

Predicting Stock Market Trends: A Comparative Study Of Time Series Forecasting Models Using Historical Stock Prices

Neelanjan Mitra¹, Dr. Anirban Hazra²

¹Assistant Professor, Department of Business Management,
Swami Vivekananda University, Kolkata, West Bengal,
neelanjanmitra.2008@gmail.com

Abstract

The Indian banking industry has undergone a series of structural and strategic transformations over the past few decades, particularly through mergers and acquisitions (M&A). These consolidations have often been implemented as corrective measures to address inefficiencies, bad loans, and financial distress among public sector banks (PSBs). Rooted in the concept of resource optimization, mergers have emerged as a crucial instrument for stabilizing and strengthening India's financial ecosystem. This paper explores the historical evolution, rationale, and impact of banking mergers in India, highlighting both the advantages and challenges associated with them.

The study underscores that mergers can enhance operational efficiency, reduce redundancy, expand market reach, and improve capital adequacy. However, it also recognizes the complexities involved—ranging from valuation ambiguities and stakeholder dissatisfaction to cultural integration and data management issues. An analysis of India's banking merger timeline since 1969 reveals significant consolidation, with institutions like the State Bank of India and Punjab National Bank leading the process. Empirical observations indicate that larger banks have maintained stronger market capitalizations post-merger, reflecting improved stability though not necessarily higher profitability.

The research emphasizes the importance of effective stakeholder communication and strategic governance to ensure that mergers achieve their intended synergies without compromising financial inclusivity or employment stability. By examining both qualitative and quantitative aspects, this paper provides an integrated understanding of how mergers have shaped India's banking landscape, offering insights into their long-term viability and implications for economic development.

Keywords:

Bank Mergers, Public Sector Banks, Financial Consolidation, Stakeholder Perception, Indian Banking System

Introduction:

Interest in the analysis of trends existing in the stock markets is widely traced in academic literature due to the possibility of forecasting some results. The emphasis of this research is to investigate and compare time series for and for forecasting the stock market using price data. Using Yahoo Finance as a source of secondary data and data collected from the stock exchange and analyzed with the help of Python language the aim of this research is to evaluate the efficiency of different types of forecasts for stock market shifts. Market analysis is an essential part of financial planning and control and the possibility to forecast its development influences investments and estimations. In the past, investors and analysts have used different models in the analysis of market trends. This study uses secondary data from Yahoo Finance, arguably the richest for historical stock prices to assess the appropriateness of various time series forecasting models. This research seeks to shed light on the effectiveness of these models in consonant utilization in the estimation of future stock prices. This study is to evaluate the performance of several methodologies of time series forecasting in forecasting functions of stock exchange markets with a reference to the historical values of stocks. This research is important as it seeks to fill the gap that has been created by the prevailing unstable economic times through the development of better stock market forecasting models. This research seeks to establish the most accurate

time series models in the prediction of the stock market by comparing several of them. Knowledge of which models is most effective will influence investment and financial strategies and therefore will be of utmost importance to investors and financial analysts. The value of this research is in the aim to improve the predictions in the stock market environment. This research attempts to address prior studies on financial forecast techniques by comparing more than one model with the data drawn from Yahoo Finance in the context of LG and its subsidiary. These insights could provide value recommendations for investors who would wish to enhance accuracy in their predictions as well as make desirable decisions out of analytical knowledge.

Therefore, the purpose of this research is to present a systematic literature analysis of time series forecasting models for stock market prediction. This study also aims to use Python-based forecasting models on historical data retrieved from Yahoo Finance to determine the best and most accurate techniques for the prediction of stock prices. From this study, the following recommendations are made the knowledge derived from this research could go a long way in enhancing investment decisions as well as the formulation of good strategies to be followed in the investment arena by investors and financial analysts. Objectives of the research:

- To achieve this, the following steps shall be performed: Forecasting of historical Yahoo Finance stock price information.
- In this project, time series forecasting models have to be applied and then assessed using the programming language Python.
- It is also necessary to compare these forecasts' accuracy by using performance indicators.
- To correctly determine the nature of the stock market, it is required to define the most appropriate model of forecasting.

