The Study on the Application of Machine Learning Algorithms for Stock Prices Prediction During Special Periods

Wei Li

Finance, Dongbei University of Finance and Economics, Dalian, 116025, China
Li-W12@ulster.ac.uk

Abstract. Stocks are a favored investment channel, with their fluctuations closely linked to macroeconomic conditions. In recent years, global events such as COVID-19, the Russo-Ukrainian war, and the Israeli-Palestinian conflict have caused significant volatility in international stock markets. This paper investigates the application of machine learning algorithms for stock price prediction during these turbulent times. Models such as Random Forest, Support Vector Machine, Linear Regression, Convolutional Neural Networks, Artificial Neural Networks, and Long Short-Term Memory Networks are frequently employed during such periods. Many of these models integrate policy and news indicators that reflect social changes and investor sentiment, helping improve prediction accuracy in times of high uncertainty. Additionally, some models introduce methods for optimizing hyperparameters to enhance forecasting performance further. However, existing machine learning models face challenges such as low interpretability, limited applicability, and high sensitivity to external factors. These issues often lead to reduced investor confidence, increased training costs, and inconsistent results across different stocks and special periods. Looking forward, the integration of new approaches, such as expert systems, alongside traditional machine learning methods, could help mitigate these challenges and improve prediction outcomes in future periods of economic uncertainty.

Keywords: Machine Leaning, Deep Learning, Stock Price Prediction.

1 Introduction

Stocks have been in development for over 400 years and are an important way for corporations to acquire funding and reduce investment risk for investors. Investors are interested in choosing promising companies to invest in based on their growth expectations. In the short term, a company's share price is determined by investor sentiment; in the medium term, it is determined by the country's fiscal and monetary policies, and in the long term, it is determined by the company's earnings growth. Therefore, share prices are affected by multiple factors such as investor sentiment, market conditions, and macroeconomic conditions. Both institutional and individual investors in the market hope to be able to predict the price of a stock so that they can

© The Author(s) 2025
J. Lu (ed.), *Proceedings of the International Workshop on Navigating the Digital Business Frontier for Sustainable Financial Innovation (ICDEBA 2024)*, Advances in Economics, Business and Management Research 315,
https://doi.org/10.2991/978-94-6463-652-9_68

buy low and sell high to make a profit. Investors used to predict stocks by analyzing company earnings data, share price movements, and macro conditions. Now with the advancement of artificial intelligence, a growing number of studies are exploring the application of machine learning models and the establishment of multiple models to predict stock values.

Nowadays, numerous machine learning methods are extensively employed for stock price forecasting such as Random Forest (RF), Neural Networks, Long and Short Term Memory (LSTM) Models and Support Vector Machine (SVM) Models and so on. The random forest model was first proposed by Breiman [1] in 2001 and has since been used in stock forecasting. Machine learning algorithms have made great progress in financial time series data forecasting compared with previous statistical measurement models, but their forecasting effect has to be further improved. Many scholars have respectively conducted optimisation studies on forecasting models from the levels of algorithm improvement, model fusion, feature selection and so on. In addition, with the development of neural networks particularly in the domains of textual analysis and computer vision, many scholars have tried to introduce neural network models into the forecasting research in the field of finance and achieved better results. Sellvin [2] uses three distinct models: Convolutional Neural Network (CNN), LSTM, and Recurrent Neural Network (RNN), to predict the short-term trends of equities in Indian stock exchanges. Patel et al. [3] use a two-stage fusion approach when predicting stock prices using machine learning models. The initial phase employs the SVR model, whereas the subsequent phase utilizes ANN, RF, and SVR. It was found that the accuracy of the results predicted using the two-stage fusion approach was generally higher than the prediction accuracy using the single-stage machine learning model. The SVR-ANN model had the highest prediction accuracy.

