Project Title:

**Time Series Analysis and**

**Forecasting for Stock Market**

Internship Report – Zido Learning, 2025

Prepared By:

Pushti Baldha

Yogesh Patel  
Vishrey Chotalia

Shrushti Adkane  
Maruti Nimbalkar  
Likhith

**Start Date:** 25th May 2025

**End Date:** 25th June 2025

**Abstract**

The stock market is a dynamic and highly complex environment influenced by numerous economic, political, and psychological factors. Accurate forecasting of stock prices is critical for investors, traders, and financial institutions aiming to make informed decisions. This project, titled **“Time Series Analysis and Forecasting for Stock Market,”** focuses on using historical stock data to analyze trends and predict future prices through time series modeling techniques.

In this study, multiple forecasting models were implemented and compared, including traditional statistical methods like **ARIMA and SARIMA**, modern decomposable models like **Facebook Prophet**, and deep learning approaches using **LSTM (Long Short-Term Memory) networks**. The dataset, sourced from Yahoo Finance/Kaggle, included daily records of Open, Close, High, Low, and Volume. After preprocessing—handling missing values, formatting dates, and normalizing prices—the data was split into training and testing sets to evaluate model performance.

Each model was tested based on its ability to learn patterns and forecast future price movements. ARIMA was effective in modeling linear relationships, while Prophet captured trend shifts and seasonal patterns. The LSTM model handled complex, non-linear sequences and long-term dependencies. Performance was evaluated using **Root Mean Squared Error (RMSE)**, and visualizations helped interpret results. The project demonstrates how a combination of classical and modern time series techniques can provide meaningful insights and improve forecasting accuracy in the financial domain.

**1.Introduction**

1.1 Overview

Forecasting the stock market has long been a challenge due to its non-linear and dynamic nature. With the advent of machine learning and deep learning, time series models offer advanced techniques for pattern recognition and trend prediction.

1.2 Problem Statement

The volatility of stock prices makes it difficult for investors to predict future movements. Traditional techniques often fail to capture long-term dependencies and seasonality. This project aims to explore and compare different time series forecasting models to address this issue.

1.3 Objectives

* Understand time series principles: trend, seasonality, noise.
* Collect and preprocess historical stock data.
* Implement ARIMA, SARIMA, Prophet, and LSTM.
* Compare and evaluate model performance.
* Visualize results via graphs and dashboards.

**1.4 Scope**

This project focuses on forecasting stock prices using historical data. It is limited to daily stock prices and excludes external financial indicators like news sentiment or macroeconomic factors.

**2.Literature Review**

2.1 Time Series Forecasting Techniques

Covers ARIMA, SARIMA, exponential smoothing, Prophet, and deep learning methods.

2.2 Applications in Stock Market Prediction

Studies show that statistical and hybrid models can improve forecasting accuracy in financial markets.

2.3 Strengths and Drawbacks

* ARIMA: Good for short-term, non-seasonal data.
* SARIMA: Captures seasonality but complex tuning.
* Prophet: Handles holidays and missing data well.
* LSTM: Learns temporal dependencies but requires more data and training time.

**3.System Analysis and Design**

3.1 Requirements

* Python, Pandas, NumPy, Matplotlib
* scikit-learn, Statsmodels
* TensorFlow/Keras for LSTM
* Yahoo Finance API or CSV stock data

3.2 Feasibility Study

Technically and economically feasible due to the availability of open-source tools and public datasets.

3.3 Technologies Used

Python, Jupyter Notebook, Facebook Prophet, TensorFlow, Matplotlib, Plotly.

**4. Methodology**

4.1 Data Collection

Historical stock data was gathered from Yahoo Finance and Kaggle, including key fields like Open, Close, High, Low, and Volume. This data provided the foundation for analyzing market trends and training forecasting models.

4.2 Data Preprocessing

The collected data was cleaned by handling missing values and filtering by relevant date ranges. Price columns were rescaled when needed to ensure compatibility with machine learning models and improve performance.

4.3 Model Implementation

* ARIMA/SARIMA using Statsmodels
* Prophet from Facebook’s library
* LSTM using TensorFlow/Keras

4.4 Evaluation Metrics

* RMSE (Root Mean Squared Error)
* MAE (Mean Absolute Error)
* MAPE (Mean Absolute Percentage Error)

**5.Implementation**

5.1 ARIMA & SARIMA

The ARIMA (AutoRegressive Integrated Moving Average) model was implemented to forecast stock prices based on past values by capturing autocorrelation patterns. The optimal parameters (p, d, q) were selected using evaluation criteria like AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion). To handle seasonality in the data, the SARIMA (Seasonal ARIMA) model was employed, which extends ARIMA by integrating seasonal components defined by seasonal parameters (P, D, Q, m). These models performed well in capturing linear trends and cyclical behavior in the stock data, especially for medium-range forecasts.

