ECEN 360- Final Report

Stock Price Prediction

Leisha Khapre
Texas A&M
leisha@tamu.edu

Executive Summary

This project aimed to develop a predictive model for stock price forecasting, focusing on the S&P 500 index. Using deep learning techniques, specifically Long Short-Term Memory (LSTM) networks, the objective was to accurately predict future stock prices based on historical data. The methodology involved preprocessing the data, building and training the LSTM model, and evaluating its performance using data visualization. Key findings revealed the model's ability to capture temporal dependencies but highlighted challenges in predicting stock prices accurately due to market volatility and uncertainty. Despite all the efforts the model's performance fell short of expectations. Major conclusions include the need for a holistic approach which combines quantitative analysis with qualitative judgment and market expertise to make informed investment decisions. Additionally, further research and refinement of the model architecture, incorporation of additional data sources, and continuous evaluation and iteration are recommended to better prediction accuracy and reliability of the model because there are several aspects to price prediction than a mere study of the historical prices.

Introduction

Stocks are a share of a company. When the company does well, the price of the stock appreciates and when the company does poorly, the value of the stock depreciates.

The S&P 500 is an index of the largest 500 companies in the stock market. Since it is one of the most followed equity indices in the world, it is a popular and a relatively safe choice among Americans.

The primary objective of this project was to develop a predictive model for stock price forecasting, focusing specifically on the S&P 500 index. Today, the stock market, specifically the S&P 500, is important for individuals to build wealth in order to support themselves throughout retirement. Stock prices reflect the way the economy is doing which is why they are extremely crucial. The prediction of stock prices would ultimately help in predicting the economic state of the country as well. Predicting stock prices is a challenging task due to the complex and dynamic nature of financial markets. Stock prices are influenced by a wide range of factors, which include but are not limited to company performance, economic indicators, geopolitical events, and investor sentiment. Traditional methods of stock price prediction have limitations and may not always provide accurate forecasts.

This project aimed to leverage LSTM networks to develop a predictive model capable of capturing temporal patterns in historical stock price data and generating forecasts for future price movements.

This model should be prefaced with the fact that no one can accurately predict the future and this model will definitely not. Events in society have a large impact on stock prices that cannot be accurately accounted for by just viewing trends in historical stock prices. It is important to do the due diligence and understand the risk and volatility of investing in the stock market. Monitoring the news and researching into a company's financial reports is crucial to support any predictions that this model makes.

Methodology

The methodology used in this project encompassed several steps to develop and evaluate the predictive model for stock price forecasting. These steps include data preprocessing, model development, training, evaluation, and optimization. Below is a detailed description of each component:

Data preprocessing:

- The project began with the collection of historical stock price data for the S&P 500 index. Data was obtained from Yahoo Finance which is a built in library in python.

- The raw data underwent preprocessing steps to handle missing values, remove outliers, and normalize the features. Time series data was organized into sequential samples suitable for training the LSTM model.

Model Development:

- The core of the project involved the implementation of an LSTM neural network architecture for time series forecasting. LSTM networks work well for capturing temporal dependencies in sequential data, which made them an ideal choice for stock price prediction tasks and this project.
- Keras and TensorFlow were employed to achieve the LSTM model.
- The architecture of the LSTM model included input layers- 2 were used for this project- LSTM layers with specified units, activation functions, and output layers. Hyperparameters such as learning rate, batch size, and number of epochs were fine-tuned to optimize model performance and make it as efficient as possible.
- Techniques to mitigate overfitting, such as dropout and regularization, were incorporated into the model architecture to improve generalization to unseen data.

Model Training:

- The dataset was split into training, validation, and testing sets to train and evaluate the LSTM model. The training set was used to optimize model parameters, while the validation set was employed to monitor model performance and prevent overfitting.

- The model was trained using the training data with a specified number of epochs. The training and validation loss was plotted to make sure the model was working as expected. The predictions were then made on the training and testing data set to ensure proper working of the model.