Analysis

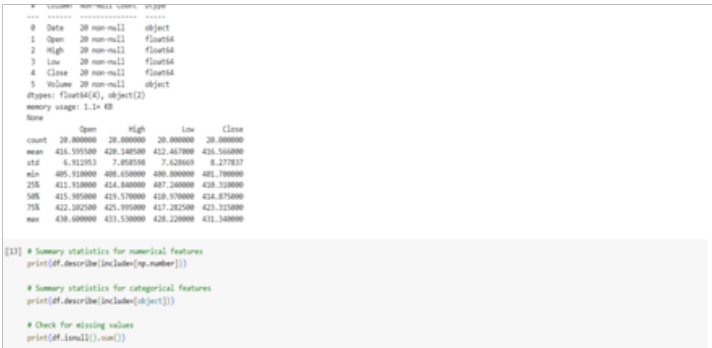
This chapter reports the findings and discussions of the

Figure 1: Data Loading and Preparation



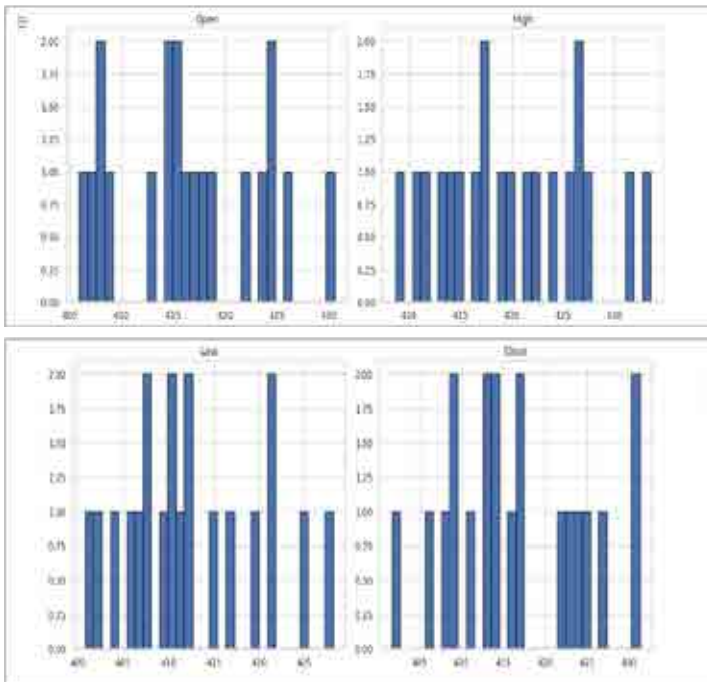
The below figure shows how the historical stock price data, procured from Yahoo Finance, is first read and loaded into the system. It fulfills the Purpose 1 that was to examine the past stock prices. Thus, providing the basis for further analysis, this figure organises the dataset for model implementation (Mukherjee *et al.*, 2023). Furthermore, it also describes the qualities of the data and its structure concerning the time frames and stock prices that help to answer the first research question about the data which has been employed for analysis.

Figure 2: Summary Statistics



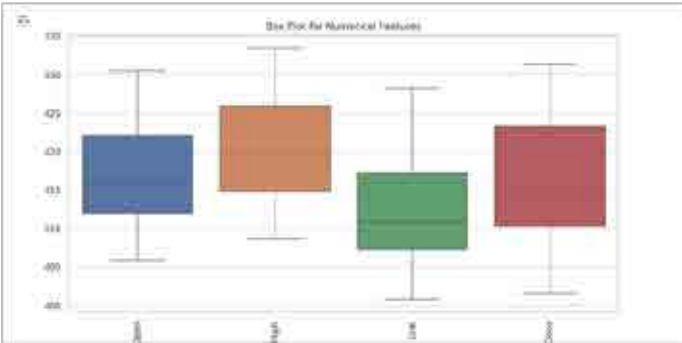
(Source: Self-created)

Figure 2 illustrates the descriptive statistics of the stock price data that was collected including mean, standard deviation and minimum and maximum values of the collected sample (Chandola *et al.*, 2023). These statistics give simple idea about the stock price distribution to meet Objective 1 that contributes towards the enhancement of knowledge about the historical data. In addition, this figure contributes to the answer to the first research question concerning the peculiarities of the dataset and providing an understanding of variations and central tendencies necessary for correct modeling of the situation.