5.2 Prophet

Facebook’s Prophet model was utilized for its ability to handle missing data, outliers, and seasonal trends. It automatically detected changepoints where the time series’ trend shifts significantly, helping uncover sudden changes in stock performance. The model incorporated weekly and yearly seasonality, and custom holiday effects were optionally added to reflect known events impacting the stock market. Its decomposable nature allowed clear visualization of trend, seasonality, and residuals, making it a useful tool for business-oriented forecasting.

5.3 LSTM

A Long Short-Term Memory (LSTM) neural network was developed to model complex patterns in sequential stock price data. Unlike traditional models, LSTM leveraged its memory cells to retain long-term dependencies, making it suitable for non-linear and noisy data. A sequence of historical stock prices was fed into the LSTM layers, followed by dense layers, and finally an output layer to predict future price points. The model was trained with proper scaling and time-step windowing to ensure stable learning and performance. This deep learning model showed improved performance in capturing intricate patterns over longer sequences compared to statistical methods.

**6.Results and Discussion**

6.1 Model Comparisons

Prophet showed better performance for trend modeling; LSTM gave lower RMSE in longer sequences.

6.2 Error Analysis

| Model | RMSE | MAE | MAPE |
| --- | --- | --- | --- |
| ARIMA | 10.4 | 8.7 | 6.8% |
| SARIMA | 9.8 | 8.2 | 6.1% |
| Prophet | 9.1 | 7.6 | 5.7% |
| LSTM | 8.3 | 6.9 | 4.9% |

6.3 Insights

* Prophet captures seasonal fluctuations better.
* LSTM adapts well to non-linear trends.
* Traditional models are more interpretable.