There are many well-established and more accurate models for predicting stock prices using machine learning models. However, most of these models are currently based on data obtained in stable stock markets for research. Since 2020 the occurrence of special events such as COVID-19 and geopolitical conflicts have exerted a significant influence on the stock market. Predicting stock market trends from past data is challenging owing to the chaotic environment and significant volatility [4]. Machine learning models of stock price predictions are more difficult during periods when special events occur. This paper discusses the capability of machine learning models to predict stock price movements during periods of high uncertainty, such as geopolitical conflicts or pandemics, and examines the challenges associated with their current use. The focus is on whether these models can maintain their predictive accuracy in such special periods compared to times of market stability.

2 Methods

2.1 Traditional Machine Learning Models

Random Forest. RF is a basic approach in machine learning when investors seek to forecast the stock price. In stock prediction, researchers always use random forests by constructing multiple decision tree models and then using the votes or averages of these decision tree models to predict outcomes. For instance, Rouf, Malik and Arif [5] first used Pearson Correlation to study the correlation between the overall amount and each day's number of confirmed cases and the Nifty50 index of India during COVID-19. It was concluded that market volatility is positively correlated with the number of cases. Predicted the stock price using many machine learning models, among which the RF approach predicted with the highest accuracy.

Support Vector Machine. SVM is a machine learning algorithm frequently employed to address classification and regression issues. In forecasting for the stock market, SVM models usually learn patterns in historical data to make predictions. Niu et al. [6] use SHapley Additive exPlanations (SHAP) to identify the types of geopolitical risk that have the greatest impact on realised volatility forecasts. Subsequently, it used several machine learning models to forecast volatility, with SVR forecasts having the highest accuracy.

Linear Regression. LR is a basic supervised learning method. The LR model in forecasting stock prices mostly relies on the analysis of historical data and making predictions by constructing a function to predict the class of new data points. When using the machine learning model for forecasting, Adán Reyes Santiago and González [7] added two indicators Williams %R and the 13-week T-bills into the model for forecasting. The uncertain COVID-19 period was divided into pre-COVID-19, COVID-19 high volatility period and COVID-19 low volatility period. The prediction found that linear regression has the highest prediction accuracy in the COVID-19 low volatility period. Also, they found that using any of the machine learning models will have a higher yield compared to passive investing strategies in uncertain periods.

2.2 Deep Learning Models

Convolutional Neural Network. CNN model is a deep learning model mostly utilized in activities such as image recognition and object detection. The process of stock prediction usually uses CNN to extract features from financial data. In predicting stock prices during periods of uncertainty, Zouaghia et al. [8] combine CNN and Halving Grid Search (HGS) algorithm to get a new method called AutoCNN. This model first used a one-dimensional CNN model to automatically extract features and then used the filtered features into the CNN model again to predict stock price. If after n cycles the model does not satisfy the requirement of error acceptance, the HGS model will be used to get new hyperparameters and selected features to re-enter the CNN to predict the stock price. It will get more accurate results. This model predicting the SSE index and NASDAQ Composite

index price during COVID-19 in 2020 and the recent geopolitical conflict between Russia and Ukraine in 2022 both had more accurate outcomes than others.

Artificial Neural Network. ANN is a deep learning model that emulates neurons in the human brain. Every neuron in the model calculates a linear amalgamation of inputs and weights before a non-linear transformation of the activation function to produce an output. The model has better nonlinear processing suitable for periods of uncertainty. ANN model predicts stock price mainly by analysing the historical data to identify the factors of stock price movement to forecast the future trajectory of stock prices. Napitupulu et al. [9] used the Government Stringency Index (GSI) and GT value from Google as predictors to input into the ANN model. This model includes the people's emotional reactions and the government's response to events in the short term. Predictions can be based on realities that are more closely aligned with the uncertain period and the psychological state of the population.

Jamous et al. [10] introduced a novel ANN-particle swarm optimization with a centre of gravity (ANN-PSOCoG) prediction model for price estimation. Traditional ANN models expend considerable time and effort within the process of manually searching for adjustments to get suitable hyperparameters (including the number of hidden layers (NHL), number of neurons per hidden layer (NNPHL), activation function (ACTFUN) type, and learning rate (LR)). In the ANN-PSOCoG model, to find the optimal hyperparameters, researchers use the centre of gravity (CoG) to enhance the PSO algorithm. This model uses a four-dimensional search space utilized that each hyperparameters have one dimension search space. The PSO model is then used to search for the optimal configurations by manoeuvring the particles in each dimension. A distinct ANN model is developed for each particle and the position and velocity are continuously refined using the fitness function to achieve the minimal fitness value. Reaching the minimum fitness value means that the ANN model has selected the best hyperparameters.