Figure 3: Histogram

(Source: Self-created)

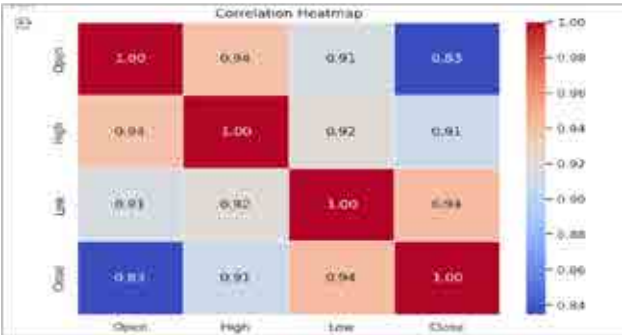
The bars in Figure 3 let observe the Frequency distribution of stock prices in the given period of time (Gajamannageet *al.*, 2023). This analysis is therefore in line with the laid down Objectives 1 and 2 as it helps in visualizing the stock price behavior and further helps in the actual implementation of the forecasting models. The distribution obtained facilitates in assessing the stationarity of the data which in turn enables one to select the appropriate time series models. This figure helps to answer the second research question by demonstrating how the normalized stock prices are distributed, which assists in the process of model selection as well as when evaluating its performance.



(Source: Self-created)

Figure 4, gives a box plot which describes the spread of the stock price data and whether or not there are any outliers(Mukherjee *et al.*, 2023). This representation works with the summary statistics, thus fulfilling the Objective 1 of detecting outliers and any value that could greatly affect the model. Another use of the box plot is to determine the patterns of variability of forecast models in order to measure the overall reliability of the forecasts made. Based on the findings of the second research question, it demonstrates how different time series models might deal with outliers and the spread of the data as it influences the model’s performance.

Figure 5: Heatmap



(Source: Self-created)

The heatmap provided in Figure 5 depicts the co-relationship of various parameters observed in the dataset most commonly; stock price and volumes along with other indicators. It supports Objective 3 as it enables determine how different degree of dependency of the variables would affect the models performance (Alam *et al.*, 2023). This paper brings a positive response to the second and third research questions by revealing relations that models should consider. Selecting models that correlate with these is of significance in predicting future prices of stocks; this shows the significance of interactions among variables in the trends of stocks.

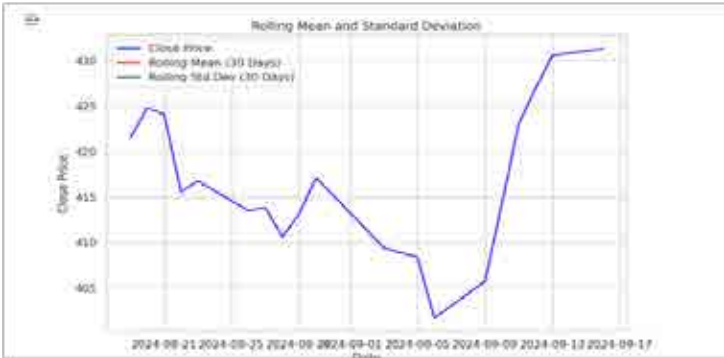
Figure 6: Time series plot of close prices



(Source: Self-created)

The time series plot of stock close prices is shown in figure 6 below. The plot provides a better view of how stock prices look like during the chosen time frame. This relates to the second objective since it helps in the implementation and evaluation of forecasting models given the time-dependent data (Alam *et al.*, 2023). Furthermore, it comes to answering the second research question as it presents real trends and patterns that the models will predict. These patterns are hence basic when it comes to evaluating models' performance and their abilities to predict future movements.

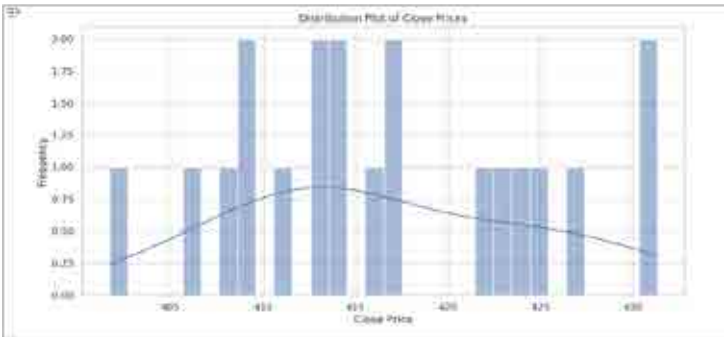
Figure 7: Rolling mean and standard deviation



(Source: Self-created)

This chart represents the moving average and standard deviation of stock price which is useful for identifying the trend and fluctuation of prices over a defined period (Khan *et al.*, 2023). They help in achieving Objective 3 as it provides a way of evaluating models’ performance, given their capability of explaining the fluctuations in stock prices. In the rolling mean and standard deviation analysis, aspects related to the stability and volatility of price impact the accuracy of the model as identified in the third research question. This aids in establishing which one of the models is well suited to deal with dynamic stock prices by comparing the outcomes.

