Proceedings of ICFTBA 2026 Symposium: Strategic Management & Business Analytics: Data-Driven Performance Models
DOI: 10.54254/2754-1169/2026.LD38255

A Comparative Investigation of Linear and Non-Linear Methods for Stock Price Prediction

Hellen Jiang

Georgia School, Ningbo, China
hellen.jiang@georgiaschool.cn

Abstract. Predicting stock prices is a common topic in finance and data science. Investors and companies are interested in possible future price movements, but making accurate predictions is difficult. Stock prices can be affected by trading activity, company information, economic conditions, investor decisions, news, and unexpected events. This paper reviews several approaches for stock price prediction, with a focus on both linear and non-linear methods. The methods discussed are Linear Regression, ARIMA, Support Vector Regression (SVR), Decision Trees, Random Forest, Artificial Neural Networks (ANN), and Long Short-Term Memory (LSTM). Previous studies are compared based on their datasets, prediction targets, evaluation measures, and main findings. In general, simpler models are easier to understand and use, while non-linear models can deal with more complicated patterns. However, complex models may require more data, computing power, and careful training. The paper also considers problems related to external information, model interpretability, overfitting, and changes in financial markets. Future studies could combine market data with news and sentiment, improve the explainability of AI models, and allow prediction systems to update as market conditions change.

Keywords: Stock Price Prediction, Linear Regression, Machine Learning, Deep Learning, LSTM

1. Introduction

The stock market connects companies that need funding with people who want to invest. Because stock prices can change quickly, investors are interested in estimating what may happen next. However, this is a difficult task. Price movements can be related to previous prices, trading volume, company information, economic conditions, investor behavior, news, and unexpected events. Financial data also contain a large amount of noise, and the relationships between variables may change over time. These characteristics make stock price prediction different from many ordinary prediction problems.

Because of these difficulties, researchers in finance, statistics, and data science have developed many ways to study future stock prices and returns. Some of the earlier approaches use statistical models. Linear Regression is useful when a relatively simple relationship between variables is expected, while ARIMA is designed for time-series patterns. More recently, machine learning and deep learning have received more attention. In a review of 138 studies, Kumbure et al. found that

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2754-1169/2026.LD38255

neural networks, support vector machines, and other machine learning techniques were frequently used in stock market forecasting [1]. Other reviews have also reported growing interest in deep learning, especially methods such as LSTM for financial time-series prediction [2-4].

The development of these methods does not mean that newer models always replace older ones. Researchers have found different results under different conditions. Kara et al., for example, examined Artificial Neural Networks and Support Vector Machines for predicting the direction of the Istanbul Stock Exchange index [5]. Chen et al. applied LSTM to stock returns in the Chinese market and found that it could identify patterns beyond random prediction [6]. Studies comparing ARIMA and LSTM have also reached different conclusions. Li found better results from ARIMA for the stocks in the study, while Zhang reported better performance from LSTM for five major indexes [7, 8]. These differences suggest that the choice of model depends on the data and the prediction task.

This paper compares many commonly used linear and non-linear methods for stock price prediction. The section 2 introduce the main methods and examples from previous studies, section 3 compares their strengths and weaknesses and discusses issues such as external information, explainability, overfitting, and changing market conditions. And for the final part, it considers possible future developments, including the use of different data sources, explainable AI, and adaptive prediction systems.

2. Methods used for stock price prediction

Stock prediction research has gradually expanded from traditional statistical models to machine learning and deep learning. Depending on the study, the inputs may include previous prices, returns, trading volume, technical indicators, and other financial information. Researchers normally use part of the data to train a model and another part to test how well it works on unseen data. Common evaluation measures include Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and R^2. Kumbure et al. examined 138 studies and identified 2,173 variables that had been used in stock market prediction research [1]. This shows that researchers have considered a wide range of information when building prediction models.

2.1. Linear regression

Linear Regression is one of the simpler choices for stock prediction. It estimates a relationship between a target value and one or more input variables. For example, previous prices, returns, trading volume, or technical indicators can be used as inputs.

Chibli et al. tested Linear Regression together with Random Forest and LSTM using data from 29 companies and the S&P 500 index. The data covered the period from January 2000 to June 2024. MSE, RMSE, and R^2 were used to evaluate the models. Random Forest produced better results than Linear Regression, while LSTM gave the strongest results among the three models [9].

