Time Series Analysis of ICICI Bank Stock Prices

Using Holt-Winters Method & SARIMA Model

Authors:

Sahil Lakhmani

Ashish Ranjan

Supervisor:

M.R. Srinivasan

Adjunct Professor, Chennai Mathematical Institute

November 11, 2024

Introduction

Background

Time series analysis is a statistical technique used to analyze and interpret data points collected or recorded at specific time intervals. The goal of time series analysis is to identify patterns, trends, and seasonal components in the data, allowing for accurate forecasting of future values. It is widely applied in various fields, including finance, economics, healthcare, and meteorology, due to its capability to model dynamic and evolving processes.

In the context of financial markets, time series analysis plays a crucial role in understanding and predicting the movement of stock prices. Stock prices are inherently volatile and influenced by various factors such as company performance, market sentiment, economic indicators, and geopolitical events. By examining historical price data, analysts aim to capture underlying trends and seasonal patterns that can inform investment decisions.

For this project, we focus on analyzing the stock prices of ICICI Bank, one of India's leading private sector banks. ICICI Bank's stock is a key component of major stock indices in India, making it an attractive choice for investors and analysts. The ability to accurately forecast its stock price movements can provide valuable insights for market participants, helping them to make informed trading decisions.

To model the stock price data, we employ the Holt-Winters Method as our primary forecasting approach, given its effectiveness in capturing both trend and multiplicative seasonal components in financial time series. Additionally, we conduct a comparative study using the Seasonal Autoregressive Integrated Moving Average (SARIMA) model, a robust extension of the ARIMA model that accounts for seasonality.

Motivation

The motivation for this project stems from the need for reliable forecasting methods in the financial sector. Accurate predictions of stock prices can help investors optimize their portfolios and mitigate risks. Given the volatility of the stock market, having a model that can capture both short-term fluctuations and long-term trends is essential. ICICI Bank's prominence in the banking sector and its significant market capitalization make it an ideal candidate for analysis.

In our study, a comparative analysis was conducted between the Holt-Winters Method and the SARIMA model. While SARIMA initially demonstrated better performance in capturing certain seasonal components, the Holt-Winters Method was ultimately selected as the final model due to its ability to effectively handle multiplicative seasonality and provide more stable forecasts for this dataset.

Objective

The main objective of this project is to develop a predictive model for forecasting the stock prices of ICICI Bank using the Holt-Winters Method. We aim to:

- Analyze historical stock price data to identify trends and seasonal patterns.

- Build a robust Holt-Winters model that captures the underlying structure of the time series data.
- Conduct a comparative analysis with the SARIMA model to validate model performance.
- Evaluate the model's accuracy in forecasting future stock prices.

Scope of the Project

This project focuses on time series forecasting of ICICI Bank stock prices using data from 2012 to 2024. We employ the Holt-Winters Method as the primary forecasting method due to its capability to model multiplicative seasonal patterns effectively. Additionally, a comparison is made with the SARIMA model, a widely used method for financial time series forecasting. The data used in this analysis includes monthly closing prices of ICICI Bank, sourced from publicly available financial datasets. The scope of the project is limited to statistical time series analysis, and we do not incorporate external factors such as economic indicators or market news in the model.

Outline of the Report

The remainder of this report is organized as follows:

- **Methodology:** This discusses the data collection process, preprocessing steps, and the models used for analysis, including the Holt-Winters Method and SARIMA.
- **Results:** The results of the analysis, including the comparative study, are presented in this, along with visualizations and performance metrics.
- **Future Work:** This suggests possible directions for future research and improvements to the model implemented in the paper.

Methodology

Data Collection

The dataset used in this project comprises the monthly closing prices of ICICI Bank stock, sourced from Yahoo Finance. The data covers the period from January 2012 to October 2024. The ticker symbol `ICICIBANK.NS` was used to fetch the data for ICICI Bank listed on the National Stock Exchange (NSE) of India.

