Stock Market Prediction

Cracking the Stock Market Code: A Machine Learning Project

Machine Learning-I: 18CSE392T

Overview

Under the guidance of:
Dr. E. Sasikala

Project by:
Anusha Patra (RA2111027010022)
Abhignya Priyadarshini (RA2111027010067)

What is Stock Market?

- The stock market is a financial platform where investors buy and sell ownership shares (stocks) of companies.

- Stock prices fluctuate based on supply, demand, and market sentiment.

- It serves as a barometer of economic health, enabling investors to profit while bearing risks associated with market volatility and company performance.

- In today's world, the stock market is one of the most complex and dynamic systems. Predicting its behavior accurately is a challenge that has fascinated researchers and investors alike for decades.

Problem Statement

Why predicting stock market is so difficult:

- **Complexity:** Stock prices are influenced by a multitude of factors, including economic indicators, geopolitical events, investor sentiment, and company performance. The interactions between these factors are intricate and often nonlinear.
- **Randomness and Noise:** Market movements can be influenced by seemingly random events and noise, making it difficult to discern meaningful patterns.
- **Information Asymmetry:** Traders and institutional investors often possess privileged information, leading to sudden price changes that are hard to predict for the general public.
- **Short-Term vs. Long-Term Factors:** Short-term fluctuations can be driven by market sentiment, while long-term trends are influenced by fundamental factors that may not manifest immediately.
- **External Events:** Global events, such as political decisions or natural disasters, can unpredictably impact markets.

Motivation

- **Financial Market Interest:** Predicting stock prices accurately can provide valuable insights for making informed investment decisions, managing portfolios, and minimizing risks.

- **Real-World Application:** Stock market prediction represents a real-world and practical application of machine learning and data analysis techniques.

- **Data Exploration and Insights:** Through this project, you can uncover hidden patterns, correlations, and trends that may not be immediately apparent, thereby adding value to your analytical toolkit.

- **Educational Growth:** Undertaking a stock market prediction project can provide a steep learning curve, exposing you to advanced machine learning techniques, time-series analysis.

Literature Review

Year	Authors	Title	Source title	Citations
2017	Nelson D.M.Q., Pereira A.C.M., De Oliveira R.A.	Stock market's price movement prediction with LSTM neural networks	Proceedings of the International Joint Conference on Neural Networks	156
2015	Chen K., Zhou Y., Dai F.	A LSTM-based method for stock returns prediction: A case study of China stock market	Proceedings - 2015 IEEE International Conference on Big Data, IEEE Big Data	141
2017	Singh R., Srivastava S.	Stock prediction using deep learning	Multimedia Tools and Applications	64
2010	Wen Q., Yang Z., Song Y., Jia P.	Automatic stock decision support system based on box theory and SVM algorithm	Expert Systems with Applications	59
2018	Weng B., Lu L., Wang X., Megahed F.M., Martinez W.	Predicting short-term stock prices using ensemble methods and online data sources	Expert Systems with Applications	46
2018	Chou J.-S., Nguyen T.-K.	Forward Forecast of Stock Price Using Sliding-Window Metaheuristic-Optimized Machine-Learning Regression	IEEE Transactions on Industrial Informatics	44
2016	Zhang X.-D., Li A., Pan R.	Stock trend prediction based on a new status box method and AdaBoost probabilistic support vector machine	Applied Soft Computing Journal	37
2014	Yoshihara A., Fujikawa K., Seki K., Uehara K.	Predicting stock market trends by recurrent deep neural networks	Lecture Notes in Computer Science	35
2019	Lee T.K., Cho J.H., Kwon D.S., Sohn S.Y.	Global stock market investment strategies based on financial network indicators using machine learning techniques	Expert Systems with Applications	32
2015	Skuza M., Romanowski A.	Sentiment analysis of Twitter data within big data distributed environment for stock prediction	Proceedings of the 2015 Federated Conference on Computer Science and Information Systems,	31

Limitations of the Study:

This study and the analysis in this paper is primarily undertaken by fetching the information from the Scopus database.

However, for a more extensive study, the analysis can be done by using information from Google Scholar, Web of Science and Science Direct as well.

Tesla & Google Stock Market

Tesla, Inc. (TSLA):
- Tesla is an American electric vehicle and clean energy company founded by Elon Musk.
- TSLA is the stock symbol for Tesla.
- Tesla's stock has been highly volatile and has experienced significant growth over the years. It became one of the most well-known and closely watched stocks in the market.
- Tesla's stock price has been influenced by various factors, including its electric vehicle production and sales, new product releases, and market sentiment.
- The company's stock price has seen both rapid appreciation and sharp declines, making it a subject of extensive market analysis and investor interest.