This comparison suggest that a simple linear relationship may not describe all the patterns in stock data. However, the Linear Regression is still useful for comparison, because it is fast and relatively easy to explain. It can provide a basic result that more complicated models can be compared against.

DOI: 10.54254/2754-1169/2026.LD38255

2.2. ARIMA

ARIMA, or Autoregressive Integrated Moving Average, is a statistical model made for time-series data. Instead of treating observations as completely separate, it uses information from earlier observations to estimate later ones. Data may also need to be differenced or transformed before the model is used so that the time series is more suitable for analysis.

Li compared ARIMA with LSTM using the closing prices of Pfizer, Alibaba, EA, and NVIDIA [7]. In this study, ARIMA gave better prediction results than LSTM. This is an important result because it shows that a traditional model can still perform well against a more complex deep learning approach.

A different result was reported by Zhang. The study used daily opening, closing, high, and low prices from five major stock indexes from 2015 to 2022. LSTM performed better than ARIMA in that comparison [8]. The difference between these two studies suggests that the performance of a model can change with the market, dataset, and time period.

ARIMA is useful when the main information comes from patterns in a time series. Its weakness is that it may have difficulty dealing with strong non-linear patterns or sudden market movements caused by unexpected events.

2.3. Support Vector Regression

Support Vector Regression (SVR) is a machine learning method for predicting continuous values, so it can be applied to stock prices. One useful feature of SVR is its use of kernel functions. Different kernels allow the model to deal with relationships that are not simply linear.

Henrique et al. examined SVR with both daily and high-frequency stock price data from different markets [10]. Their study included Brazilian stocks as well as companies such as Bank of America, Google, and Exxon Mobil. They looked at both normal daily observations and prices recorded at shorter time intervals. The researchers compared SVR with a random walk model and evaluated the prediction errors using RMSE and MAPE.

Several kernel types were tested, including linear, radial, and polynomial kernels. The study found that SVR could provide useful predictions, especially when the model was updated regularly [10]. This suggests that SVR can be useful when stock prices have more complicated relationships that a basic linear model may not capture.

At the same time, SVR requires decisions about its parameters and kernel. Different choices can produce different results. Its performance may also change as market conditions change.

2.4. Decision Trees and Random Forest

A Decision Tree divides data according to a series of conditions and uses the resulting groups to make predictions. In financial prediction, these conditions can involve variables such as previous prices, returns, trading volume, or technical indicators. A single tree is relatively easy to follow, but it can fit the training data too closely.

Random Forest takes a different approach by combining many decision trees. Each tree can work with different samples and features, and their outputs are combined into a final prediction. This makes the method less dependent on one individual tree and allows it to capture non-linear relationships.

In the study by Chibli et al., Random Forest, Linear Regression, and LSTM were tested using data from 29 companies and the S&P 500 between 2000 and 2024 [9]. MSE, RMSE, and R^2 were

DOI: 10.54254/2754-1169/2026.LD38255

used for the comparison. Random Forest outperformed Linear Regression, while LSTM produced the best overall results among the three models [9].

The result shows that Random Forest can find relationships that may be missed by a basic linear model. It is also able to work with many variables without assuming that they have a linear relationship. However, Random Forest is not specifically designed to remember long sequences of past observations, which is one of the main features of LSTM.

2.5. Artificial Neural Networks

Artificial Neural Networks (ANNs) use connected layers to learn relationships in data. Unlike a simple linear model, an ANN can learn non-linear patterns and can use several financial variables together.

Kara et al. studied ANN and Support Vector Machines for predicting whether the Istanbul Stock Exchange index would move upward or downward [5]. Therefore, the goal was not to predict an exact stock price but to classify the direction of the market. The researchers trained the models using historical market information and compared their classification results. Their work showed that ANN and SVM could be applied to stock market direction prediction [5].

Kumbure et al. also found that neural networks were among the frequently used approaches in the 138 studies included in their review [1]. This indicates that ANN has been an important part of financial prediction research for some time.

The main advantage of ANN is its ability to learn complicated relationships without requiring the researcher to specify the exact form of those relationships. The trade-off is that neural networks generally require more data and computing resources than simpler models. It can also be difficult to explain exactly how the model reaches a prediction.

2.6. Long Short-Term Memory Networks

Long Short-Term Memory (LSTM) is a type of recurrent neural network that is designed for sequential data. This makes it suitable for financial time series because observations from earlier periods can provide information for later predictions.