In addition, Goel et al. [11] projected the Indian stock market via an ANN model, incorporating macroeconomic indices and global stock market variables as input factors. Besides, they used the Scaled Conjugate Gradient (SCG) algorithm to train the ANN model, which can avoid line search iteration-wise to save time.

Long Short-Term Memory. LSTM models forecast future stock price movements by capturing patterns and trends in time series data. Sharaf et al. [12] propose a system consisting of stock historical price data analysis and social news sentiment analysis. The model represents the influence of COVID-19 on the stock market by performing sentiment analysis on COVID-19 news and articles. Upon completion of the sentiment analysis and historical data analysis, the data is allocated to a Random Forest classifier, while a Stacked-LSTM classifier is employed in the last phase of the system to forecast future inventory trends.

While introducing investor sentiment in the machine learning model, Chou et al. [13] add an attention mechanism to filter the key information. They increase an attention layer after the LSTM layer and before the output layer. The probability of

information assignment for each time step is derived from the attention layer and then its information is merged into the LSTM layer.

3 Discussion

In recent years, both traditional machine learning and deep learning have been regularly employed for stock price prediction during specific periods. However, the overall trend is that investors are now using methods that are gradually shifting from traditional machine learning to deep learning models for forecasting. In contrast to deep learning models, traditional machine learning is more comprehensible and necessitates fewer data. However, it also has the problem of lower prediction accuracy. In the special period of the stock market, numerous uncertainties prevail, deep learning models can deal with complex non-linear relationships and based on a large amount of data to predict. They are more suitable for use in periods of high uncertainty.

Limitations. Machine learning methods have inherent limitations in predicting stock prices. The first problem faced by machine learning is the problem of interpretability of results. Often the results of stock price predictions will be strongly correlated with the decisions of investors or managers. Traditional machine learning is based on statistics and math and uses human-set features to make predictions that are easy to understand. On the contrary, because of the excessive number of parameters and the complexity of the methods used to extract features deep learning models are often difficult for decision makers to understand. When there is a deviation between the model prediction result and the actual, its overly complex process makes it difficult to trace the cause of the deviation and impossible to improve the model.

The second problem faced by machine learning is the applicability problem. Currently, there are many types of models available in academic circles to forecast stock prices. In special periods of high uncertainty, multiple models are often combined or new methods are introduced to combine with the original model to improve the precision of the forecast. There is a large gap between the responses of research subjects in different regions and industries to the occurrence of special events. Therefore, it is likely that the same model will not be used for other types of research subjects or will not be applicable when special events fluctuate. It is costly to design different models for research subjects with different characteristics.

The third problem of machine learning is the impact of extrinsic variables such as media reports and investor sentiment. The occurrence of special events usually creates nervousness and makes people more cautious in their investment behaviour. Social media will also convey different special emotions to people. These result in the market feeling that most investors have completely different investment strategies and specific plans than in normal times. In machine learning model design, it is important to take into account external information and investor psychology during special times to get more accurate results. At the same time, this will also enhance the difficulty of model design and operation.

Future Prospects. In the future machine learning can solve the current problems in the following ways. For interpretability, expert systems can be introduced in combination with machine learning models to solve the problem. Expert systems can provide specialized knowledge such as market dynamics, economic indicators, and company fundamentals to guide parameter tuning of machine learning models. Introducing an expert system can give the investor an explanation of the error in improving the methodology when the machine learning model makes an error. For example, Hossain et al. build Belief Rule-Based Expert Systems based on predicting a company's pricing history [14], which have higher predictive accuracy than ordinary machine learning models in predicting stock prices five days after a special event. Another approach known as Local Interpretable Model-agnostic Explanations (LIME) can enhance the interpretability of machine learning. LIME selects some of the model's data samples and generates new sample points by perturbing them in their vicinity to obtain the predictions of the original model. The training results obtained by LIME using the new datasets are a good local approximation of the machine learning model. During specific intervals, news and investor sentiment can be transformed into quantifiable variables suitable for integration into machine learning models for predictive purposes. At present, several scholars have used the method of extracting word frequency to quantify social news to add it to the prediction model. In the future, it may be possible to assess the external factors of special periods through other new methods and add them to the model to enhance the predictive accuracy.