The data includes the adjusted closing prices, which reflect the stock's market value after adjustments for corporate actions like stock splits and dividends. The time period chosen provides a comprehensive view of the stock's performance over more than a decade, allowing us to capture long-term trends and seasonal patterns.

Data Preprocessing

After downloading the raw stock price data, the following preprocessing steps were performed:

- **Resampling:** The daily stock prices were resampled to a monthly frequency using the last trading day of each month. This step reduces noise in the data and focuses on capturing long-term trends.
- **Feature Selection:** Only the closing prices were retained for the analysis, as they represent the final traded price of the stock for each period.
- **Handling Missing Values:** Any missing values in the dataset were dropped to ensure the continuity of the time series.

The first few rows of the processed dataset are shown below:

```
Date          Close
2012-01-31   164.027267
2012-02-29   164.781815
2012-03-31   161.854538
2012-04-30   160.427277
2012-05-31   142.409088
```

Exploratory Data Analysis

Exploratory Data Analysis (EDA) was performed to understand the underlying patterns in the ICICI Bank stock price data. The goal of EDA is to identify trends, seasonalities, and any potential outliers in the time series.

The first step in the analysis was to visualize the monthly closing prices of ICICI Bank. The plot below shows the time series of the stock prices from January 2012 to October 2024. This visualization helps in identifying any visible trends or cycles.

From the plot, we can observe the following:

- A general upward trend in stock prices over the years, suggesting long-term growth.

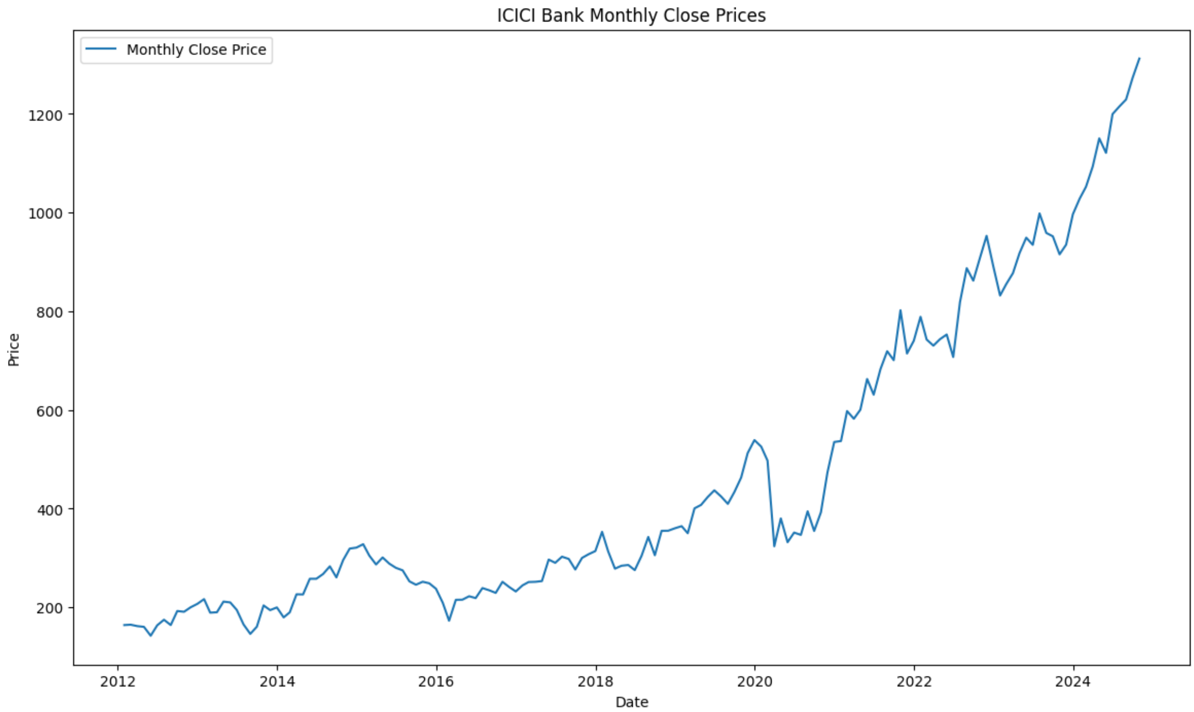

Figure 1: ICICI Bank Monthly Close Prices

- Periodic fluctuations, likely indicating market cycles, suggesting the presence of seasonality.
- A few sharp drops in prices, which could be indicative of market shocks or significant events affecting the bank's stock price.

