Reading Report 3

Stock price forecasting using data from Yahoo finance and analysing seasonal and nonseasonal trend

Introduction

- Need:

  Accurate stock price prediction is challenging due to the diverse seasonal and nonseasonal trends they display. Most investors make high-risk decisions because they overlook patterns and trends, relying instead on historical price data.

- Goal:

  The objective is to reduce risks and produce a strong forecast range by using time series models

- Insight:

  Advanced forecasting techniques are necessary because traditional models often fail to account for dynamic phenomena like volatility clustering and autocorrelation structures.

Data Overview

- Data Source:
  - Yahoo Finance – Monthly stock prices of Apple Inc. from January 2000 to January 2018
- Data Characteristics:
  - Open, high, low, close, adjusted close prices, and trading volume are essential for analyzing stock trends and volatility, with close price being the key for forecasting models
- Preprocessing:
  - Handling missing values, applying log transformation for variance stabilization, and splitting data into training (2000-2016) and test (2016-2018) sets

ARIMA vs Holt-Winters Model

Criteria	Arima Model	Holt-Winters Model
Trend Type	Non-Seasonal	Seasonal
Model	Autoregression, differencing and moving averages	Triple exponential smoothing to model level
Parameter Tuning	ACF and PACF plots	Alpha, beta and gamma are fine-tuned
Dependencies	Long term dependencies	Short term dependencies
Model Fit	ARIMA(2,1,0) selected based on minimum AIC/BIC scores for best predictive performance	Applied to training data, effectively capturing periodic stock price fluctuations with high accuracy
Specifications	Applied to address non-stationarity by stabilizing the mean	Identified with a 12-month cycle based on the monthly stock data

Results and Observations

- Seasonal Patterns in Apple Stock:

  Apple stock prices exhibit strong seasonal effects, characterized by consistent peaks at the end of each year.

- Forecast Uncertainty:

  As the forecast horizon extends, prediction intervals widen due to increased uncertainty, reflecting the challenges of long-term forecasting.

- Model Validation:

  The accuracy of the model was confirmed using test data, resulting in a prediction interval with a confidence level of 95%.

- Combined Forecast Benefits:

  The combined forecast indicates that price fluctuations within the predicted range are safer for investors, especially during periods of high volatility.

- Optimal Forecasting Approach:

  The combination of ARIMA and Holt-Winters models provided an optimal price range for stock investments, offering a robust forecasting strategy

Conclusions

- Limitations of Current Models:

  The ARIMA and Holt-Winters models have limitations, as they assume previous price patterns will continue, potentially failing to account for sudden market changes.

- Future Directions: External Inputs:

  To enhance the predictive capability of the model, future research could explore incorporating external factors such as macroeconomic data and news sentiment analysis.

- Future Directions: Machine Learning:

  Future studies may utilize machine learning models, like LSTM, to capture complex, nonlinear patterns in stock prices, potentially improving forecast accuracy.

- Effectiveness of Current Approach:

  This study effectively demonstrated the benefits of integrating ARIMA and Holt-Winters models for stock price forecasting.

- Potential for Model Enhancement:

  By addressing the limitations and exploring new methodologies, future research can further refine and improve the forecasting capabilities of these models.

Quiz Questions

- **Question:**

  How do ACF and PACF differ in their roles when specifying the parameters of an ARIMA model?

- **Answer:**
  - In the process of determining ARIMA model parameters, the Autocorrelation Function and the Partial Autocorrelation Function serve complementary roles. The ACF provides insight into how past values collectively influence current values, which is crucial for identifying the moving average component. In contrast, the PACF isolates direct correlations between data points, effectively filtering out the impact of intermediate values. This makes it essential for determining the autoregressive component. By analyzing these functions, researchers can make informed decisions about the ARIMA parameters, leading to more accurate model specifications and improved forecasting capabilities.