IBM – Stock Market Prediction

Group 6

Mohith Harish Sethupandian - 22018914
Praveenkumar Croos Anthonimuthu - 22033745
Samson Raj Babu Raj - 22013145
Sathish Balachandran - 22021050
Shanmugapriya Ramakrishnan - 22026376

University of Hertfordshire UH

1

1

- Introduction
- Stock Market Selection

2

- LSTM
- LSTM - Architecture

3

- Methodology
- Prediction
- Ethical Implication
- Conclusion

University of Hertfordshire UH

Introduction

The purpose of the data science group project is to explore the world of financial markets and forecasting. By using deep learning, we aim to predict IBM stock prices in the future. As one of the world's leading technology corporations, IBM's stock price movement can provide valuable insights and potentially lucrative investment strategies.

To solve this problem, we use Long Short-Term Memory (LSTM) networks, a special type of Recurrent Neural Network (RNN) that works well with time-series data such as stock prices. LSTM networks can learn patterns over time and predict the future based on this information.

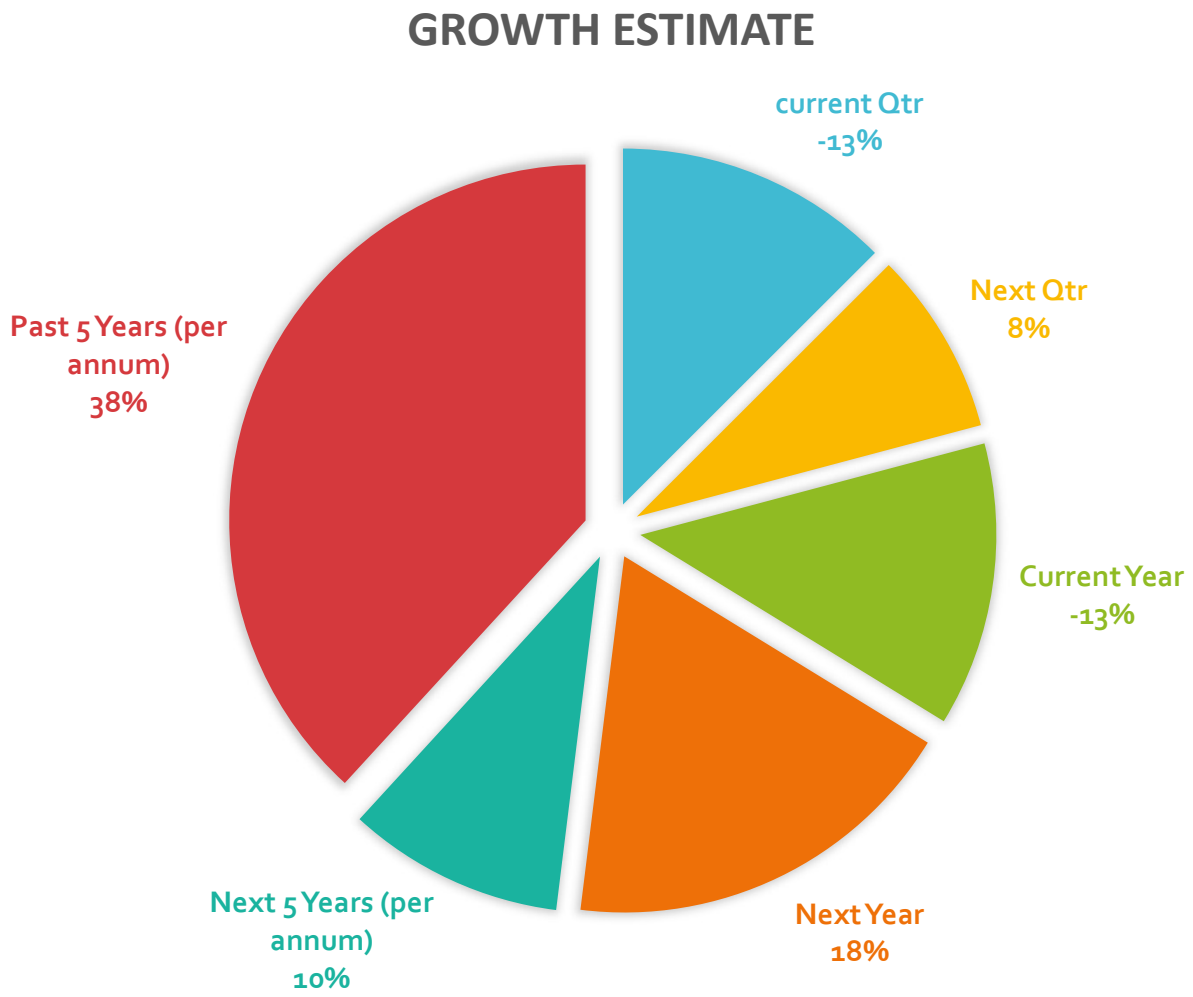
From data preprocessing, model building, and training to prediction and visualization, this project covers it all.