This initial analysis guided the choice of the SARIMA model, which is well-suited for handling time series data with both trends and seasonal components.

One of the important steps in EDA was calculating the moving averages (MAs) for different window sizes (3 months, 6 months, and 12 months) to smooth the time series data and highlight underlying trends. The following plot shows the monthly closing prices of ICICI Bank along with the calculated moving averages.

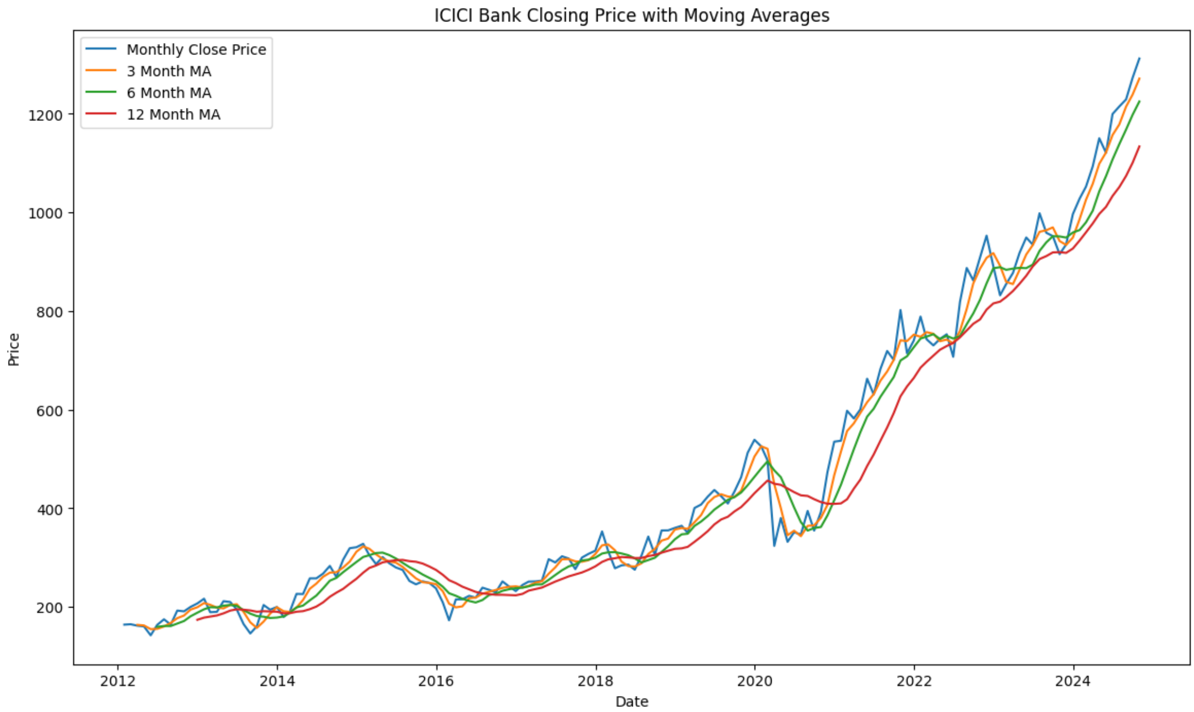

Figure 2: ICICI Bank Closing Price with Moving Averages

From the plot, we observe the following:

- The 3-month moving average (MA) closely follows the monthly closing prices but is less volatile, smoothing out the short-term fluctuations.
- The 6-month moving average provides a clearer indication of medium-term trends and shows a more stable pattern compared to the 3-month MA.
- The 12-month moving average captures the long-term trend of the stock price, providing a smoother curve that highlights the overall direction of the stock's movement.