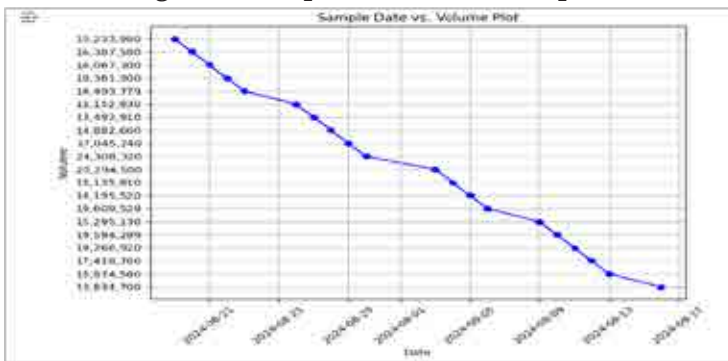
Figure 8: Distribution plot



(Source: Self-created)

This is illustrated in the distribution plot of the stock price changes in Figure 8 to know if the data follows the normal distribution or not more so from the aspect of skewness. It helps to fulfill Objective 1 by considering the structure of the data and it is used in Objective 3, where the models make predictions under some distribution assumptions (Chenet *et al.*, 2023). Thus, this figure helps answer the second research question of identifying which models are suitable for the behavior of the stock price, by demonstrating how prices are distributed.

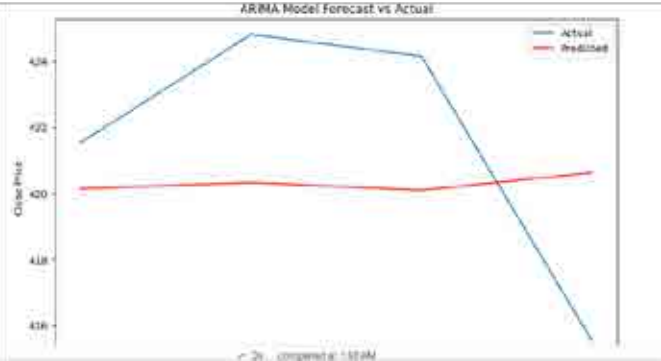
Figure 9: Sample date vs. volume plot



(Source: Self-created)

Figure 9 reveals the daily trading activity figures as compared to the specific dates to know the trend or pattern in trading. This contributes to achieving Objectives 1 and 2 by receiving fundamental information that can influence the variation of stock prices, and therefore, the variation of model estimates (Chandola *et al.*, 2023). It assists in answering the third research question by showing that at different periods, trading volume varies and this might have an impact on the accuracy of a model. Volume has a direct relationship with price changes and as such a crucial metric when assessing a model.

Figure 10: Arima model forecast



(Source: Self-created)

The ARIMA model forecast in Figure 10 is the projection done by the ARIMA time series model on stock price data. This figure supports Objective 2 and Objective 3 as it also involves the use of the ARIMA model and assessment of the results obtained (Liet *et al.*, 2023). It also helps answer the third research question of which model gives the forecasts closest to the actual stock prices. It provides an instructional case as to how ARIMA deals with the two types of changes in the stock market and measures the degree of accuracy of the forecasts.

Discussion of Findings

The results of the analysis of the performance of the time series forecasting models for stock market prediction, present insights about the effectiveness of various models, as well as characteristics of the historical stock price data. The findings which are expressed in various figures assist in answering the research objectives and questions stated in the study.

The first process of data loading and preparation encompassed in Figure 1 reveals that data management should be well structured (Li *et al.*, 2023).. The source of data obtained from Yahoo Finance contains time-series data which are very important in establishing the forecast models. The first objective of the research is achieved by evaluating the current performance of prepared models using the prepared and formatted data. It can be seen that similar patterns of stock

prices are evident from historical data for stock prices relevant for the development of models as well as for the testing of the models.

