and fundamental factors offered broader economic and company information that may improve prediction.

Despite these improvements, several challenges remain. Sentiment information may contain noise and misinformation, technical indicators may behave differently under changing market conditions, and macroeconomic data is often updated less frequently than market data. Furthermore, many advanced machine learning and deep learning models require large datasets.

A trend is observed through this literature. Recent studies showed a transition from single-factor prediction models to multi-factor prediction systems. Relying on a single source of information is not comprehensive. Different factors can help the model catch different features of the same market behavior. In order to process multiple features, researchers used hybrid models that integrate multiple data sources, which can achieve better prediction accuracy and stability.

Overall, the literature suggests that future research should pay attention to developing hybrid and ensemble models that are able to integrate diverse information sources on the stock market. Such approaches may further improve the reliability and practical application of machine-learning-based stock prediction systems.

7. Conclusion

This paper reviewed recent developments in machine-learning-based stock prediction from four perspectives: traditional machine learning models, sentiment analysis, technical indicators, and macroeconomic and fundamental factors.

The studies of Machine learning methods proved that machine learning models can improve stock prediction performance. Research has evolved from using historical market data alone to combining different aspects of information sources such as investor sentiment, technical indicators, macroeconomic and fundamental variables. These developments have enabled prediction models to fit the stock market better with higher accuracy.

DOI: 10.54254/2755-2721/2026.GL35291

An important finding of this review is that relying on a single factor still has limitations. Different factors such as historical data, sentiment analysis, technical indicators, and macroeconomic factors, each provide unique patterns and features. Therefore, combining multiple sources of information may be considered the most effective approach for improving prediction accuracy and robustness.

In conclusion, machine learning should continue to be used in stock market forecasting. Future studies should focus on building more adaptive and robust prediction systems that can analyze different features from different sources of information. Discovering more factors that can affect the stock markets and Hybrid models with large, open and transparent datasets contribute to improvements and can help investors, enterprises and policymakers make their decisions.

References

[1] Kara, Y., Boyacioglu, M. A., & Baykan, Ö. K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange. Expert Systems with Applications, 38(5), 5311–5319.
[2] Ballings, M., Van Den Poel, D., Hespeels, N., et al. (2015). Evaluating multiple classifiers for stock price direction prediction. Expert Systems with Applications, 42(20), 7046–7056.
[3] Patel, J., Shah, S., Thakkar, P., et al. (2015). Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. Expert Systems with Applications, 42(1), 259–268.
[4] Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. Journal of Computational Science, 2(1), 1–8.
[5] Gu, W., Zhong, Y., Li, S., et al. (2024). Predicting stock prices with FinBERT-LSTM: Integrating news sentiment analysis. In Proceedings of the 2024 8th International Conference on Cloud and Big Data Computing (pp. 67–72).
[6] Davidovic, M., & McCleary, J. (2025). News sentiment and stock market dynamics: A machine learning investigation. Journal of Risk and Financial Management, 18(8), 412.
[7] Mostafavi, S. M., & Hooman, A. R. (2025). Key technical indicators for stock market prediction. Machine Learning with Applications, 20, 100631.
[8] Chong, E., Han, C., & Park, F. C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. Expert Systems with Applications, 83, 187–205.
[9] Wang, C. (2024). Stock return prediction with multiple measures using neural network models. Financial Innovation, 10(1), 72.
[10] Wang, L. R., Fu, H., & Fan, X. (2023). Stock price predictability and the business cycle via machine learning. arXiv preprint arXiv: 2304.09937.
[11] Zeng, Q., Lu, X., Xu, J., et al. (2024). Macro-driven stock market volatility prediction: Insights from a new hybrid machine learning approach. International Review of Financial Analysis, 96, 103711.
[12] Huang, Y., Capretz, L. F., & Ho, D. (2021). Machine learning for stock prediction based on fundamental analysis. In 2021 IEEE Symposium Series on Computational Intelligence (SSCI) (pp. 01–10). IEEE.