Github code – https://github.com/bksat90/StockPrediction/blob/main/StockPrediction.ipynb

University of Hertfordshire UH

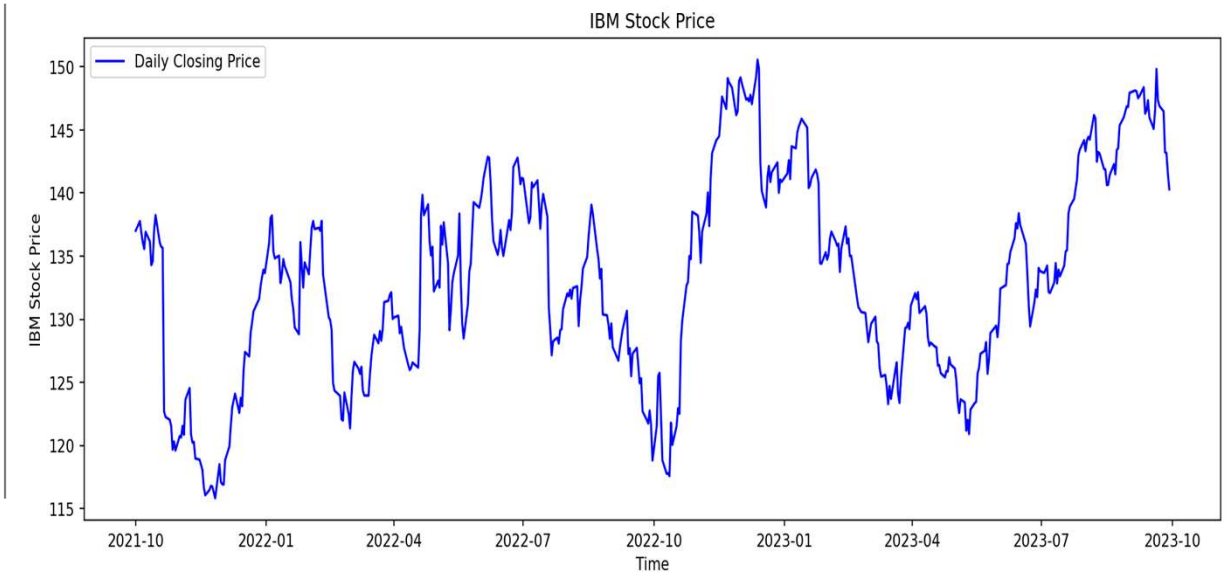
Stock Growth Estimate

IBM Stock For the corresponding Year(2011-2022)



Data

- IBM Stock details were captured from https://finance.yahoo.com/
- “yfinance” library is used to procure the data for this project.
- Data for two years from 1st Oct 2021 to 30th Sep 2022 is used. This period was arbitrarily chosen.
- This data has open price, high price, low price, close price, adjusted close price and volume of the stocks.
- Here the close price of the stocks is used.