The use of moving averages allows us to better understand the general trend in the stock price and is an essential step before applying forecasting models, such as SARIMA, which can model both trend and seasonality in the data.

After downloading and cleaning the data, the next step is to ensure that the time series is stationary, which is a necessary condition for many time series models, including ARIMA and SARIMA.

Testing stationarity

The stationarity of the data was tested using the **Augmented Dickey-Fuller(ADF) test**.
The null hypothesis of the ADF test is that the series is non-stationary, while the alternative hypothesis is that the series is stationary. A p-value less than or equal to 0.05 indicates that the series is stationary.
Initially, the ADF test was performed on the raw closing prices, and the results indicated that the series was non-stationary:

```
ADF Statistic: 2.1957653455033173
p-value: 0.9988790026277173
Series is Non-Stationary
```

Since the series was non-stationary, first differencing was applied to make the series stationary. After differencing, the ADF test was repeated on the differenced series, and the results showed that the series became stationary:

```
ADF Statistic: -13.774036912354227
p-value: 9.564034472009168e-26
Series is Stationary
```

The differenced series was used for further analysis and model building.

Seasonal Decomposition

Seasonal decomposition helps to break the data into these components, making it easier to understand any patterns or cycles in the stock price. To decompose the data, we used the 'seasonal_decompose' function with a multiplicative model, assuming that the seasonal effect scales with the trend. The decomposition was performed with a period of 12, corresponding to the monthly data, which assumes that the seasonality is annual (i.e., repeating every 12 months).

The seasonal decomposition of ICICI Bank's monthly closing prices is shown in the following plot:

