TIME SERIES ANALYSIS AND FORECASTING FOR STOCK MARKET

Submitted in partial fulfilment

by

Sohansingh Pardeshi
Akshay Jankar
Uday Katkar
Yuvraj Tamrakar
Vansh Raina

Under the Guidance of

Mr. Chandan Mishra

Acknowledgement

In successfully completing our project, I am indeed thankful to my group and my guide who had encouraged and helped me in a variety of ways. It's definitely a privilege to express my gratitude to the people involved. My profound and sincere thanks to Mr Chandan Mishra who consented to be our project guide. I have a scarceness of words to express my courtesy towards him for their keen interest, valuable guidance, important judgement and constant encouragement during the entire course of my study. I fall short to express my feelings towards Mr Chandan Mishra for his timely and important help in completing this project report.

Project Members:

Sohansingh Pardeshi
Akshay Jankar
Uday Katkar
Yuvraj Tamrakar
Vansh Raina

Abstract

This project, titled **“Time Series Analysis and Forecasting for Stock Market”**, aims to explore and implement various forecasting techniques to predict stock price movements based on historical data. The unpredictable nature of the stock market makes forecasting a valuable yet challenging task. By leveraging time series modelling techniques such as ARIMA, SARIMA, Facebook Prophet, and LSTM, the project attempts to identify patterns, trends, and seasonality in financial time series data. These models are applied and evaluated to determine their effectiveness in predicting future stock prices. The project provides practical exposure to real-world financial data and develops skills in data preprocessing, statistical modelling, deep learning, and result visualization. Through a comparative analysis of model performance, the project concludes with insights into which approaches work best under specific conditions, making it an enriching experience for anyone aiming to understand financial forecasting.

List Of Figures

Contents

Chapter 1: Introduction

The stock market is one of the most dynamic and complex financial systems in existence, influenced by a variety of economic, social, political, and psychological factors. Investors, traders, and analysts constantly seek ways to anticipate market behaviour to make informed decisions. Forecasting stock prices using historical data is one such method, rooted in time series analysis. Time series data contains valuable patterns such as trends, cycles, and seasonality, which can be leveraged to predict future outcomes. In this project, we delve into various forecasting techniques, starting from classical statistical methods like ARIMA and SARIMA to modern machine learning-based models like Facebook Prophet and deep learning models like LSTM. These models are evaluated on their ability to forecast stock prices with accuracy and efficiency. By systematically analysing the data and comparing models, the project bridges theoretical concepts with practical applications in the domain of financial analytics.

Chapter 2: Methodology

The methodology for this project begins with data collection from trusted sources such as Yahoo Finance, Kaggle, and Alpha Vantage. The dataset generally consists of daily stock values including Open, Close, High, Low, Volume, and Date. Once the data is acquired, it undergoes rigorous preprocessing. This involves handling missing values, converting date columns to datetime format, and setting them as indices. To ensure compatibility with models like ARIMA, the data is tested for stationarity using the Augmented Dickey-Fuller test and transformed using differencing or log scaling if necessary.

The next step involves the application of four core models. ARIMA, a traditional time series model, is used for forecasting based on autoregressive and moving average components. SARIMA extends ARIMA by including seasonal components to capture repeating patterns in the data. Facebook Prophet is employed for its ability to manage missing data, abrupt trend changes, and seasonality, making it particularly suitable for business-like stock time series. Lastly, LSTM—a deep learning model—is used for its ability to learn long-term dependencies in sequences, making it a powerful tool for complex and nonlinear stock price patterns.

Each model is trained on a portion of the historical data and tested on a holdout set. Performance is evaluated using metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). Visual comparisons between predicted and actual values are also performed to better interpret the forecasting accuracy. The entire implementation is done using Python and libraries such as Pandas, NumPy, Matplotlib, Seaborn, Statsmodels, Scikit-learn, TensorFlow/Keras, and Facebook Prophet. Optional deployment of results is facilitated through Streamlit or Flask-based dashboards.