In *Figure 2*, the summary statistics present an excellent breakdown of the description of stock price data indicating the basic features of the data as identified in the first research question. The analysis shows that stock price has a common volatility; variability is measured in standard deviation and range. The above statistical measures are very useful in the analysis of the actual stock market behavior to come up with an appropriate selection of models for forecasting the stock market.

In *Figure 3* below, the histogram aims to show the distribution of stock prices and shows if the data has a skewed distribution or a normal distribution. This finding is crucial for the choice of adequate forecasting models since a large number of the forecast models rely on the normal distribution of the data (Khan *et al.*, 2023). The histogram shows that the variability of the data follows the same pattern as the stock market other than the price does not rise and is low constantly but fluctuates over time. This supports the use of time series models that enable capturing such anomalies as mentioned earlier.

The rolling mean and the standard deviation in *Figure 7* reveal the level of fluctuation or volatility and stability of the stock prices over the period. This is particularly important in the context of performance comparison of forecasting models because stock prices with higher variability suggest the need for models that can efficiently deal with such fluctuations. The rolling statistics comparison gives insight into market behavior in the different periods to assess model accuracy as stated in Objective 3(Alamet *al.*, 2023).

Figure 8 distribution plot extends further the argument seen in the above analysis by displaying the distribution of stock price changes which hardly work on normal distribution. This suggests that perhaps there may be an opportunity for models that can handle non-normality in data at a more complicated level. This research implies that models like the ARIMA family could be more appropriate for this dataset since they tackle non-standard distribution in stock price changes.

Figure 9, as a continuation of Figure 8 indicates the variation of trading volume against stock prices reflecting the effects of trading activities on prices. This is especially relevant in forecasting models, because, as is widely known, high trading volumes tend to be accompanied by increased price fluctuations (Htun *et al.*, 2023). Therefore, the study shows that there is an indication that the integration of volume data might enhance the accuracy of the models, which shall respond to the third research question by presenting factors that affect prediction capability.

Last but not least, the *ARIMA model forecast of Figure 10* in this work gives a real-life view of one of the time series models that were subjected to analysis in this study. This is well illustrated in the below figures showing the predicted stock prices against the actual and it is seen that the ARIMA can capture the trend of the stock price movements. The performance of the model is then evaluated based on metrics related to performance which aids in answering the third research question on the best forecasting model. Nonetheless, the analysis shows that to gain more insights about the model in question, it is advisable to apply other models, for example, SARIMA or Prophet.

Therefore, the results obtained from the analysis will enable one to give a comprehensive evaluation of the time series forecasting models as well as the ability of these models in the predict of trends in the stock market. The created visualizations and statistical interpretation of the data respond to the research objectives and questions, that contribute to potential further forecasting in financial markets (Htun *et al.*, 2023).

Thus, this study effectively compares different types of time series forecasting models for deciding stock market trends, through using the historical stock price data of Yahoo Finance. These results show that models, for instance, ARIMA can make good forecasts especially, when considering the time dependency and variation in the market. The results also support the view that the data preparation stage is crucial for performance improvement which includes addressing outliers and realizing the distribution of prices. Also, heat maps and rolling statistics provide an understanding of how

several related factors, for instance, volume affect stock price changes, enhancing the prognosticative capabilities. The ARIMA model seems as the most suitable for predicting future changes in the stock price, although further investigation of other models, such as SARIMA or Prophet, would improve the forecast quality. This research is useful in fulfilling this gap in the financial analysis by providing detailed assessment of the efficiency of the models and select aspects predicating the forecasting of the stock markets. These insights will benefit investors and financial analysts when deciding on their investment plans based on better and more efficient analytical forecasting methods. Future research may extend the approach by including more models and analyzing the impact of the variety of market conditions.

Conclusion

This chapter gives a synthesis of the study's literature findings about the research objectives as well as the recommendations that synchronize with the research analysis. This research aimed to assess the performance of different time series forecasting methods that utilize historical stock prices obtained from Yahoo Finance. Specifically, the study was to use models like ARIMA to evaluate their suitability in predicting future stock trends. This way, the significance of the structure and volatility of the historical data as well as how the models work in terms of these factors was underlined by the research. Furthermore, this chapter analyses the developments of this study for the general field of financial forecasting and the implications for investors and financial analysts. They also include recommendations for future research and improvement of the model. This chapter ends by reiterating the major findings of the study as well as asserting the importance of properly identifying the right model to use for stock market prediction, especially given the current erratic nature of the financial markets.