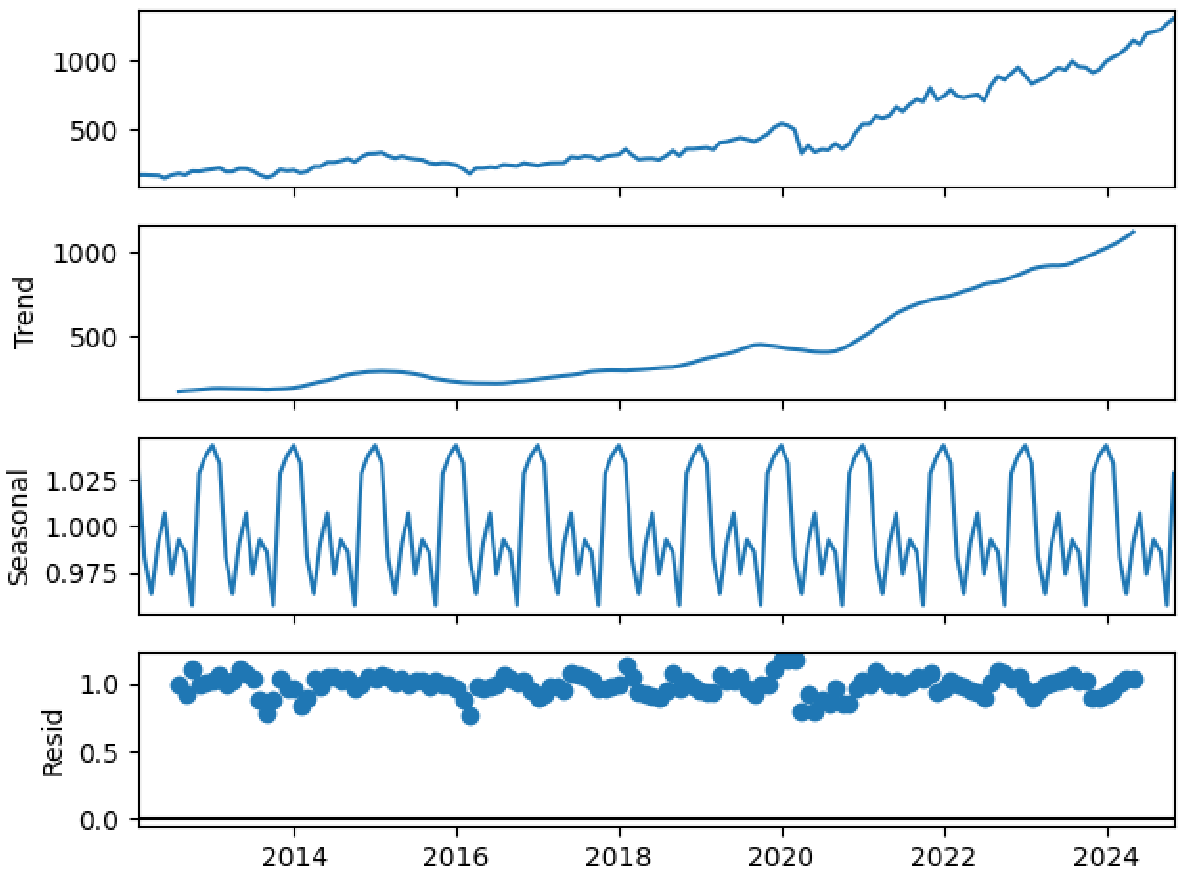

Figure 3: Seasonal Decomposition of ICICI Bank Monthly Close Prices

From the decomposition, we observe the following components:

- **Trend**: The overall upward trend in the stock price is visible in the trend component.
- **Seasonality**: The seasonal component captures the annual fluctuations in stock prices, which may correspond to specific market events, financial reports, or economic conditions.
- **Residuals**: The residual component represents the noise or random fluctuations after accounting for trend and seasonality.

Plotting ACF and PACF

In this section, we examine the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) of the time series before and after differencing to understand the underlying patterns and stationarity.

We begin by plotting the ACF and PACF for the original time series data. The ACF and PACF plots help us assess the autocorrelation and partial autocorrelation structure of the data, which can guide us in determining the parameters for the ARIMA model.

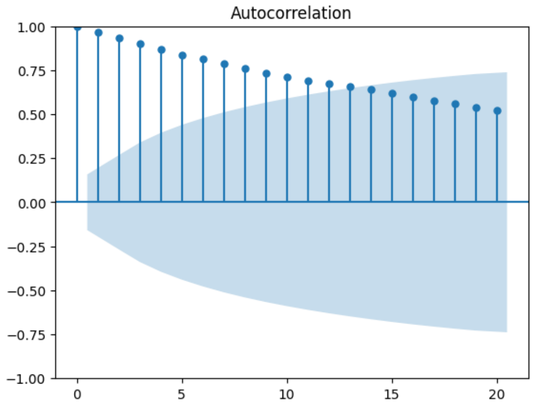

Figure 4: ACF of ICICI Bank Closing Prices (Before Differencing)

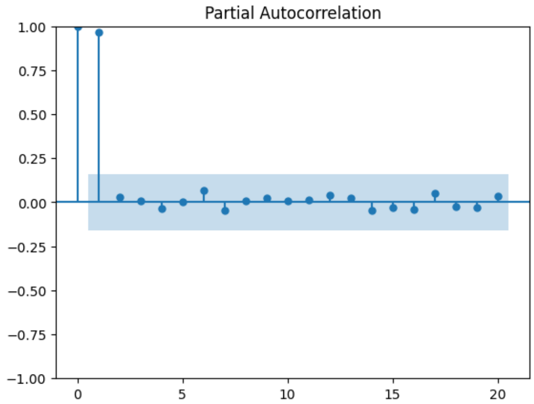

Figure 5: PACF of ICICI Bank Closing Prices (Before Differencing)

To make the series stationary, we perform first differencing on the data. Differencing removes trends and helps stabilize the mean, making the series more suitable for modeling. We then plot the ACF and PACF for the differenced series to observe the changes in the autocorrelation structure.