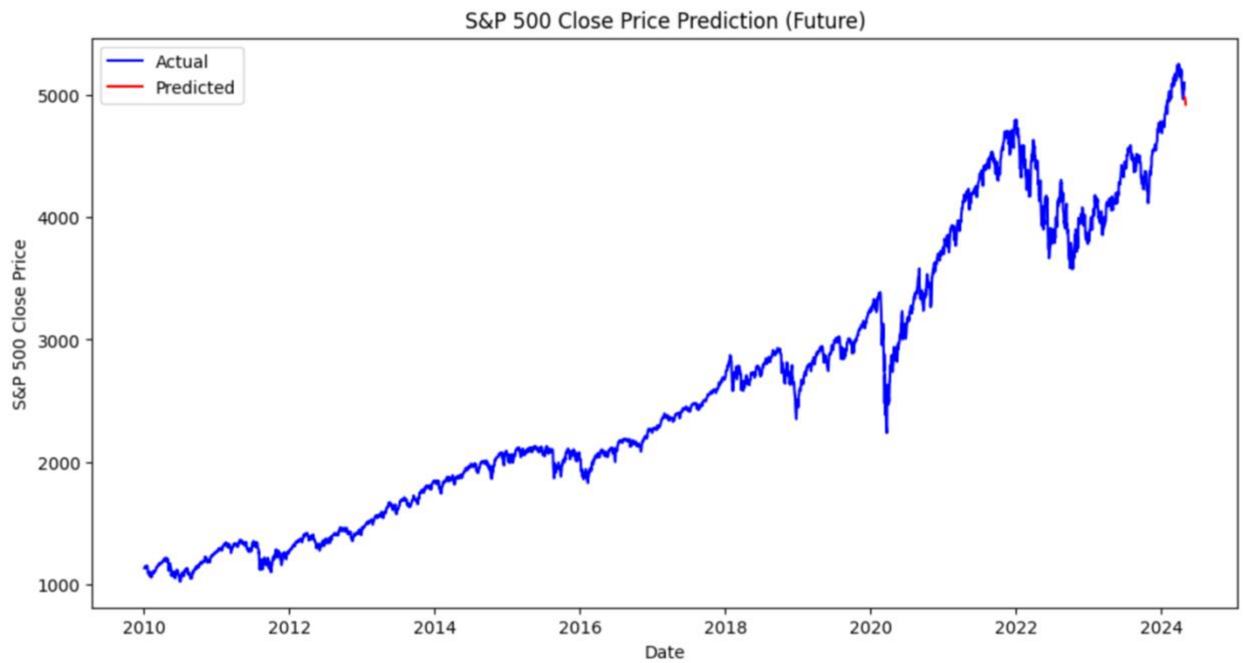
Model evaluation and prediction:

- Predictions were made for the entire dataset which were then inverse scaled to restore them to their original scale for meaningful interpretation and comparison with the actual values.
- A set number of days was defined to predict in the future for and the prediction was visualized using matplotlib along with the actual valued of the data.

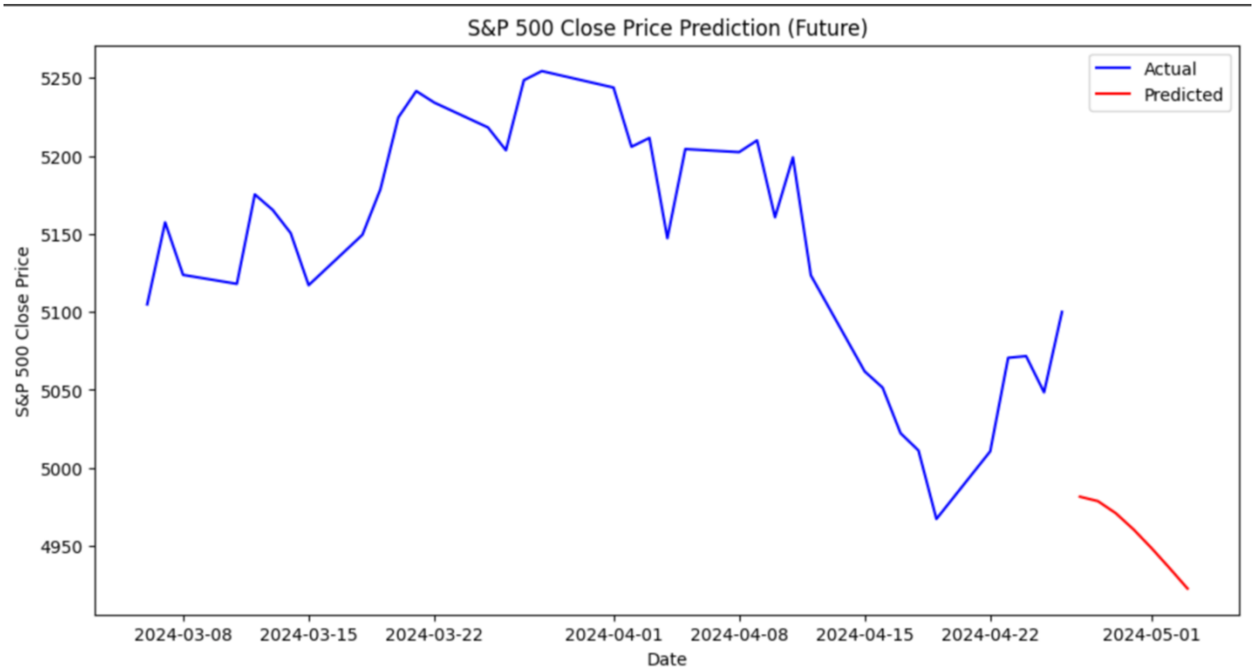
Results

The findings of the project are presented here, supported by data visualizations to provide insights into the performance of the predictive model for stock price forecasting.

