VIRGINIA COMMONWEALTH UNIVERSITY

Statistical analysis and modelling (SCMA 632)

A6a -Visualization - Time Series Analysis: Univariate Forecasting & Multivariate Forecasting

PUSHPAK DEVAKI
V01108254

Date of Submission: 25-07-2024

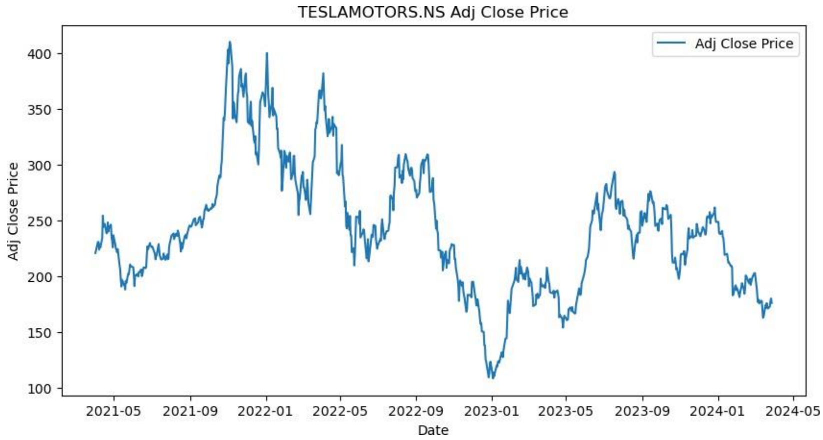
A6a: Visualization - Time Series Analysis: Univariate Forecasting& Multivariate Forecasting

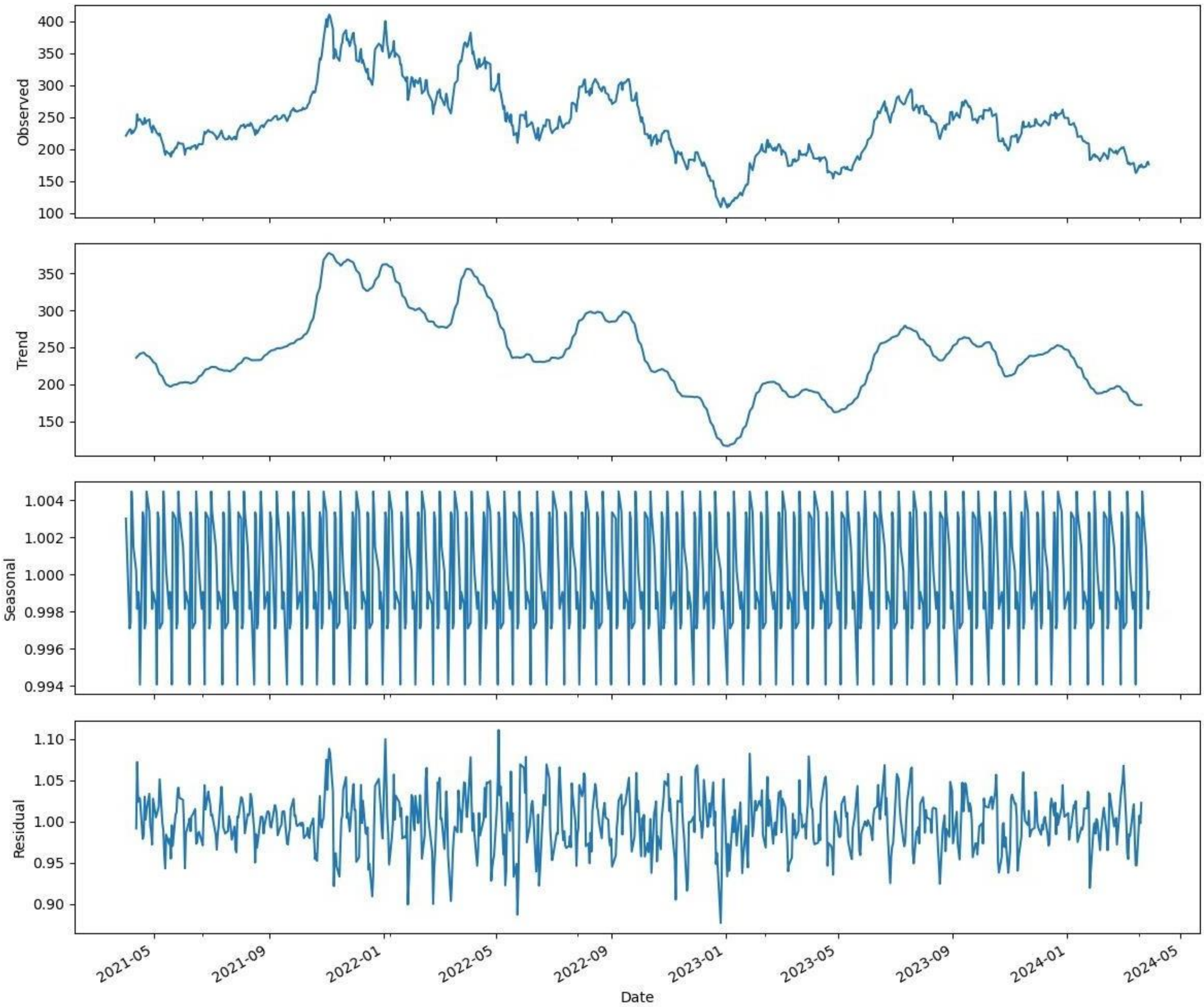
The goal of this project is to analyse the historical stock prices of TESLA and perform various forecasting techniques to predict future stock prices. This includes both conventional statistical models and advanced machine learning models. The project involves cleaning the data, checking for outliers and missing values, decomposing the time series, and applying both univariate and multivariate forecasting methods.