Chen et al. used an LSTM-based approach to predict stock returns in the Chinese stock market [6]. The model used historical stock information and learned from patterns across different time periods. The study reported that accuracy increased from 14.3% for random prediction to 27.2% with the LSTM model [6]. Although the accuracy was still limited, the result showed that the model could obtain useful information from the stock-return data.

LSTM has also been combined with information outside normal price data. Usmani and Shamsi developed an LSTM model that used categorized and weighted financial news for stock prediction [11]. Their approach therefore included text related to market events and investor expectations instead of depending only on historical prices.

LSTM is attractive for stock prediction because it can deal with both non-linear relationships and the order of observations. However, this flexibility also comes with costs. LSTM models can require large datasets, careful parameter choices, and considerable computing resources. Their internal decision-making process is also harder to explain than that of simpler models.

Taken together, these studies show that stock prediction methods range from simple statistical approaches to more flexible machine learning and deep learning models. Linear Regression and ARIMA are relatively easy to use and understand, while SVR, Random Forest, ANN, and LSTM can represent more complicated patterns. The studies also show that greater model complexity does

DOI: 10.54254/2754-1169/2026.LD38255

not automatically lead to better prediction. The result can depend on the dataset, prediction target, time period, and evaluation method.

3. Discussion

3.1. Comparison between linear and non-linear models

Linear and non-linear models approach financial data in different ways. Linear Regression assumes a relatively direct relationship between the inputs and the predicted value. ARIMA is also a traditional approach, but its main focus is the behavior of observations over time. These methods can be useful when the data have patterns that remain fairly stable.

Models such as Random Forest, SVR, ANN, and LSTM provide more flexibility. They can represent relationships that are difficult to describe with a simple linear equation. However, greater flexibility should not be confused with guaranteed higher accuracy.

The research reviewed in this paper gives a good example. Li reported that ARIMA performed better than LSTM for the stocks examined in that study [7]. Zhang reached the opposite result when comparing the two models using five major indexes [8]. Chen et al. found that LSTM could identify useful patterns in Chinese stock returns [6]. Chibli et al. found that Random Forest performed better than Linear Regression, while LSTM had the strongest results among their three tested models [9].

These findings make it difficult to name one method as the best overall choice. A model may work well for one market or period and perform less well in another. The prediction target and evaluation method also matter. Non-linear models can be useful for complicated data, but they normally require more training and computational resources. Simpler models may not capture every pattern, but they are easier to interpret and can provide a useful baseline. The model should therefore match the research question and the type of data being used.

3.2. Main challenges

Historical prices are not the only things that move financial markets. Company announcements, economic reports, political developments, financial news, and investor sentiment may all influence prices. A model that depends only on past market data may have difficulty responding to a sudden event. Kumbure et al. showed that stock prediction research has used a very large range of variables [1]. Usmani and Shamsi also demonstrated that financial news can be included in an LSTM model [11]. Ashtiani and Raahemi reviewed the use of news, text mining, and alternative information in financial market prediction [12].

However, adding outside information is not automatically helpful. News and social media posts may contain opinions, incomplete information, or contradictory signals. Researchers also need to decide which information is relevant and when it should enter the model. As a result, simply adding more variables may increase complexity without improving prediction.

Another concern is interpretability. Some models make their reasoning relatively easy to follow. For example, the coefficients in Linear Regression can help show the relationship between an input and the prediction. Deep learning models such as LSTM are much harder to inspect because they contain many internal parameters and layers. Černevičienė and Kabašinskas reviewed different explainable AI approaches that can be used in finance [13]. This area is important because a useful prediction is not always enough; users may also want to understand why a model produced it.

Overfitting is another problem. A model can learn details that fit its training data but fail when it receives new observations. This problem can be especially serious with complex models because

DOI: 10.54254/2754-1169/2026.LD38255

financial data contain noise. Market behavior can also change over time. A pattern found during one period may disappear during another period, especially when a major economic event or market crisis occurs. Researchers therefore need careful training and testing procedures and should check model performance over more than one period.

3.3. Future prospects

One possible direction is to combine several types of information. Future models could use prices, returns, volume, and technical indicators together with financial news, company reports, social media, and sentiment. Kumbure et al. showed how many different variables have already been considered in previous research [1]. Usmani and Shamsi provided an example of combining financial news with LSTM [11], while Ashtiani and Raahemi discussed the use of alternative data and sentiment information in financial prediction [12].