LSTM

Long Short Term Memory

- Long short- term Memory network is a special kind of RNN network introduced by **Hochreiter** & **Schmidhuber** which has the capability to solve long-term dependency real time problems.
- The **feedback loop** allow RNN network to be a better pattern recognition compared to another neural network.
- LSTM model is explicitly used in sequence learning problem including language modeling , speech recognition , sequence prediction.

LSTM - Architecture

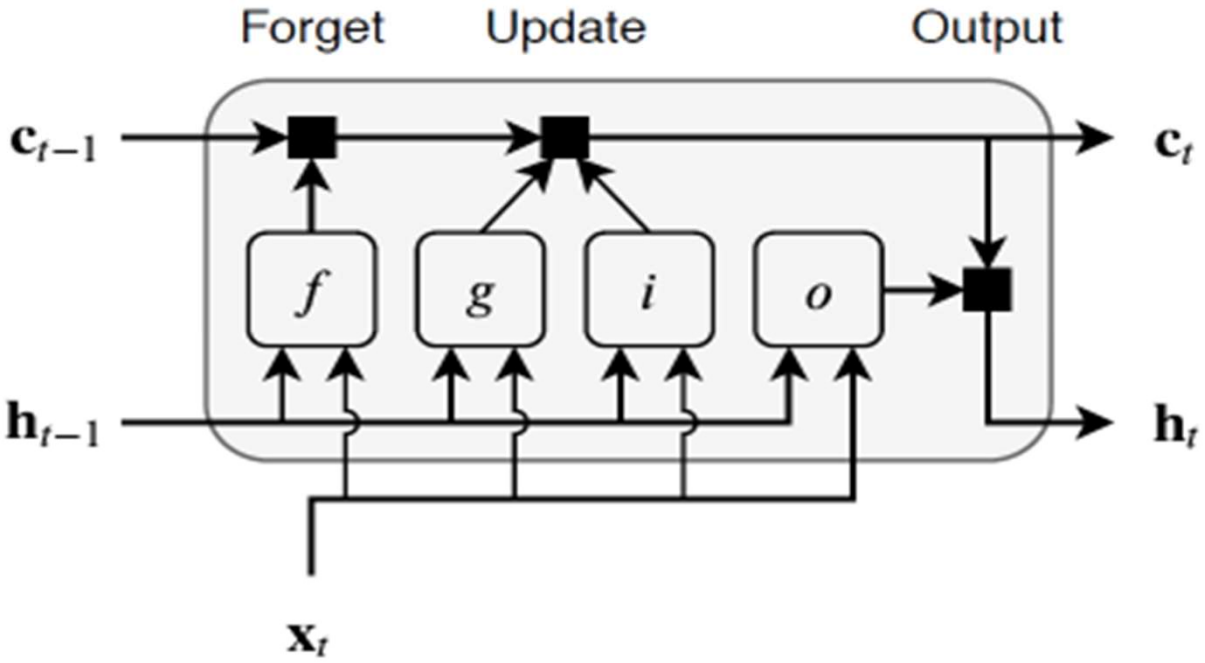
Component

Input gate (i) - Control level of cell state update.

Forget gate (f) - Control level of cell state forget

Update gate (g) -
Add information to cell state.

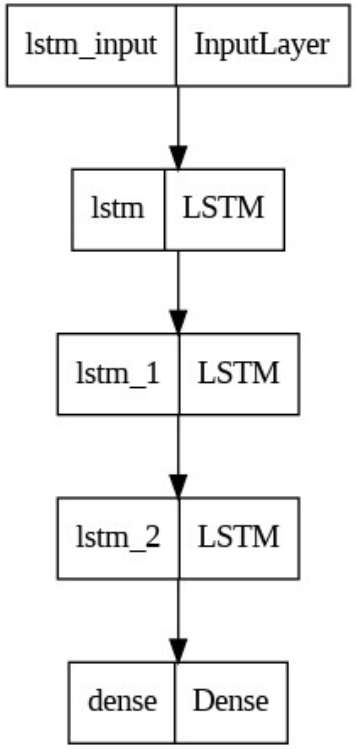
Output gate (o) - Control level of cell state added to the hidden state.