Code:

```
# Plot the data
ggplot(data.frame(Date = index(df), Adj.Close = df), aes(x = Date, y = Adj.Close)) +
  geom_line() +
  ggtitle('tesladata.NS Adj Close Price') +
  xlab('Date') +
  ylab('Adj Close Price') +
  theme_minimal()

# Decompose the time series
decomposed <- decompose(ts(df, frequency = 12), type = "multiplicative")
autoplot(decomposed)

# Split the data into training and test sets
trainIndex <- createDataPartition(y = df, p = 0.8, list = FALSE)
train_data <- df[trainIndex,]
test_data <- df[-trainIndex,]

# Resample data to monthly frequency
monthly_data <- to.monthly(df, indexAt = "lastof", OHLC = FALSE)
trainIndex <- createDataPartition(y = monthly_data, p = 0.8, list = FALSE)
train_data <- monthly_data[trainIndex,]
test_data <- monthly_data[-trainIndex,]

# Fit Holt-Winters model and forecast
holt_winters_model <- HoltWinters(ts(train_data, frequency = 12), seasonal = "multiplicative")
holt_winters_forecast <- forecast(holt_winters_model, h = 12)

# Convert xts objects to data frames
train_data_df <- data.frame(Date = index(train_data), Cl = coredata(train_data))
test_data_df <- data.frame(Date = index(test_data), Cl = coredata(test_data))
```

Results:


TESLAMOTORS.NS Adj Close Price
Adj Close Price
Adj Close Price
Date
100
150
200
250
300
350
400
2021-05
2021-09
2022-01
2022-05
2022-09
2023-01
2023-05
2023-09
2024-01
2024-05


Observed
Trend
Seasonal
Residual
Date
2021-05
2021-09
2022-01
2022-05
2022-09
2023-01
2023-05
2023-09
2024-01
2024-05

