Time Series Analysis of Stock Closing Data

Submitted by: Pragya Gupta

PGDM – Research & Business Analytics

Abstract

This project focuses on forecasting the weekly closing prices of the S&P 500 index using different statistical and smoothing models such as Naive Forecast, Moving Averages, Exponential Smoothing methods, and ARIMA. The primary goal is to compare these models, evaluate their accuracy, and identify the most suitable approach for reliable stock market forecasting. The analysis revealed that the Triple Exponential Smoothing (Holt-Winters) model performed best, effectively capturing both trend and seasonality.

1. Introduction

Forecasting stock prices is critical in financial analytics as it aids investors and analysts in making informed decisions. Stock prices often exhibit patterns such as trends, cycles, and seasonality, making time series analysis a powerful tool. This project applies classical statistical forecasting techniques to analyze and predict the S&P 500 weekly closing prices and evaluates their performance.

2. Dataset Description

- Source: Yahoo Finance

- Stock: S&P 500 Index (^GSPC)
- Variable: Weekly closing price
- Frequency: Weekly observations
- Period: Multi-year historical data

The dataset was pre-processed by cleaning missing values, setting the date index, and visualizing trends and autocorrelations using line plots, ACF, and PACF.

3. Methodology

The following forecasting models were implemented and compared:

- Naive Forecast – baseline model using the last observed value.
- Simple Moving Average (SMA) – smooths short-term fluctuations.
- Weighted Moving Average (WMA) – assigns higher weights to recent data points.
- Simple Exponential Smoothing (SES) – incorporates exponential weighting.
- Double Exponential Smoothing (DES / Holt's Trend Method) – accounts for trends.
- Triple Exponential Smoothing (TES / Holt-Winters) – models both trend and seasonality.

- ARIMA (1,1,1) – autoregressive integrated moving average model for autocorrelation patterns.

Evaluation Metrics

- Mean Squared Error (MSE)
- Root Mean Squared Error (RMSE)
- Mean Absolute Percentage Error (MAPE)

Residual diagnostics were also performed to validate assumptions.

4. Results & Evaluation

- Naive Forecast & Moving Averages: Served as simple benchmarks but lacked predictive power.
- SES & DES: Improved fit but failed to capture seasonality.
- TES (Holt-Winters): Outperformed others by modeling both trend and weekly seasonality, providing the lowest error metrics.
- ARIMA (1,1,1): Competitive accuracy, but residual analysis indicated issues with non-normality and heteroskedasticity.

Best Model: Triple Exponential Smoothing (Holt-Winters)

- Effectively captured seasonality and trend.
- Produced stable, interpretable forecasts.
- Achieved the lowest RMSE and MAPE among all tested models.

5. Conclusion & Insights

This project demonstrated that classical time series models can effectively forecast stock prices when patterns are present in the data. While ARIMA captured autocorrelation well, the Holt-Winters method proved most effective for weekly S&P 500 forecasts due to its ability to model both trend and seasonality.

Future Work

- Implementation of SARIMA for enhanced seasonality.
- ARCH/GARCH models to capture volatility clustering.
- Exploration of machine learning models (LSTM, Random Forest, XGBoost) for nonlinear forecasting.