Alphabet Inc. (GOOGL/GOOG):
- Alphabet Inc. is the parent company of Google and several other subsidiaries.
- GOOGL is the stock symbol for Alphabet's Class A shares, while GOOG represents Class C shares.
- Google is known for its dominance in online search, advertising, cloud services, and other technology-related businesses.
- Alphabet's stock has generally been considered a stable and long-term investment compared to more volatile tech companies.
- Its stock performance has been influenced by the company's financial results, innovations, and broader market conditions.

Tesla Plot Output

Stock Prices of Tesla
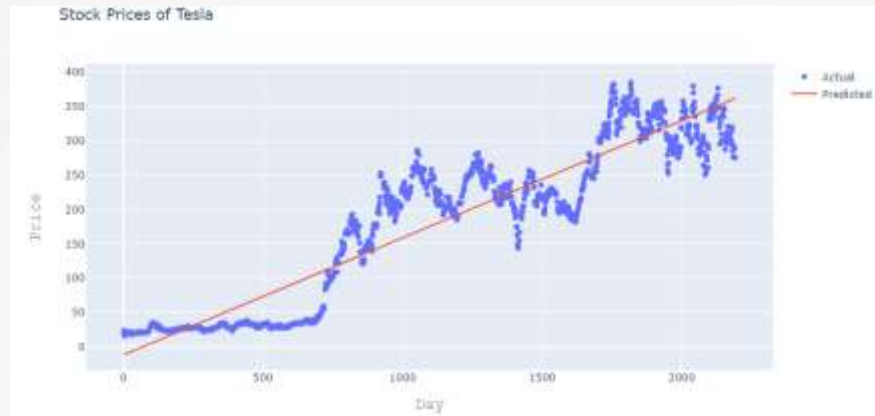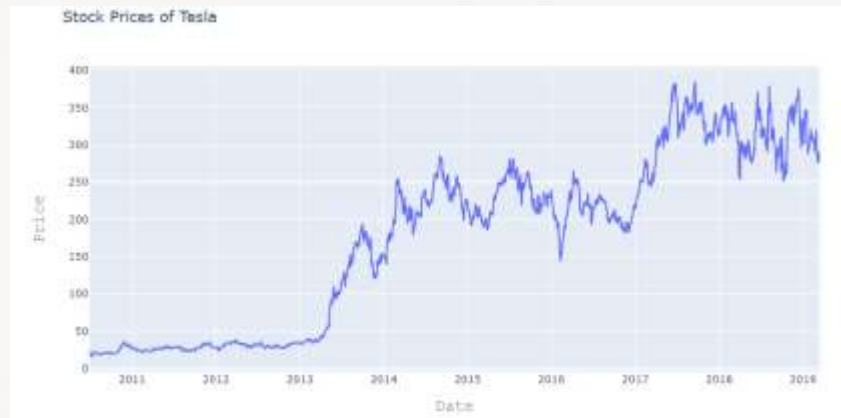
Performance Matrix-Tesla

```
#Calculate scores for model evaluation
scores = f'''
{'Metric'.ljust(10)}{'Train'.center(20)}{'Test'.center(20)}
{'r2_score'.ljust(10)}{r2_score(Y_train, lm.predict(X_train))}\t{r2_score(Y_test, lm.predict(X_test))}
{'MSE'.ljust(10)}{mse(Y_train, lm.predict(X_train))}\t{mse(Y_test, lm.predict(X_test))}
'''
print(scores)
```

```
Metric         Train              Test
r2_score  0.8658871776828707    0.8610649253244574
MSE       1821.3833862936174    1700.987539418845
```