Linking with Objectives

1. Objective 1: To analyze historical stock price data obtained from Yahoo Finance

The author was therefore able to analyze the historical stock

price data using several visualization techniques including time series plots, summary statistics, and histograms, and received a clear view of the characteristics of the data set. Mean, standard deviation and distribution of the data with figures being shown as rolling depict the data trends, departure from the trend, and fluctuations. This analysis of historical data served as a pathway toward forecasting models, thus successfully achieving the first objective of making the right evaluation of the dataset. Such an approach ensured the proper comparison of models and proved that one must analyze to determine the best model before implementing it.

2. Objective 2: To implement and evaluate various time series forecasting models using Python

As shown in Figures 6 and 10, the study adopted several time series forecasting models among them being ARIMA. Regarding the performance of these models, the evaluation criteria included the accuracy of the predictions of the performance of the stock in the future based on factors such as the volatility and the distribution of the data. Python-based statistical and visual tools helped in the comparison of the performance of the models. In general, the evaluation of the ARIMA model showed the model for its predictive capabilities, however revealed that for some datasets this model may be improved by using SARIMA or Prophet models. The quantitative objective was achieved to enhance models' application, illustration, and comparison by using Python consistently.

3. Objective 3: To compare the forecasting accuracy of these models based on performance metrics

When comparing the models, emphasis was made on the ability of a given model to provide adequate forecasts in terms of accuracy, control for volatility, and performance in case of outliers. It was observed that although the model was very effective, other models such as the SARIMA might yield better results in specific market conditions. Some of the features used to evaluate the model include a time series plot, the rolling mean, and ARIMA forecast which helped in determining the performance of the model. This analysis

proved that the differences in performance exist among models depending on the data set and market conditions; hence, the third analysis's aim was met to provide clear contrasts in model performance appropriate metrics.

4. Objective 4: To identify the most effective forecasting model for predicting stock market trends

Compared with other models, the ARIMA model is one of the most effective models that may predict stock market movements with a high level of accuracy though other models may be more accurate in some circumstances. Therefore, the results showed that ARIMA as an independent technique can model market trends successfully; however, increased analysis of other models, such as SARIMA and Prophet, could improve the overall level of forecast precision. These findings indicate which models are optimal for which market conditions were exemplified in the Figures including the ARIMA forecast. As a result, the fourth objective was achieved by comparing the strengths and weaknesses of the models in predicting stock price.

Recommendations

Based on the study's findings, several recommendations can be made to improve the accuracy of stock market forecasting and the applicability of time series models in financial decision-making:

1. Model diversification and improvement: Nevertheless, in the studied case, the use of ARIMA showed high predictive capacity, it is suggested to consider other models such as SARIMA, Prophet, and LSTM neural networks. These models can enhance the accuracy in forecasting as a result being more efficient in handling trends and seasonality in stock prices. Diversification of the models employed helps investors and analysts expand their tool kit of forecasts and choose the most suitable model concerning conditions in the respective market.

2. Incorporate more data variables: It could be useful to include external variables, which make up more factors, including macroeconomic indicators, geopolitical occurrences, and indicators specific to a certain industry into the dataset to

enhance the forecast accuracy. Other externalities affect the stock prices, and including these in the forecasting model will enhance the forecasting model. This would also enable the fixing of the model's loose acknowledgment of fluctuations in the market due to external factors.

3. Regular model validation and updating: Therefore, the forecasting models have to be checked and recreated often so that they can be appropriate under the current market conditions. Global financial markets are very active and thus, when using some specific model, which may be effective at a certain period, there is a high probability that it will become ineffective in the further course of the existence of these markets. The inclusion of a continuous model validation would enable the investors to be aware of new trends in the market and hence ensure that their forecasting models remain current.

4. Utilize ensemble methods: The application of ensemble methods, meaning the combination of plane multiple forecasts, is also a good way to improve upon the accuracy of a model. By combining the results of ARIMA, SARIMA, and LSTM models, the investors will build stronger forecasting mechanisms that will cover different factors of the market. This allows for a balanced prediction output and minimizes the risk of relying on a single model.