University of Hertfordshire UH

LSTM - Methodology

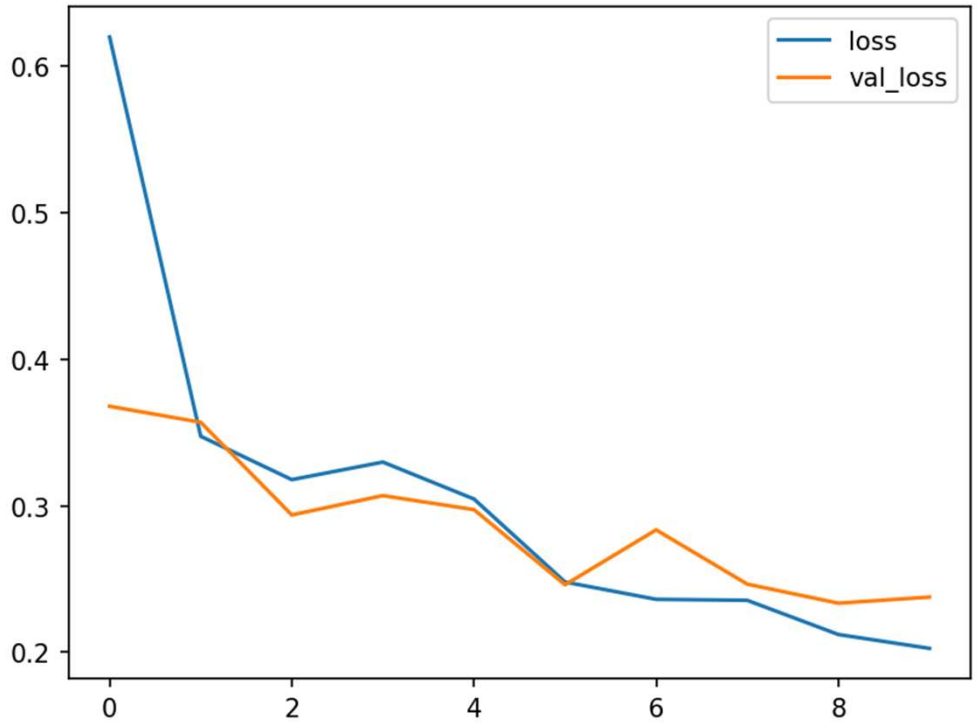
- LSTM is used in this model.
- Keras library is used to model the prediction in TensorFlow.
- This model uses three LSTM layers and one dense layer.




```
[ ] # display the model summary
    model.summary()

Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 lstm (LSTM)                 (None, 40, 50)            10400

 lstm_1 (LSTM)               (None, 40, 100)           60400

 lstm_2 (LSTM)               (None, 100)               80400

 dense (Dense)               (None, 1)                 101

=================================================================
Total params: 151301 (591.02 KB)
Trainable params: 151301 (591.02 KB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

Loss vs Value Loss

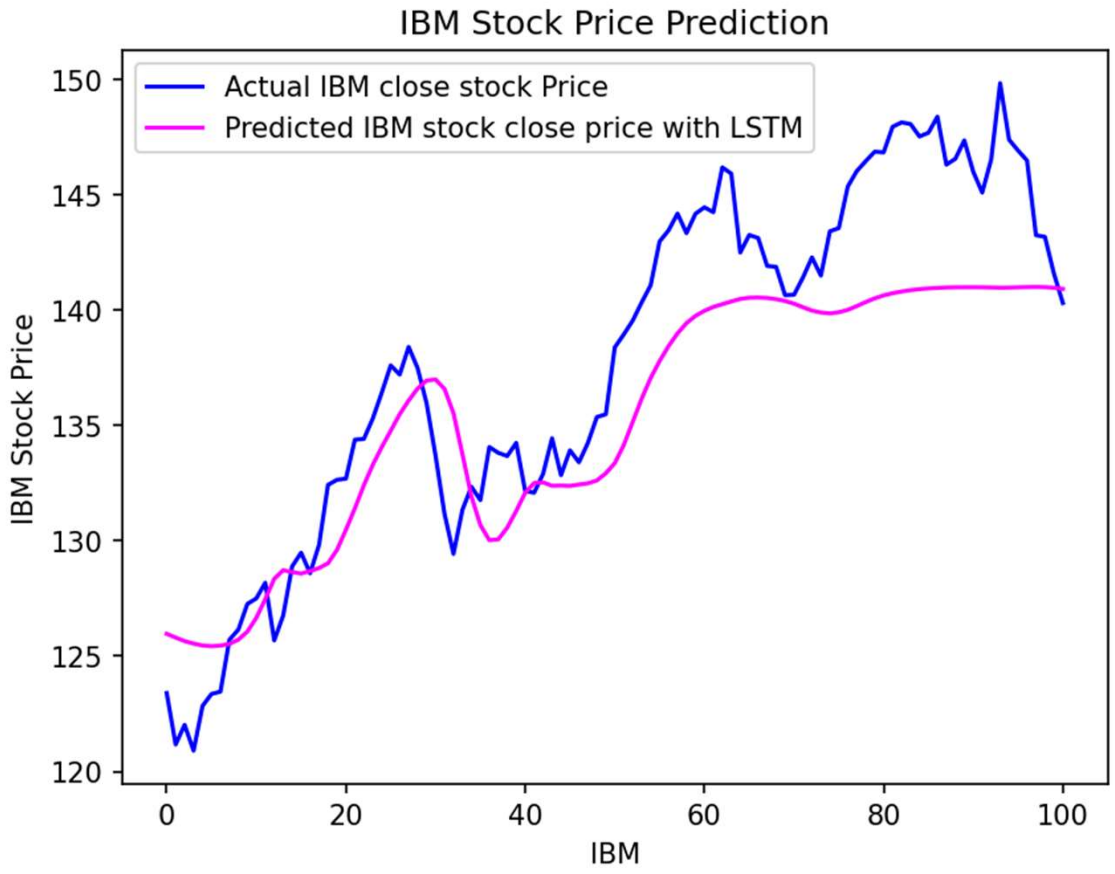
- Learning curve is plot with loss values and value loss for each epochs.
- Training loss drops suddenly and decreases gradually whereas test loss drops and then increases to the value higher than training loss. This behaviour is as expected.