Chapter 3: Results

```
--- Evaluation Complete ---
Best ARIMA(2, 2, 2) RMSE: 65.0279
```

Fig 3.1 (ARIMA)

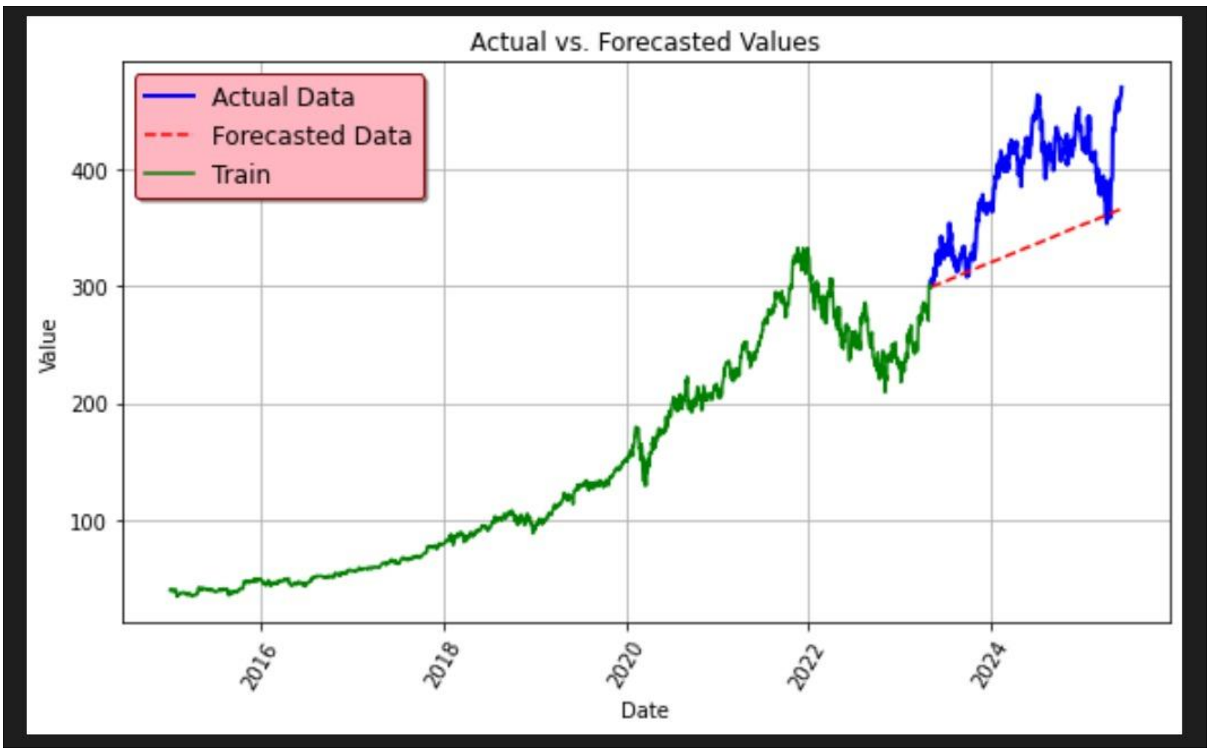

Fig 3.2 (ARIMA)

```
rmse = np.sqrt(mean_squared_error(test_data,forecast))
print(rmse)
```

```
64.98863314748812
```

Fig 3.3 (Sarima Forecast)