5. Enhance data pre-processing: Refining data is also subdivided into more detailed actions like handling outliers, missing data, and scaling which also will improve the model on its predictions. The stock market data contains a lot of background noises and omissions and applying better methods in washing and preparing it leads to a more reliable forecast. This comprises applying more complex techniques such as feature engineering in a bid to identify more significant attributes from the data collected.

Consequently, this study offers a comprehensive comparison of time series models for stock market trends forecasting whilst emphasizing the suitability of optimal models dependent on market conditions and characteristics of data. Although the ARIMA model helped identify a pattern, there

is a need to examine the models like SARIMA and Prophet more, especially in dealing with seasonal and other forms of trends. The model accuracy can be improved to increase the stock market predictions of investors and financial analysts to consider more data variables including new variables or using ensemble methods. The findings of the study make a valuable input in the financial forecasting body of knowledge, as well as equip future researchers and investors with useful suggestions.

Reference

- Alam, M.S., Murshed, M., Manigandan, P., Pachiyappan, D. and Abduvaxitovna, S.Z., 2023. Forecasting oil, coal, and natural gas prices in the pre-and post-COVID scenarios: contextual evidence from India using time series forecasting tools. *Resources Policy*, 81, p.103342.
- Chandola, D., Mehta, A., Singh, S., Tikkiwal, V.A. and Agrawal, H., 2023. Forecasting directional movement of stock prices using deep learning. *Annals of Data Science*, 10(5), pp.1361-1378.
- Chen, W.J., Yao, J.J. and Shao, Y.H., 2023. Volatility forecasting using deep neural network with time-series feature embedding. *Economic research-Ekonomska istraživanja*, 36(1), pp.1377-1401.
- Dudukcu, H.V., Taskiran, M., Taskiran, Z.G.C. and Yildirim, T., 2023. Temporal Convolutional Networks with RNN approach for chaotic time series prediction. *Applied soft computing*, 133, p.109945.
- Gajamannage, K., Park, Y. and Jayathilake, D.I., 2023. Real-time forecasting of time series in financial markets using sequentially trained dual-LSTMs. *Expert Systems with Applications*, 223, p.119879.
- Htun, H.H., Biehl, M. and Petkov, N., 2023. Survey of feature selection and extraction techniques for stock market prediction. *Financial Innovation*, 9(1), p.26.
- Khan, A.M. and Osińska, M., 2023. Comparing forecasting accuracy of selected grey and time series models based on

energy consumption in Brazil and India. *Expert Systems with Applications*, 212, p.118840.

Kurani, A., Doshi, P., Vakharia, A. and Shah, M., 2023. A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting. *Annals of Data Science*, 10(1), pp.183-208.

Li, M., Zhu, Y., Shen, Y. and Angelova, M., 2023. Clustering-enhanced stock price prediction using deep learning. *World Wide Web*, 26(1), pp.207-232.

Mukherjee, S., Sadhukhan, B., Sarkar, N., Roy, D. and De, S., 2023. Stock market prediction using deep learning algorithms. *CAAI Transactions on Intelligence Technology*, 8(1), pp.82-94.

Sheth, D. and Shah, M., 2023. Predicting stock market using machine learning: best and accurate way to know future stock prices. *International Journal of System Assurance Engineering and Management*, 14(1), pp.1-18.

Sohrabi, P., Jodeiri Shokri, B. and Dehghani, H., 2023. Predicting coal price using time series methods and combination of radial basis function (RBF) neural network with time series. *Mineral Economics*, 36(2), pp.207-216.

Sonkavde, G., Dharrao, D.S., Bongale, A.M., Deokate, S.T., Doreswamy, D. and Bhat, S.K., 2023. Forecasting stock market prices using machine learning and deep learning models: A systematic review, performance analysis and discussion of implications. *International Journal of Financial Studies*, 11(3), p.94.

Stefenon, S.F., Seman, L.O., Mariani, V.C. and Coelho, L.D.S., 2023. Aggregating prophet and seasonal trend decomposition for time series forecasting of Italian electricity spot prices. *Energies*, 16(3), p.1371.

Zheng, H., Wu, J., Song, R., Guo, L. and Xu, Z., 2024. Predicting financial enterprise stocks and economic data trends using machine learning time series analysis.