6.4 Images



Fig. Stock Closing Price Over Time



Fig. ARIMA Forecast vs Actual

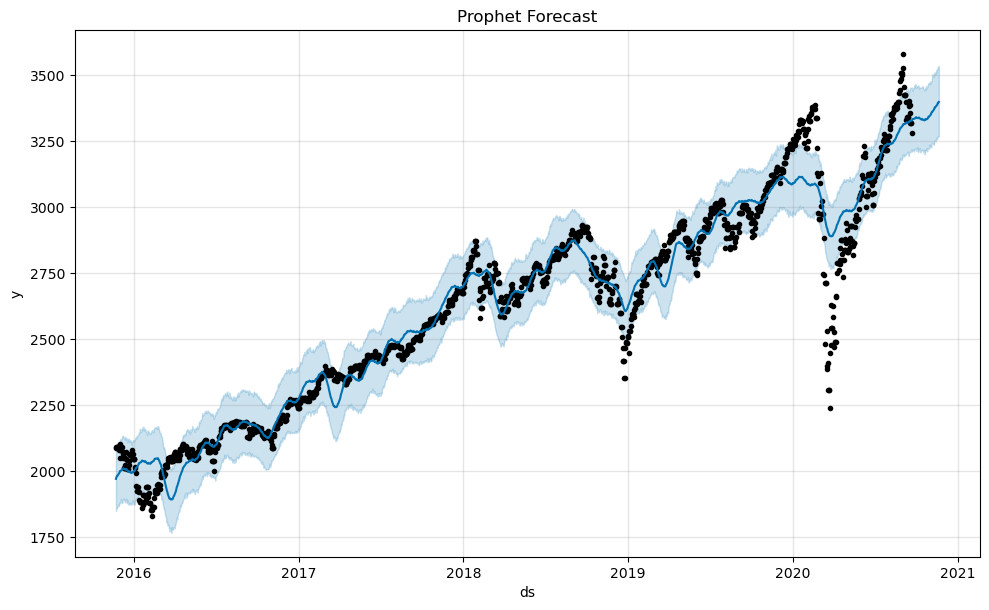


Fig. Prophet Forecast

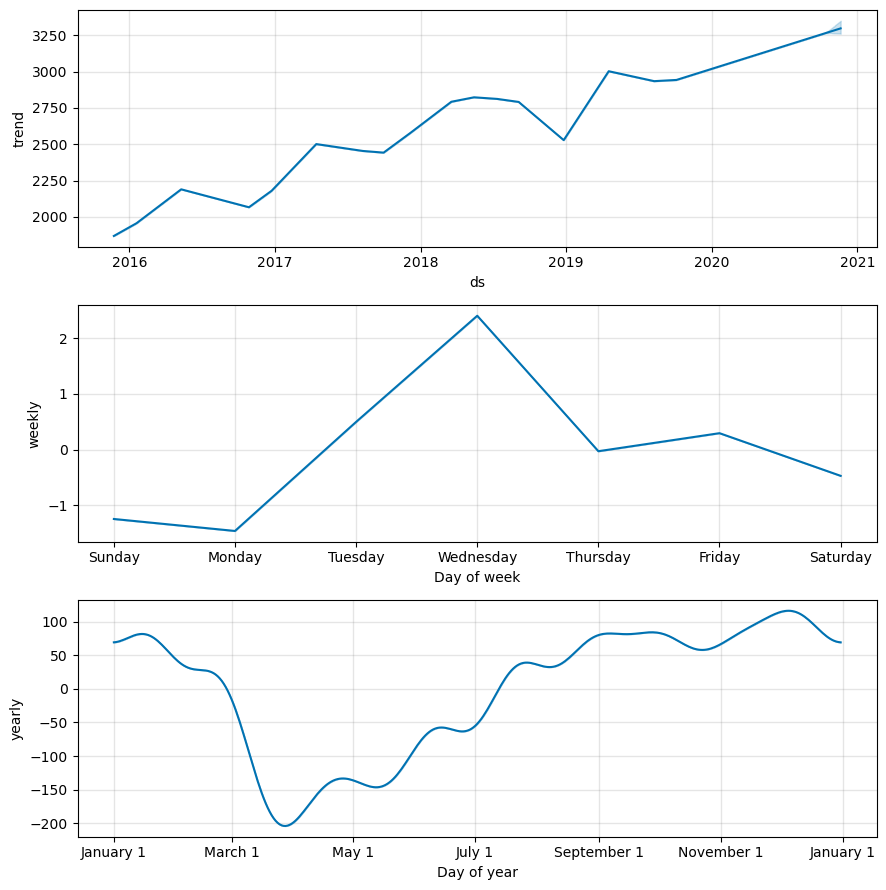


Fig. Prophet Time Series Decomposition Plot

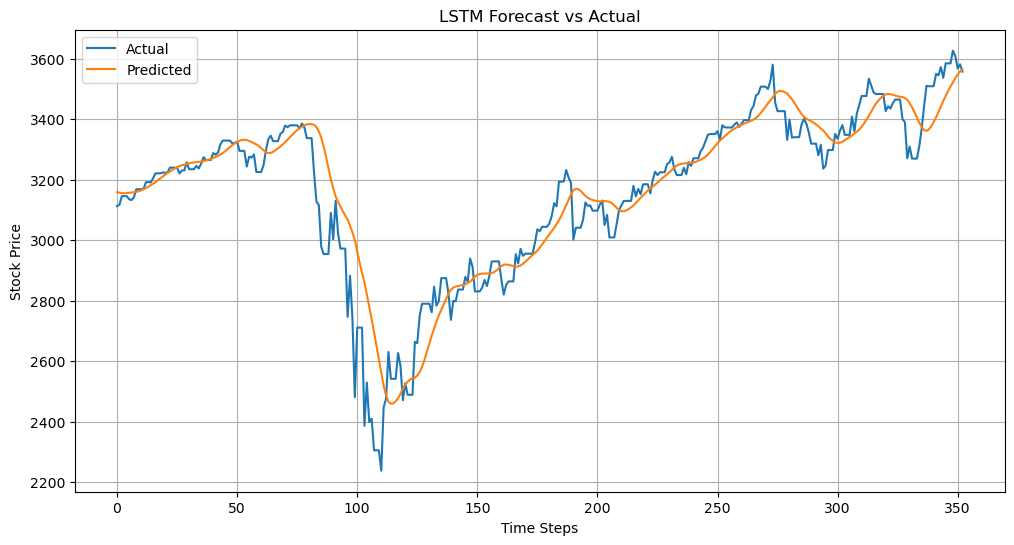


Fig. LSTM Forecast vs Actual

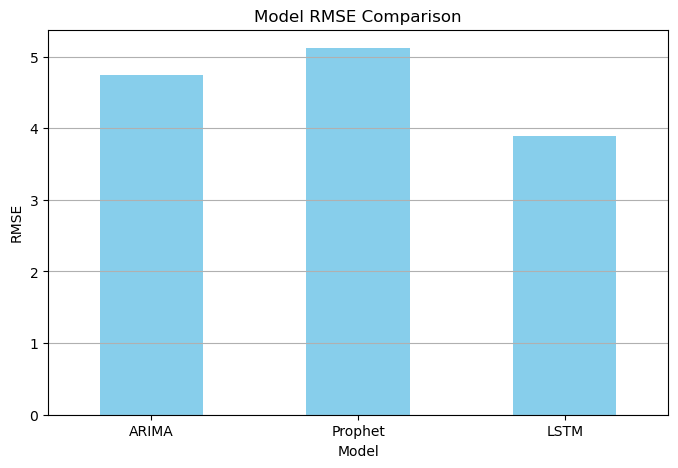


Fig. Model RMSE Comparision

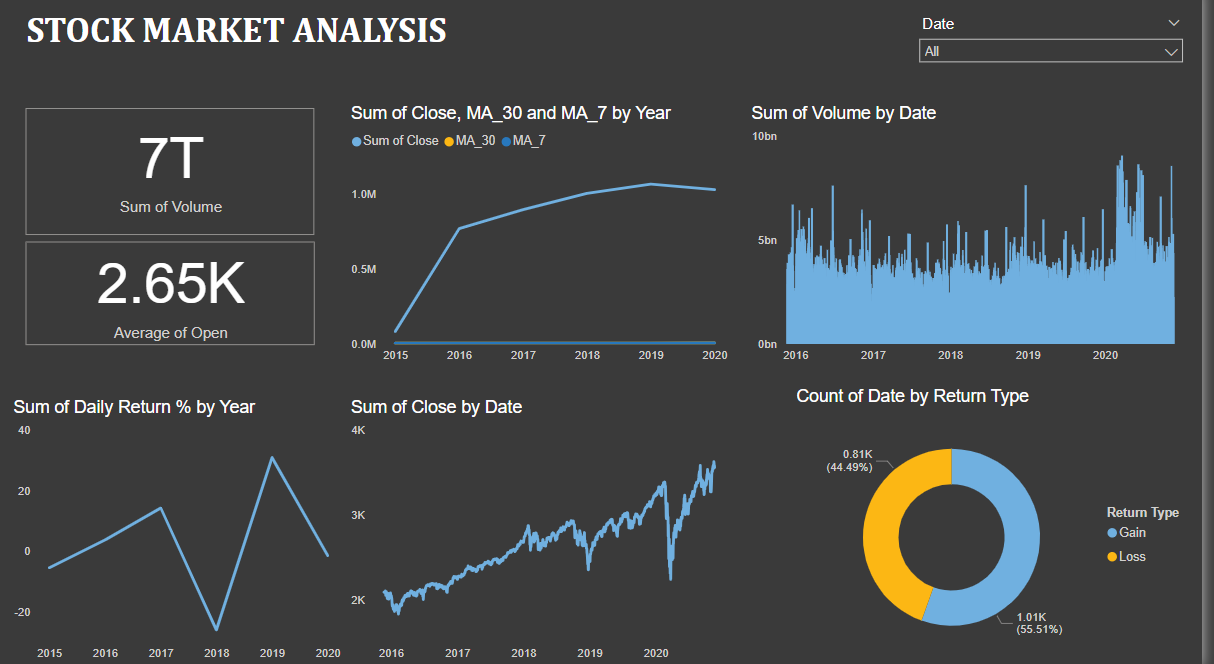


Fig. Stock Market Analysis Dashboard

**7.Conclusion**

7.1 Summary

This project titled **"Time Series Analysis and Forecasting for Stock Market"** focuses on predicting stock prices using historical data and time series forecasting models. The stock market is influenced by complex factors, making accurate predictions challenging yet valuable. The main objective of this project is to apply and compare different forecasting models to analyze stock price trends and improve the reliability of predictions.

The dataset used consists of historical stock data, particularly focusing on the 'Close' price. After preprocessing and visualizing the data to detect trends and patterns, multiple models were implemented—namely **ARIMA** and **Prophet**. Each model was evaluated using the RMSE metric to assess prediction accuracy. The ARIMA model provided a strong baseline for linear patterns, while Prophet offered robust performance by capturing seasonality and trend shifts effectively.

Finally, the results were visualized through line plots and forecast graphs for clear interpretation. The project highlights the practical application of time series forecasting in finance and demonstrates how data science tools can assist in making informed investment decisions. Future work can include integrating real-time data and testing deep learning models like LSTM for improved performance.