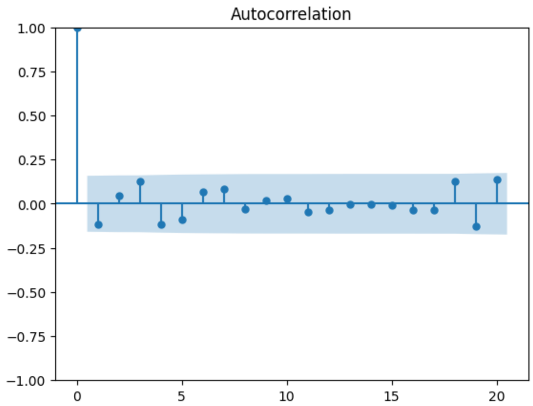

Figure 6: ACF of ICICI Bank Closing Prices (After Differencing)

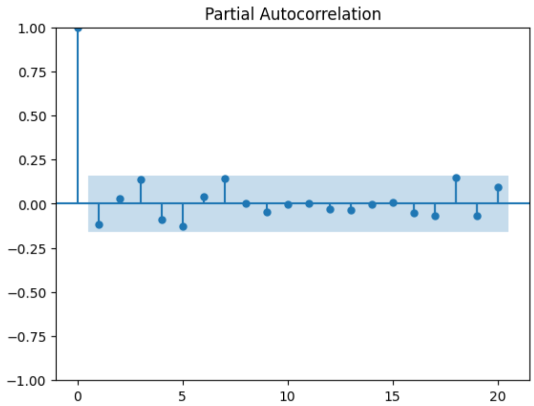

Figure 7: PACF of ICICI Bank Closing Prices (After Differencing)

Models Used

SARIMA Model

To identify the optimal parameters for the SARIMA model, an automatic grid search approach was implemented. Grid search is a method that exhaustively tries all possible combinations of model parameters within a specified range, and selects the set of parameters that minimizes the **Akaike Information Criterion (AIC)**, which is a measure of model fit and complexity.

The SARIMA model parameters consist of both non-seasonal and seasonal components. The non-seasonal parameters include the autoregressive order (p), differencing order (d), and moving average order (q), while the seasonal parameters include the seasonal autoregressive order (P), seasonal differencing order (D), seasonal moving average order (Q), and the seasonal period (S).

The grid search was conducted for a range of values for each parameter: - **p, d, q**: Ranges from 0 to 2 for the AR, differencing, and MA terms. - **P, D, Q**: Ranges from 0 to 2 for the seasonal AR, differencing, and MA terms, with a seasonal period of 12 (corresponding to monthly data with yearly seasonality).

The best combination of parameters obtained through the grid search was:
SARIMA Order:(2, 1, 2), **Seasonal Order:(0, 1, 2, 12)**, **AIC: 831.3167**

SARIMA Order (p, d, q)	**Seasonal Order (P, D, Q, s)**	**AIC**
(0, 1, 0)	(2, 1, 2, 12)	855.5350
(0, 1, 1)	(0, 1, 2, 12)	847.3011
(0, 1, 1)	(2, 1, 2, 12)	845.5294
(0, 1, 2)	(0, 1, 2, 12)	832.5362
(0, 1, 2)	(2, 1, 2, 12)	831.7073
(2, 1, 2)	(0, 1, 2, 12)	831.3168

Table 1: SARIMA Model Orders and AIC Values

The grid search attempts all possible combinations of parameters, and the combination that results in the lowest AIC is selected as the optimal set. In this study, the best SARIMA model was identified with the order (2, 1, 2) for the non-seasonal components, and (0, 1, 2, 12) for the seasonal components, resulting in the lowest AIC of 831.3167.

After identifying the optimal parameters for the SARIMA model, we use the model to generate forecasts on the test data. The forecasted values are compared with the actual test data to assess the performance of the model.