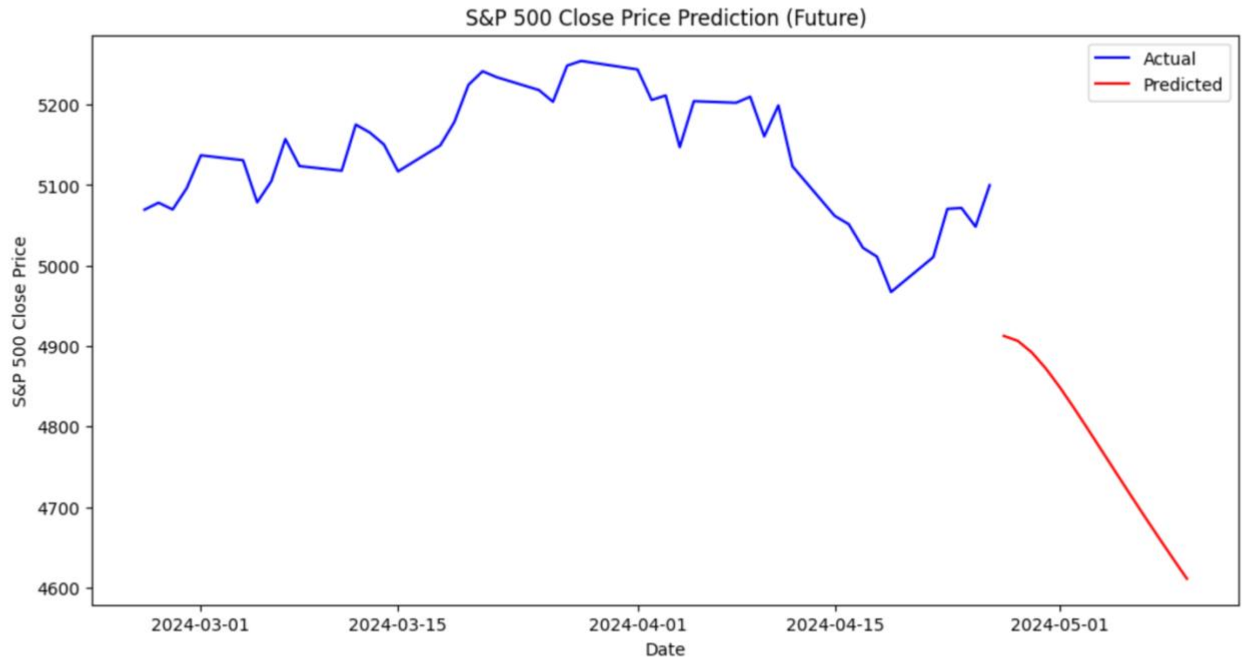
This is what the model predicted for 7 days along with the actual historical values from 2010.



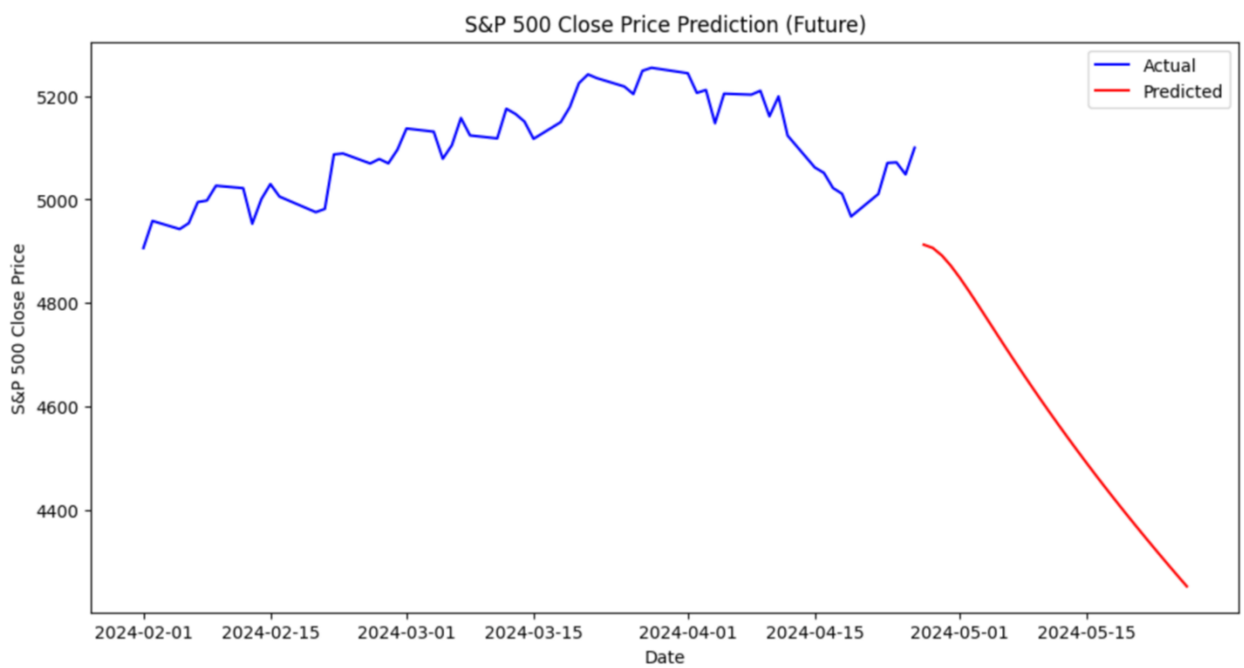
The next few graphs are a zoomed in version of the stock price data focused on the predictions.