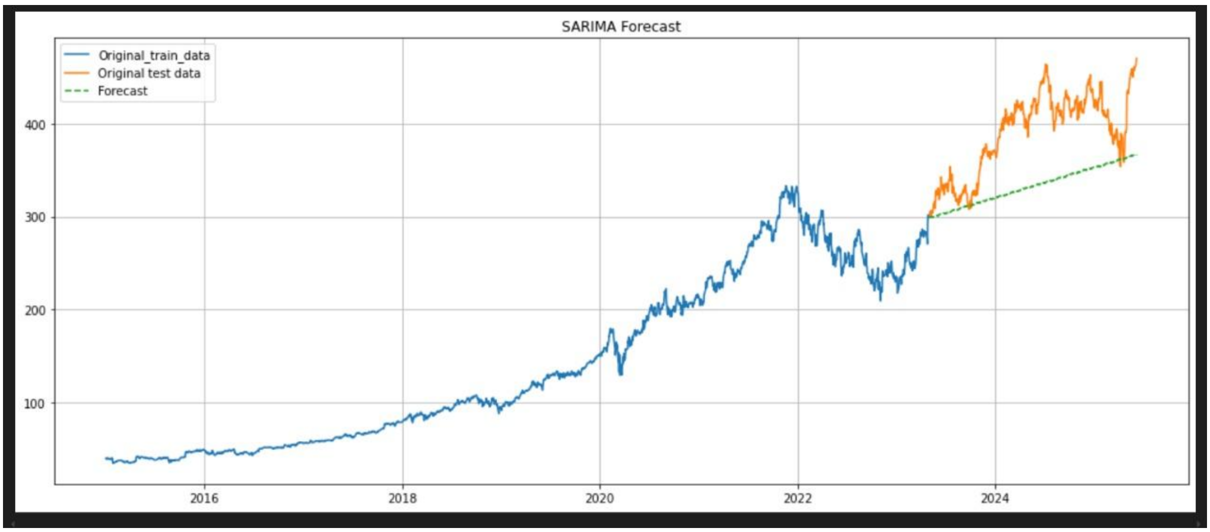

Fig 3.4 (Sarima Forecast)

```
rmse = np.sqrt(mean_squared_error(y_test,forecast))
print(rmse)
```

```
7.535450773570006
```

Fig 3.5 (Sarimax)

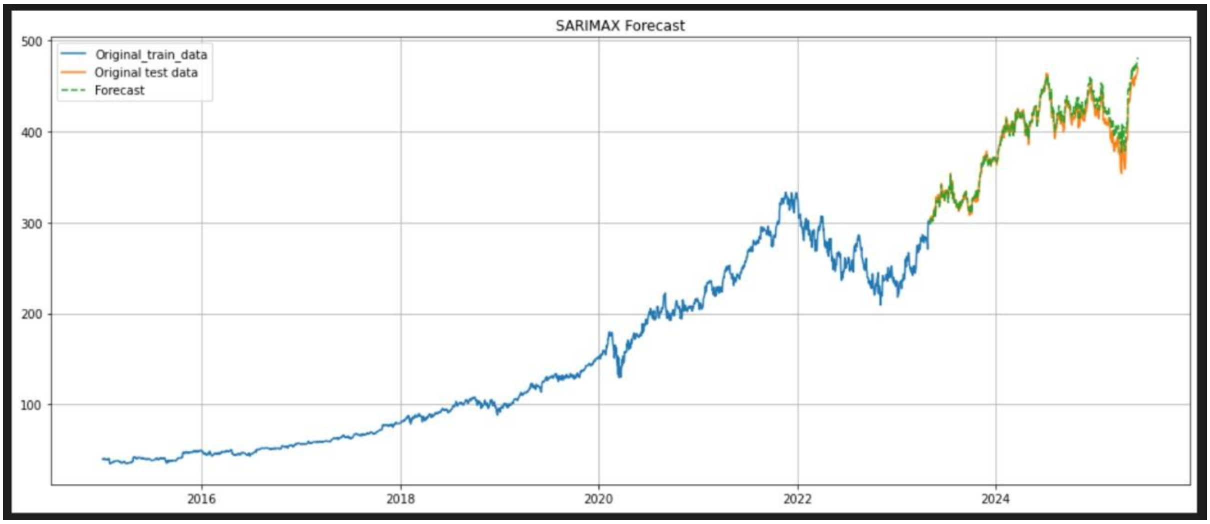

Fig 3.6 (Sarimax)

Fig 3.7 (Prophet Model)