The SARIMA model was applied to forecast the ICICI Bank stock prices for the test period. The forecast was generated using the `predict` function of the SARIMA model, where the forecast starts from the end of the training data and continues for the length of the test data. The forecast was made without dynamic updates to the model parameters during the forecast period.

The following plot compares the actual stock prices from the test set with the forecasted values produced by the SARIMA model. The training data is also included in the plot for context.

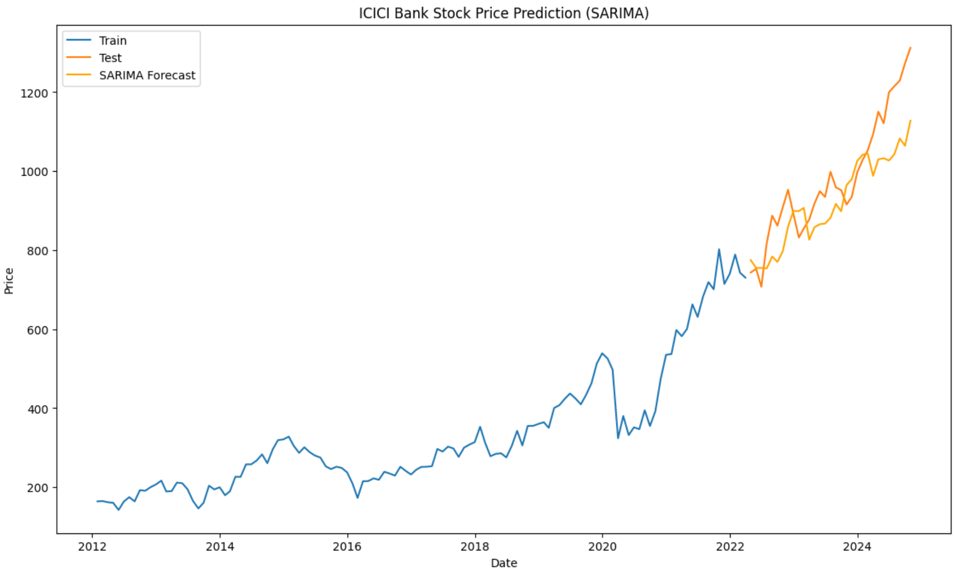

Figure 8: ICICI Bank Stock Price Prediction using SARIMA Model

Exponential Smoothing (Holt-Winters Method)

The Exponential Smoothing method is a popular technique for time series forecasting, particularly when the data exhibits both trend and seasonality. In this project, we applied the Holt-Winters method with multiplicative components to account for the varying seasonal effects in the ICICI Bank stock price data.

Model Description

The Holt-Winters model was configured with the following components:

- **Trend**: Multiplicative, to capture the changing trend in the time series.
- **Seasonality**: Multiplicative, to model the seasonal effects that increase or decrease proportionally with the trend.
- **Seasonal Periods**: 12, indicating monthly seasonality based on the yearly cycle of the stock prices.

The model was trained on the historical data, and the forecast was generated for the test period. The results are visualized in the following plot.

Forecast Plot

The plot below shows the actual stock prices for the training and test sets, as well as the forecasted values using the Exponential Smoothing model with multiplicative trend and seasonality.