Following is a graph of what the model predicted for a week into the future.



Finally, the model's prediction for a month ahead is visualized below.



Discussion

While the LSTM model attempted to forecast stock prices, it struggled to keep up with the wild swings of the stock market. It is highly unlikely that the stock prices would follow the prediction of the model and decrease in a manner that would make the graph a straight line. It is too good to be true and quite unrealistic.

Despite efforts to account for past trends and prevent overfitting, the model often missed the mark. The model's reliance on historical data may lead to limited generalization to future scenarios, especially when faced with unforeseen events or structural changes in the market. There is still much improvement to be made to the model before it can be put to any practical use.

This highlights the complexity of navigating financial markets, where quantitative analysis alone may not suffice. However, while the LSTM model may not provide perfect predictions of stock prices, it can still offer valuable insights for investors, traders, and financial analysts. Achieving reliable predictions would require a blend of insights and expert intuition which investors may have experience with to navigate the ever-changing environment of investment. Continuous refinement and adaptation of the model to evolving market conditions is essential to improving its predictive capabilities and enhancing its value.

Conclusion

In summary, the project explored the application of LSTM networks for stock price prediction, with a specific focus on the S&P 500 index. Despite the careful preprocessing of data and optimization of model parameters, the LSTM model's performance in forecasting stock prices remained suboptimal. The unpredictability and volatility of financial markets pose significant challenges for predictive modeling, highlighting the limitations of relying solely on historical data for future predictions. While the LSTM model may offer valuable insights, its effectiveness is possible only by continuous refinement and adaptation to the ever-changing market conditions. Therefore, the project underscores the importance of combining quantitative analysis with qualitative judgment and market expertise for making informed investment decisions. Future research should aim to enhance the model's validity by incorporating additional data sources and refining model architecture to improve prediction accuracy and reliability. Ultimately, while predictive modeling can serve as a valuable tool, it should be complemented with a comprehensive understanding of market dynamics and risk assessment to navigate the complexities of the financial environment effectively.

Next Steps

Moving forward, several next steps can be undertaken to improve the predictive capabilities of the model and enhance its applicability in real-world scenarios. Firstly, experimenting with different algorithms and model architectures will help to offer valuable insights into which approaches are most effective for stock price prediction tasks. This experimentation process could involve exploring variations of recurrent neural networks (RNNs), such as gated recurrent units (GRUs), or even hybrid models combining deep learning with traditional statistical methods like linear regression or multi regression models. Furthermore, evaluating the model's performance on new data and iteratively refining it based on feedback and insights gained from real-world applications is essential. This iterative process would allow for continuous learning and adaptation to evolving market conditions, thereby improving the model's accuracy and reliability over time. Additionally, integrating qualitative analysis, such as sentiment analysis of news articles or social media data, would the complement quantitative data in making more informed predictions. By combining quantitative and qualitative approaches, the model can better navigate the complexities of financial markets and support more effective decision-making for investors and traders alike.

References

- https://otexts.com/fpp2/intro.html
- https://docs.aws.amazon.com/sagemaker/latest/dg/deepar.html
- https://www.seldon.io/machine-learning-regression-explained#:~:text=Machine%20Learning%20Regression%20is%20a,used%20to%20predict%20continuous%20outcomes.
- https://youtu.be/CbTU92pbDKw?si=XcY7aq5v9vlnf9gb