4 Conclusion

This paper summarizes the use of machine learning by different scholars to forecast stock prices during periods of high uncertainty. It is found that most of the literature utilizes Random Forest, SVM, LR, CNN, ANN, and LSTM for stock prediction during the special period. To improve precision in prediction, many scholars choose to add new models based on machine learning models to make hyper-parameter predictions or formulate external factors such as government policies and news in special periods into quantifiable indicators to be added into machine learning models for prediction. However, because of the high uncertainty of special periods and the high variability of different special events, machine learning models suffer from the problems of poor interpretability, poor applicability, and high influence of results by external factors in predicting stock prices. These problems may be ameliorated in the future by combining machine learning with expert systems, and LIME models, and incorporating new quantitative external factor indicators.

References

1. Breiman, L.: Random forests. Machine Learning 45, 5-32 (2001).
2. Selvin, S., Vinayakumar, R., Gopalakrishnan, E. A., Menon, V. K., Soman, K. P.: Stock price prediction using LSTM, RNN and CNN-sliding window model. In: 2017

International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp. 1643-1647. IEEE (2017).
3. Patel, J., Shah, S., Thakkar, P., Kotecha, K.: Predicting stock market index using fusion of machine learning techniques. Expert Systems with Applications 42(4), 2162-2172 (2015).
4. Ticknor, J. L.: A Bayesian regularized artificial neural network for stock market forecasting. Expert Systems with Applications 40(14), 5501-5506 (2013).
5. Rouf, N., Malik, M. B., Arif, T.: A machine learning based approach to unleash the impact of COVID-19 on Indian stock market (2020).
6. Niu, Z., Wang, C., Zhang, H.: Forecasting stock market volatility with various geopolitical risks categories: New evidence from machine learning models. International Review of Financial Analysis 89, 102738 (2023).
7. Reyes Santiago, A., González, J.: Machine learning portfolios for US stock prices: Directional forecasting before and during the COVID-19 pandemic. Contaduría y Administración 69(4) (2023). https://doi.org/10.22201/fca.24488410e.2024.5191.
8. Zouaghia, Z., Kodia, Z., Ben Said, L.: A novel AutoCNN model for stock market index prediction. Journal of Telecommunications and the Digital Economy 12(1), 612-636 (2024).
9. Napitupulu, H., Sambas, A., Murniati, A., Kusumaningtyas, V. A.: Artificial neural network-based machine learning approach to stock market prediction model on the Indonesia Stock Exchange during the COVID-19. Engineering Letters 30(3) (2022).
10. Jamous, R., ALRahhal, H., El-Darieby, M.: A new ANN-particle swarm optimization with center of gravity (ANN-PSOCoG) prediction model for the stock market under the effect of COVID-19. Scientific Programming 2021(1), 6656150 (2021).
11. Goel, H., Som, B. K.: Stock market prediction, COVID-19 pandemic and neural networks: An SCG algorithm application. EconomiA 24(1), 134-146 (2023).
12. Sharaf, M., Hemdan, E. E. D., El-Sayed, A., El-Bahnasawy, N. A.: An efficient hybrid stock trend prediction system during COVID-19 pandemic based on stacked-LSTM and news sentiment analysis. Multimedia Tools and Applications 82(16), 23945-23977 (2023).
13. Chou, C., Park, J., Chou, E.: Predicting stock closing price after COVID-19 based on sentiment analysis and LSTM. In: 2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), pp. 2752-2756. IEEE (2021).
14. Hossain, E., Hossain, M. S., Zander, P. O., Andersson, K.: Machine learning with belief rule-based expert systems to predict stock price movements. Expert Systems with Applications 206, 117706 (2022).

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.