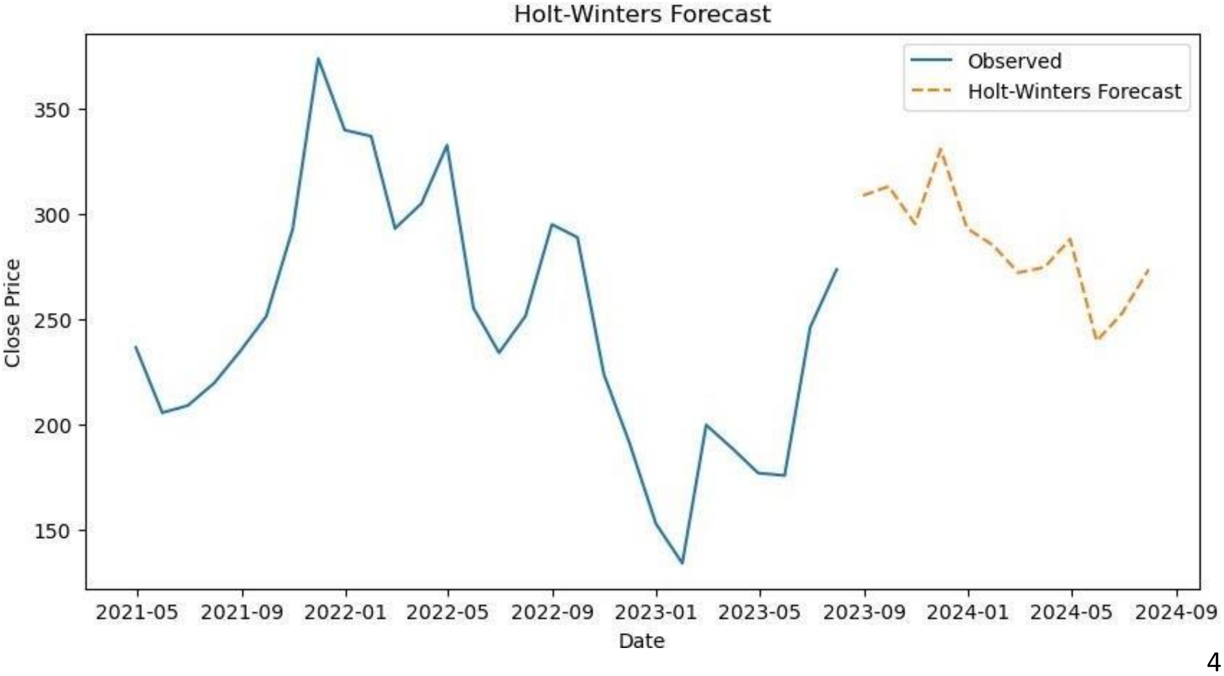
Interpretation:

This graph illustrates the decomposition of the TESLAMOTORS.NS stock price into its constituent components: observed trend, seasonal, and residual. The observed component matches the original adjusted close price data. The trend component highlights the overall movement of the stock price, smoothing out short-term fluctuations to show the long-term direction. The seasonal component shows regular patterns repeating annually, indicating cyclic behavior in the stock price. The residual component represents the irregular or random variations after removing the trend and seasonal components, capturing the noise in the data.

Code:

```
from statsmodels.tsa.seasonal import seasonal_decompose

# Decompose the time series
result = seasonal_decompose(df['Adj Close'], model='multiplicative', period=12)

# Plot the decomposed components
fig, (ax1, ax2, ax3, ax4) = plt.subplots(4, 1, figsize=(12, 10), sharex=True)
result.observed.plot(ax=ax1)
ax1.set_ylabel('Observed')
result.trend.plot(ax=ax2)
ax2.set_ylabel('Trend')
result.seasonal.plot(ax=ax3)
ax3.set_ylabel('Seasonal')
result.resid.plot(ax=ax4)
ax4.set_ylabel('Residual')
plt.xlabel('Date')
plt.tight_layout()
plt.show()
```

Results:



