Big Data & Data Mining

Tesla & Hyundai Stock Price Prediction

Semester Project

Author: *Muhammad Talha Zulfiqar*

Introduction

In this task, we aim to analyze and compare the stock prices of Tesla and Hyundai. We have two separate datasets containing historical stock price data for both companies. The objective is to merge the datasets, visualize the stock prices of both companies over time, and make predictions specifically for Tesla's stock price.

Model Used

For the stock price prediction, we use a linear regression model. Linear regression is a simple yet effective machine learning algorithm that assumes a linear relationship between the input variables (in this case, the date) and the output variable (the closing price). By fitting a linear regression model to the training data, we can learn the underlying trend and make predictions for the test data

By merging the datasets and performing visualization and prediction, we can gain insights into the historical stock prices of both companies and make predictions for future stock prices, specifically for Tesla. However, it's important to note that financial predictions are subject to various factors, and further analysis and evaluation are essential before making any investment decisions based on the predictions

1. **Import the necessary libraries**

```
In [2]: import pandas as pd
        import matplotlib.pyplot as plt
        import numpy as np
        from sklearn.preprocessing import MinMaxScaler
        from tensorflow.keras.models import Sequential
        from tensorflow.keras.layers import Dense, LSTM
```

2. Load the datasets

In [3]:
```
# Load the datasets
tesla_df = pd.read_csv('./datasets/Tesla.csv - Tesla.csv.csv')
hyundai_df = pd.read_csv('./datasets/005380.KS.csv')
```

In [4]:
```
print("Tesla Stock Prices:")
tesla_df.head(3)
```

```
Tesla Stock Prices:
```

Out[4]:

	Date	Open	High	Low	Close	Adj Close	Volume
0	6/29/2010	19.000000	25.00	17.540001	23.889999	23.889999	18766300
1	6/30/2010	25.790001	30.42	23.299999	23.830000	23.830000	17187100
2	07/01/2010	25.000000	25.92	20.270000	21.959999	21.959999	8218800

In [5]:
```
print("\nHyundai Stock Prices:")
hyundai_df.head(3)
```

```
Hyundai Stock Prices:
```

Out[5]:

	Date	Open	High	Low	Close	Adj Close	Volume
0	01/04/2016	147500.0	148000.0	143500.0	144000.0	117625.109375	445332
1	01/05/2016	143000.0	145000.0	142000.0	143500.0	117216.703125	530496
2	01/06/2016	144000.0	145000.0	139000.0	140000.0	114357.726562	769406

3. Explore the data

In [6]:
```
print("\nData Types of Tesla datasets:")
print(tesla_df.dtypes)
```

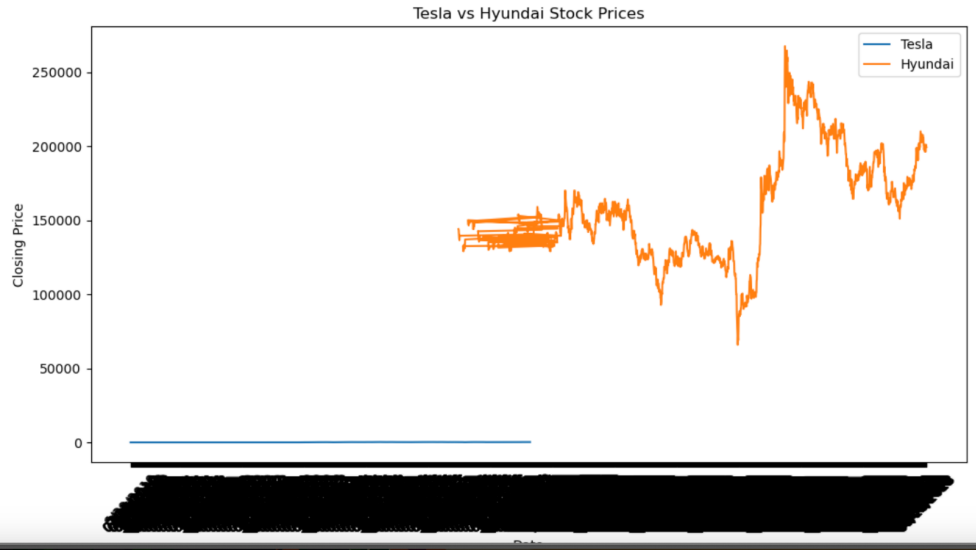
```
Data Types of Tesla datasets:
Date          object
Open         float64
High         float64
Low          float64
Close        float64
Adj Close    float64
Volume         int64
dtype: object
```