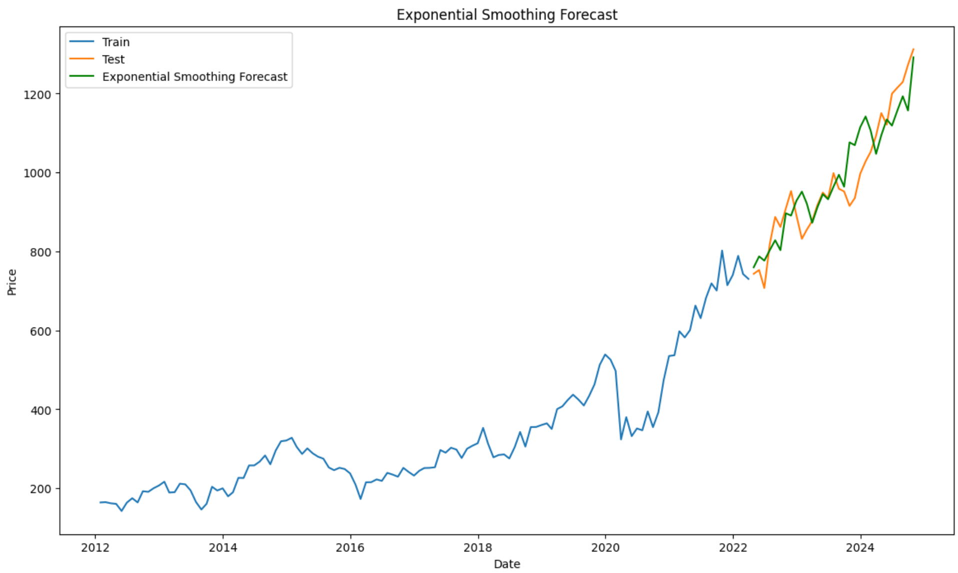

Figure 9: ICICI Bank Stock Price Prediction using Exponential Smoothing (Holt-Winters Multiplicative Model)

Model Interpretation

The multiplicative components of the Holt-Winters method were chosen due to the proportional changes in trend and seasonality observed in the data. This model type is well-suited for financial time series where the magnitude of changes varies over time, providing a more accurate fit compared to the additive model.

The forecasted values align closely with the actual stock prices, indicating the effectiveness of the model in capturing both the trend and seasonal patterns of the ICICI Bank stock prices.

Model Evaluation

Evaluation Metrics

To evaluate the forecasting performance of the models, we used the Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). These metrics are defined as follows:

- **Root Mean Squared Error (RMSE)**: Measures the standard deviation of the residuals (prediction errors). It gives an idea of how spread out the residuals are and is sensitive to large errors.
- **Mean Absolute Error (MAE)**: Represents the average absolute difference between the actual and predicted values, providing a straightforward measure of prediction accuracy.

The calculated RMSE and MAE for both models are shown below:

Model	**RMSE**	**MAE**
SARIMA	96.45	80.30
Exponential Smoothing	68.26	53.49

Table 2: Comparison of RMSE and MAE for SARIMA and Exponential Smoothing Models

Interpretation

From Table 2, it is evident that the Exponential Smoothing model outperformed the SARIMA model, achieving lower values for both RMSE and MAE. This indicates that the Exponential Smoothing model provided more accurate forecasts for the ICICI Bank stock prices in this dataset. The superior performance of Exponential Smoothing may be attributed to its ability to capture the multiplicative seasonal effects present in the data.

Future Work

This project lays a strong foundation for stock price forecasting using time series models. However, there are several avenues for future research and improvements that can be explored:

1. **Incorporating Exogenous Variables (External Factors)**: Future studies could enhance the model by including exogenous variables such as macroeconomic indicators (e.g., interest rates, inflation, GDP growth), company financial metrics (e.g., earnings reports), or market sentiment indicators (e.g., news sentiment, social media analysis). This could be implemented using models like SARIMAX, which extend SARIMA by incorporating external factors.

2. **Application of Machine Learning Models**: While traditional statistical models like SARIMA and Holt-Winters are effective, machine learning approaches such as Long Short-Term Memory (LSTM) networks, Prophet, and Random Forest Regression could be explored for capturing non-linear patterns. A hybrid approach combining machine learning with traditional time series methods might yield improved accuracy.

3. **Handling Volatility and Structural Breaks**: Financial time series often exhibit high volatility and structural breaks (e.g., market crashes or major news events). Future research could focus on integrating models that address these aspects, such as GARCH models (Generalized Autoregressive Conditional Heteroskedasticity) or change-point detection algorithms.

These suggestions aim to enhance the robustness, accuracy, and practical utility of the stock price forecasting model, providing a strong foundation for future research and potential applications in real-world financial markets.