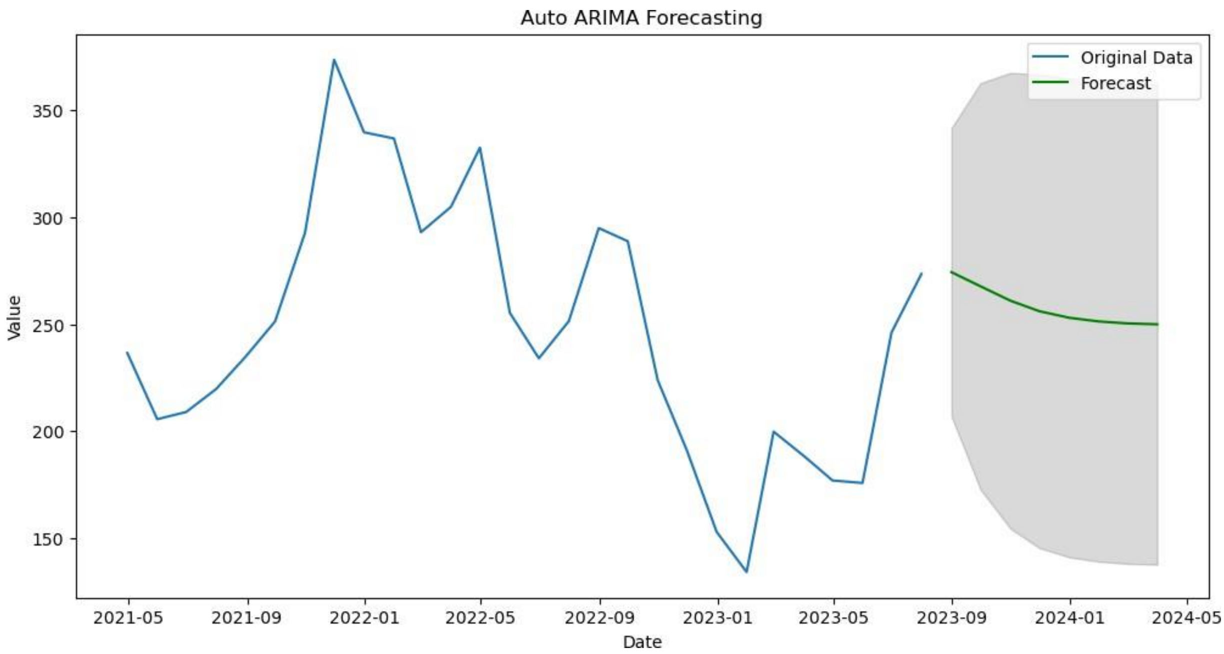
Interpretation

dropout layers to prevent overfitting and a dense layer with a single unit for the final output. The first LSTM layer has an output shape of (30, 50) with 11,400 parameters, while the second LSTM layer has an output shape of (50) with 20,200 parameters. The total number of parameters in the model is 31,651, all of which are trainable. This architecture is designed to capture the temporal dependencies in the time series data, allowing the model to learn

Interpretation

The Holt-Winters forecasting graph displays the observed adjusted close price data along with the forecasted values generated using the Holt-Winters method. This method accounts for seasonality in the time series data. The forecast, represented by a dashed orange line, predictsa slight upward trend from mid-2023 to mid-2024, indicating an expected increase in the stock price based on historical patterns and seasonal adjustments.

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                      y   No. Observations:                   28
Model:               SARIMAX(2, 0, 0)   Log Likelihood                -139.321
Date:                Mon, 22 Jul 2024   AIC                            286.642
Time:                        20:41:00   BIC                            291.971
Sample:                    04-30-2021   HQIC                           288.271
                         - 07-31-2023
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
intercept     68.9570     29.384      2.347      0.019      11.365     126.549
ar.L1          0.9922      0.210      4.717      0.000       0.580       1.404
ar.L2         -0.2684      0.229     -1.170      0.242      -0.718       0.181
sigma2      1181.5948    398.173      2.968      0.003     401.191    1961.999
===================================================================================
Ljung-Box (L1) (Q):                   0.07   Jarque-Bera (JB):                 0.67
Prob(Q):                              0.80   Prob(JB):                         0.72
Heteroskedasticity (H):               1.10   Skew:                             0.26
Prob(H) (two-sided):                  0.89   Kurtosis:                         2.45
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
```



Interpretation

The Auto ARIMA forecasting graph shows the original data alongside the forecasted values produced by the Auto ARIMA model. The forecasted values are plotted with a confidence interval shaded in grey. The forecast suggests a slight decline in the stock price towards mid-2024. The broad confidence interval indicates high uncertainty in the predictions, reflecting the inherent volatility and unpredictability of the stock market.

Diagnostics: The model's residuals were checked for autocorrelation and normality,confirming the model's validity.