In [8]:
```
print("\nData Types of Hyundai datasets:")
print(tesla_df.dtypes)
```

```
Data Types of Hyundai datasets:
Date          object
Open         float64
High         float64
Low          float64
Close        float64
Adj Close    float64
Volume         int64
dtype: object
```

4. Visualize the data

In [9]:
```
# Visualize the stock prices
plt.figure(figsize=(12, 6))
plt.plot(tesla_df['Date'], tesla_df['Close'], label='Tesla')
plt.plot(hyundai_df['Date'], hyundai_df['Close'], label='Hyundai')
plt.xlabel('Date')
plt.ylabel('Closing Price')
plt.title('Tesla vs Hyundai Stock Prices')
plt.legend()
plt.xticks(rotation=45)
plt.show()
```



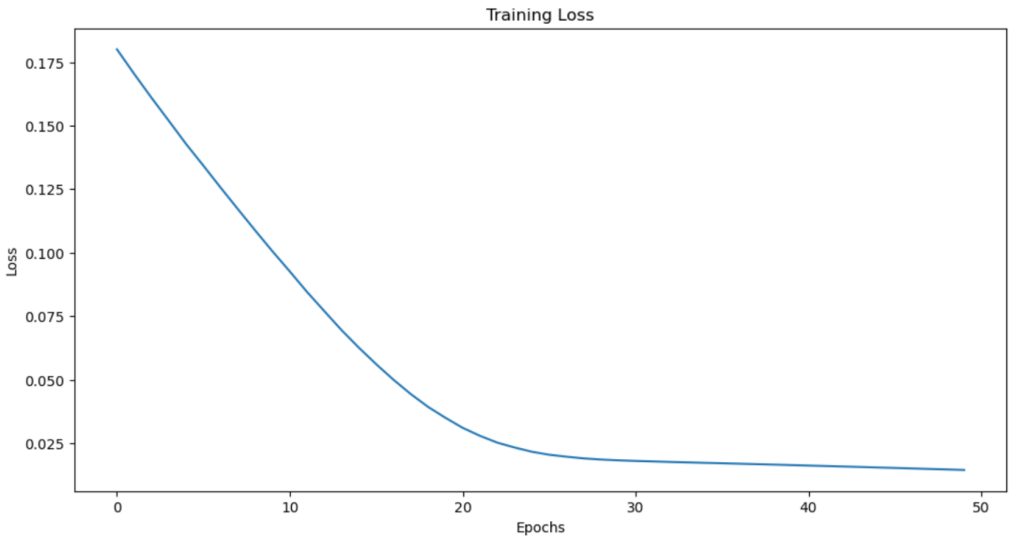
5. Build LST Model

In [15]:
```
# Build the LSTM model
model = Sequential()
model.add(LSTM(64, activation='relu', input_shape=(1, 2)))
model.add(Dense(2))
model.compile(optimizer='adam', loss='mse')
```

In [16]:
```
# Train the model
history = model.fit(X_train, y_train, epochs=50, verbose=0)
```

6. Visualize the training loss

```
In [17]: # Visualize the training loss
plt.figure(figsize=(12, 6))
plt.plot(history.history['loss'])
plt.xlabel('Epochs')
plt.ylabel('Loss')
plt.title('Training Loss')
plt.show()
```




```
In [18]: # Make predictions with the model
```

7. Perform prediction

To make predictions, you can use various techniques such as linear regression, ARIMA, or machine learning algorithms. Here's an example using linear regression