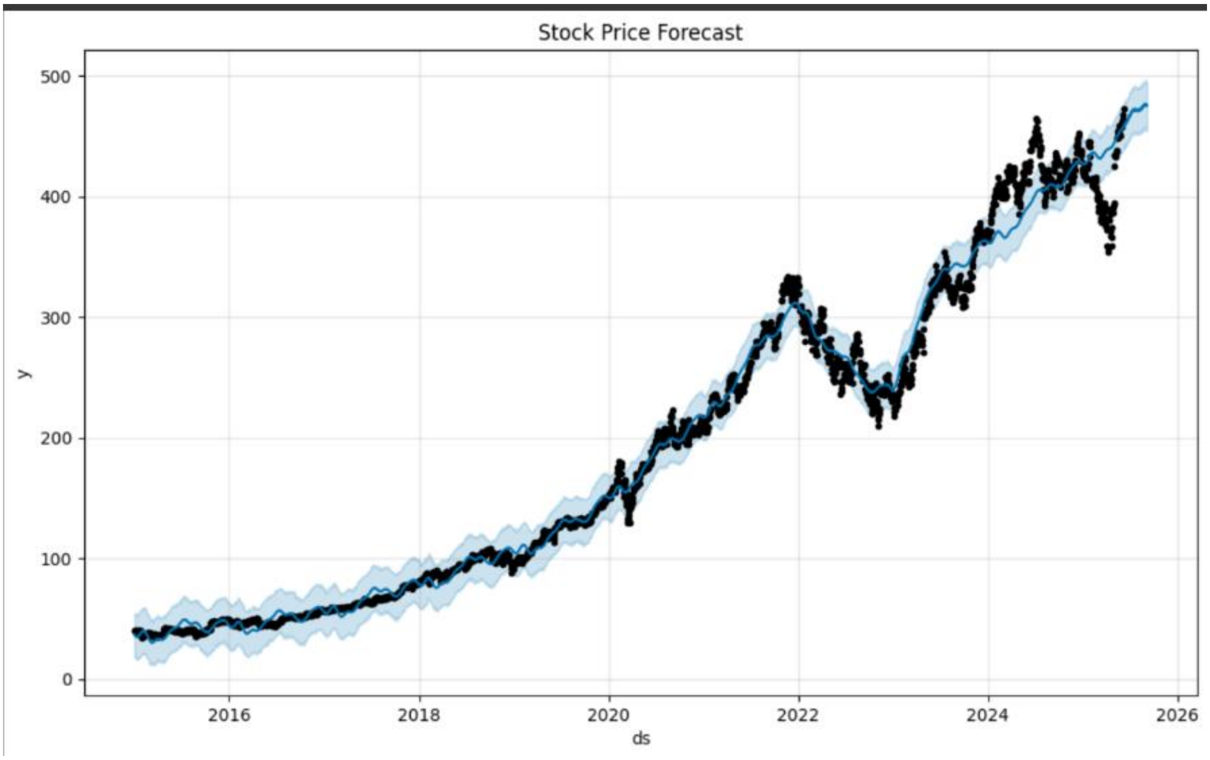

Fig 3.8 (Prophet Model)

Chapter 4: Conclusion

This project demonstrates the application of time series analysis in forecasting stock market trends and comparing the efficacy of multiple forecasting models. It was observed that classical models like ARIMA and SARIMA are best suited for datasets with linear patterns and minimal noise, while models like Facebook Prophet excel when seasonality and trend changes are prominent. LSTM, on the other hand, is highly effective for complex, nonlinear, and large-scale data. Although it requires more training and fine-tuning, the results often justify the effort. Through this project, we concluded that the model choice should depend on the specific characteristics of the stock data being analyzed. Additionally, visualizations and evaluation metrics are critical for understanding the quality of forecasts. This study offers a comprehensive framework that can be adopted or extended by financial analysts and data scientists for practical stock market forecasting.

Chapter 5: Future Scope

The potential for extending this project is vast. One future direction involves building hybrid models that combine the strengths of both statistical and deep learning approaches—for example, combining ARIMA for short-term accuracy and LSTM for long-term pattern recognition. Another enhancement could include the integration of sentiment analysis derived from news headlines, social media trends, or financial reports to improve prediction quality by adding external influencing factors. The current model could also be scaled for multi-stock or portfolio-level forecasting, which would make it more useful for institutional investors. Deployment of the model as a real-time dashboard using Streamlit or Flask could provide live predictions and alerts. Furthermore, reinforcement learning algorithms could be explored to create automated trading agents that learn from the predictions and make investment decisions. Lastly, making these AI-based forecasts explainable through tools like SHAP or LIME would increase trust and usability in real-world financial environments.

Chapter 6: References

1. Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3rd ed.). OTexts.
   [Available online at: https://otexts.com/fpp3/]
   (An essential textbook on time series forecasting using R, but the theory applies broadly.)

2. Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control* (5th ed.). Wiley.
   (A classic reference for ARIMA and statistical time series models.)

3. Yahoo Finance. (n.d.). *Historical Stock Data*. Retrieved from https://finance.yahoo.com
   (Used as a source of financial time series data.)

4. Quandl. (n.d.). *Financial, Economic and Alternative Data*. Retrieved from https://www.quandl.com
   (Reliable source for financial and economic time series datasets.)

5. Kaggle. (n.d.). *Datasets for Time Series Analysis*. Retrieved from https://www.kaggle.com/datasets
   (Repository for diverse time series datasets and competitions.)

6. Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. *PLOS ONE*, 13(3), e0194889. https://doi.org/10.1371/journal.pone.0194889
   (Comparison of statistical vs machine learning methods for time series.)

7. Brownlee, J. (2017). *Deep Learning for Time Series Forecasting: Predict the Future with MLPs, CNNs and LSTMs in Python*. Machine Learning Mastery.
   (Practical guide for neural network approaches in time series forecasting.)