Explainable AI is another area that could become more important. A complex model may produce useful predictions, but users still need some way to understand the reasons behind them. Černevičienė and Kabašinskas discussed XAI approaches for financial applications [13]. Abdullah et al. also studied explainable deep learning together with textual information for stock price forecasting [14]. Future systems could use these ideas to show which variables or pieces of text had the greatest influence on a prediction.

Models may also need to become more adaptive. Financial markets are not fixed, so a model trained on older information may lose some of its usefulness as new conditions appear. A prediction system could be updated regularly with recent prices, trading activity, news, and other information. This would allow it to respond more quickly when the market changes.

The goal should not simply be to build a more complicated model. A practical prediction system needs a balance between accuracy, reliability, explainability, and adaptability. Combining different sources of information with explainable and regularly updated models may be a more useful direction than increasing model complexity alone.

4. Conclusion

This review looked at several methods for stock price prediction, including Linear Regression, ARIMA, SVR, Random Forest, ANN, and LSTM. Traditional models are usually simpler and easier to understand, while machine learning and deep learning models can handle more complicated patterns.

The studies in this review do not point to one model as the clear winner. Li found that ARIMA performed better than LSTM in one study, while Zhang reported better results from LSTM in another. Chibli et al. also found that Random Forest performed better than Linear Regression, while LSTM gave the best result among their three models. These differences show that model performance depends on the data, market, prediction target, and time period.

There are also several challenges in stock prediction. Prices can change because of news, economic conditions, investor sentiment, and unexpected events. Complex models may be difficult to explain and can also overfit noisy financial data. In addition, a model that works well on older market data may not perform the same way when conditions change.

Future research can bring together financial data and textual information, improve the explainability of complex models, and update prediction systems as markets change. A more complicated model is not necessarily a better one. A useful prediction model should be accurate enough while also being reliable, understandable, and able to adapt.

DOI: 10.54254/2754-1169/2026.LD38255

References

[1] Kumbure, M. M., Lohrmann, C., Luukka, P., et al. (2022). Machine learning techniques and data for stock market forecasting: A literature review. *Expert Systems with Applications, 197*, 116659.

[2] Chen, W., Hussain, W., Cauteruccio, F., & Zhang, X. (2024). Deep learning for financial time series prediction: A state-of-the-art review of standalone and hybrid models. *Computer Modeling in Engineering & Sciences, 139*, 187-224.

[3] Rouf, N., Malik, M. B., Arif, T., Sharma, S., Singh, S., Aich, S., & Kim, H. C. (2021). Stock market prediction using machine learning techniques: A decade survey on methodologies, recent developments, and future directions. *Electronics, 10*, 2717.

[4] Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005-2019. *Applied Soft Computing, 90*, 106181.

[5] Kara, Y., Boyacioglu, M. A., & Baykan, O. 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*, 5311-5319.

[6] Chen, K., Zhou, Y., & Dai, F. (2015). A LSTM-based method for stock returns prediction: A case study of China stock market. *2015 IEEE International Conference on Big Data (Big Data)*. IEEE.

[7] Li, H. (2024). Comparison of ARIMA and LSTM in different industries. *Advances in Economics, Management and Political Sciences, 57*, 294-302.

[8] Zhang, Z. (2023). Comparison of LSTM and ARIMA in price forecasting: Evidence from five indexes. *2023 2nd International Conference on Economics, Smart Finance and Contemporary Trade (ESFCT 2023)*. Atlantis Press.

[9] Chibli, M., Smeureanu, I., & Haydar, M. (2025). Stock price prediction: A comparative study using linear regression, random forest, and LSTM models. *International Journal of Education, Business and Economics Research, 5*, 293-309.

[10] Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2018). Stock price prediction using support vector regression on daily and up to the minute prices. *The Journal of Finance and Data Science, 4*, 183-201.

[11] Usmani, S., & Shamsi, J. A. (2023). LSTM based stock prediction using weighted and categorized financial news. *PLOS ONE, 18*, e0282234.

[12] Ashtiani, M. N., & Raahemi, B. (2023). News-based intelligent prediction of financial markets using text mining and machine learning: A systematic literature review. *Expert Systems with Applications, 217*, 119509.

[13] Černevičienė, J., & Kabašinskas, A. (2024). Explainable artificial intelligence (XAI) in finance: A systematic literature review. *Artificial Intelligence Review, 57*, 216.

[14] Abdullah, M., Sulong, Z., & Chowdhury, M. A. F. (2024). Explainable deep learning model for stock price forecasting using textual analysis. *Expert Systems with Applications, 249*, 123740.