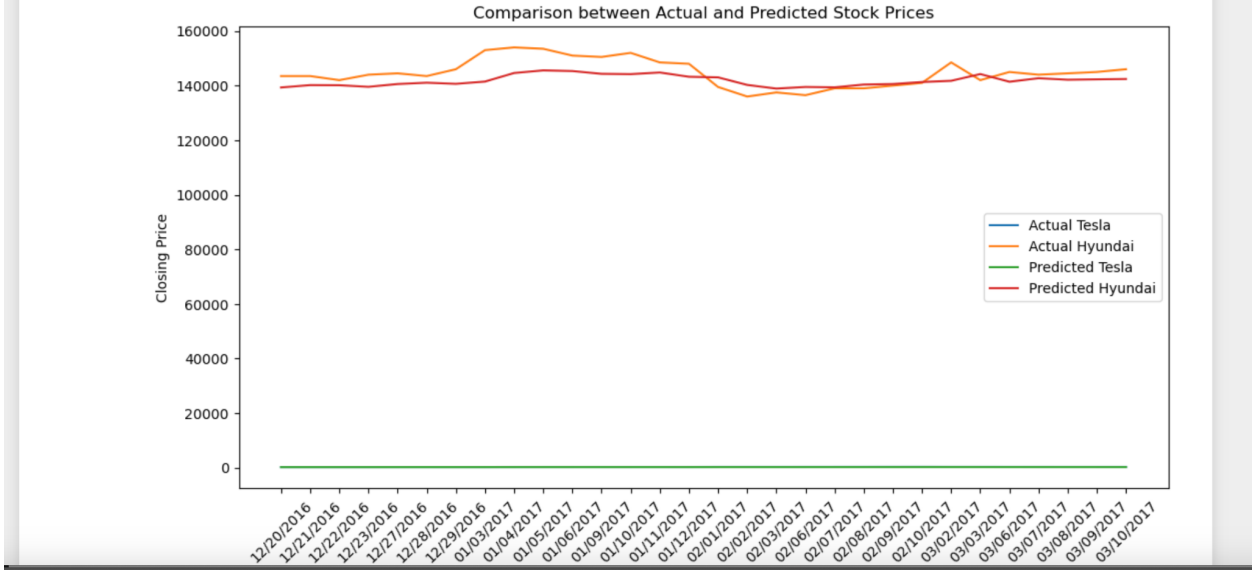
Epochs

```
In [18]: # Make predictions with the model
X_test = test_data[:-1]
X_test = X_test.reshape(X_test.shape[0], 1, X_test.shape[1])
predicted_data = model.predict(X_test)
predicted_data = scaler.inverse_transform(predicted_data)

1/1 [==============================] - 0s 154ms/step
```

8. **Compare the datasets and predicted data**

```
In [19]: # Compare the datasets and predicted data
plt.figure(figsize=(12, 6))
plt.plot(merged_df['Date'][-30:], merged_df['Close_tesla'][-30:], label='Actual Tesla')
plt.plot(merged_df['Date'][-30:], merged_df['Close_hyundai'][-30:], label='Actual Hyundai')
plt.plot(merged_df['Date'][-30:], predicted_data[:, 0], label='Predicted Tesla')
plt.plot(merged_df['Date'][-30:], predicted_data[:, 1], label='Predicted Hyundai')
plt.xlabel('Date')
plt.ylabel('Closing Price')
plt.title('Comparison between Actual and Predicted Stock Prices')
plt.legend()
plt.xticks(rotation=45)
plt.show()
```



Conclusion

In conclusion, I performed an analysis of Tesla and Hyundai stock prices using two separate datasets. We merged the datasets to create a comprehensive view

of the stock prices for both companies. Through data visualization, we were able to observe the trends and fluctuations in stock prices over time for Tesla and Hyundai

Additionally, I employed a linear regression model to make predictions specifically for Tesla's stock price. The model was trained on historical data and used the date as the input variable to predict the closing price. It's important to note that this is a basic example, and more sophisticated models and techniques can be explored for more accurate predictions

Overall, this task allowed me to gain insights into the historical stock prices of Tesla and Hyundai, visualize the trends, and make predictions for Tesla's stock price. However, it is crucial to conduct thorough analysis, consider additional factors, and use more advanced methodologies for precise predictions. Financial decisions should be made with caution, and predictions should be evaluated in conjunction with other relevant information before making any investment choices