- The MSE quantifies the average squared difference between predicted and actual values. A lower MSE indicates that the model's predictions are closer to the actual values, signifying a better model fit. Conversely, a higher MSE suggests that the model's predictions are further from the actual values, indicating a less accurate model.
- MSE is often used during the training and evaluation of regression models to assess their predictive accuracy. It's particularly useful when you want to penalize larger prediction errors more heavily, as the squared differences in the formula emphasize larger errors.
- The R-squared (R2) score, also known as the coefficient of determination, is a statistical metric used in machine learning and regression analysis to evaluate the performance of regression models. It measures how well the model explains the variance in the target variable.
- R2 score provides insight into the goodness of fit of a regression model. It represents the proportion of the variance in the dependent variable (the variable being predicted) that can be explained by the independent variables (the features used for prediction).
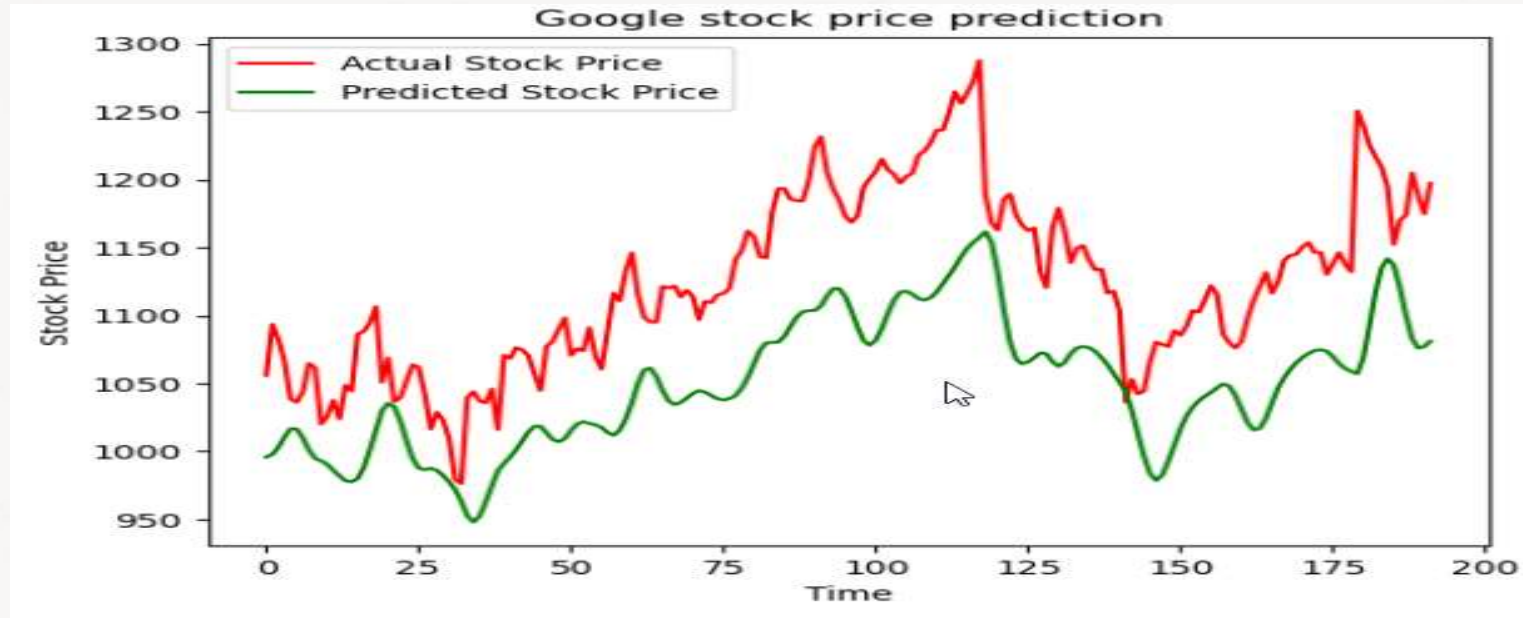
LSTM- Google

LSTM, which stands for Long Short-Term Memory, is a type of recurrent neural network (RNN) architecture used in deep learning and machine learning for various sequence-based tasks, including natural language processing and time series analysis. LSTM networks were not created by Google but have been widely adopted in the field of deep learning, and many organizations, including Google, have used them in their research and applications.

Here's how LSTM networks work and their relevance to Google:

- **LSTM Architecture**: LSTM is a specific type of RNN designed to address the vanishing gradient problem, which can limit the ability of traditional RNNs to capture long-range dependencies in sequential data. LSTM networks incorporate memory cells and gating mechanisms that allow them to store and retrieve information over extended sequences.
- **Applications at Google**: Google, being a technology giant with a strong focus on machine learning and artificial intelligence, has applied LSTM networks in various areas. For example, Google has used LSTM networks in natural language processing tasks, such as language translation and text generation, and in speech recognition systems like Google's Voice Search and Google Assistant. LSTM networks are also used in Google's DeepMind division, particularly for reinforcement learning and time series prediction tasks.
- **TensorFlow**: Google developed TensorFlow, a popular deep learning framework that provides tools and resources for implementing and training deep neural networks, including LSTMs. TensorFlow has made it easier for developers and researchers to work with LSTM networks and other deep learning architectures.
- **Research and Innovation**: Google, through its research initiatives, has contributed to the advancement of deep learning and LSTM networks. The company has published research papers and developed state-of-the-art models that incorporate LSTMs.

Prediction Plot Output-Google

Google stock price prediction

Thank You!