University of Hertfordshire UH

Prediction

- This graph represents a time series plot that contrasts the closing price of IBM shares with the estimate.
- The pink line shows the anticipated price, and the blue line shows the actual closing price.
- There may be a relationship between the two if the actual closing price often exceeds or falls short of the expected price. Furthermore, it may be a clue that something is wrong with the stock if the projected price begins to deviate from the actual closing price.



University of Hertfordshire UH

Ethical Implication

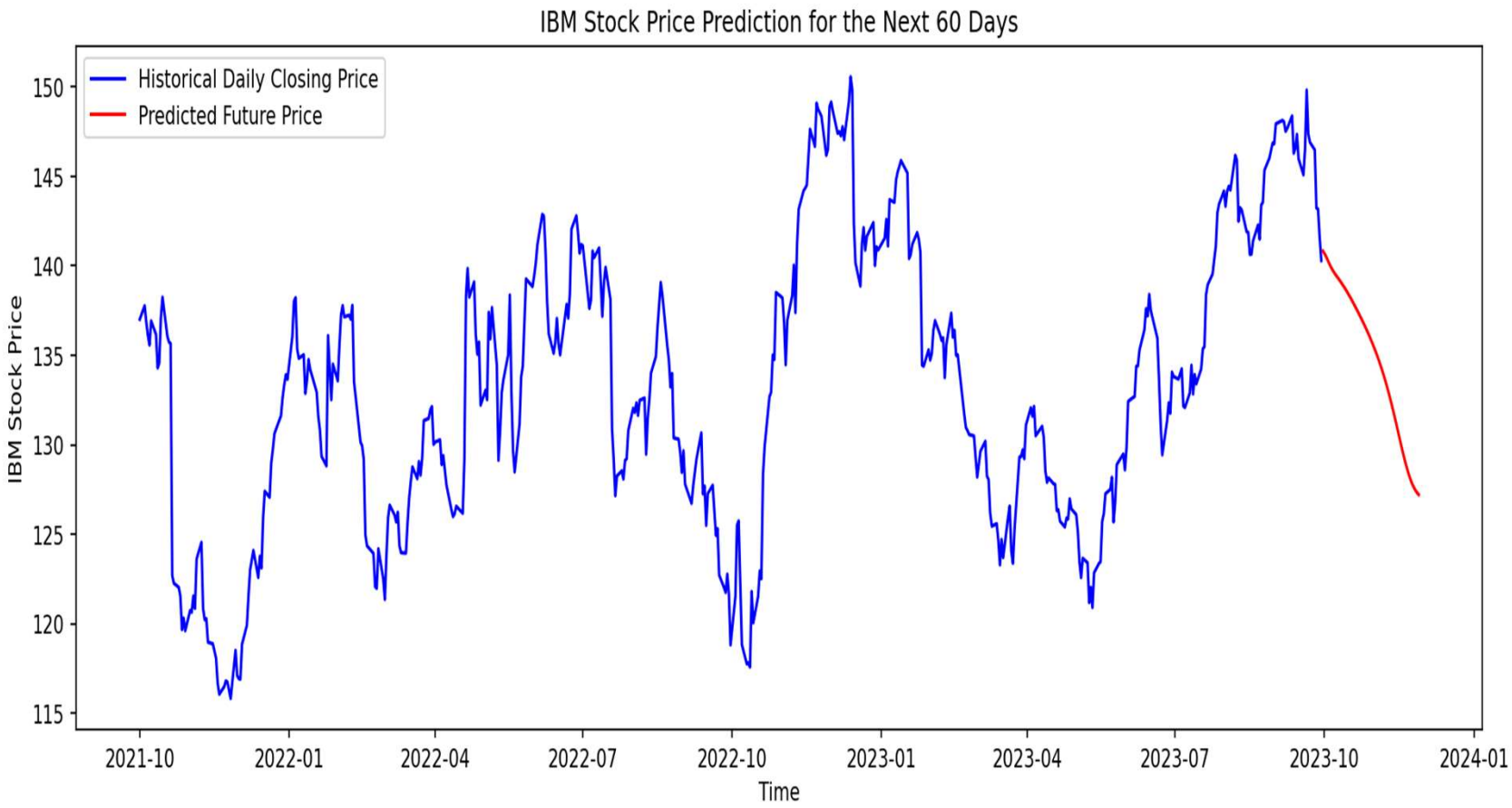
Stock Market Prediction using machine learning

- **Explainability and transparency**: Machine learning algorithms are frequently complicated and opaque, making it challenging to comprehend how they make their predictions. Investors may find it challenging to evaluate the predictability and accuracy of the data due to this lack of transparency, which could result in them making poorly informed investment decisions.
- **Algorithmic bias**: The data used to train machine learning algorithms may have biases. The algorithm's predictions may be unfair or discriminating if the training data is skewed, and the algorithm itself will be biased as well. An algorithm that has been trained using past stock market data, for instance, can have a bias in favor of particular businesses or sectors, which could have unfavorable consequences for investors who do not fall into those categories.
- **Market manipulation**: Using potent computers, high-frequency traders, or HFTs.

Conclusion

Recurrent neural networks (RNNs) and long short-term memory (LSTM) networks have shown promise in predicting IBM stock prices, demonstrating the ability to capture long-term dependencies and trends in time-series data relevant to stock price movements. While these models can achieve reasonable accuracy for short-term forecasts, it is important to acknowledge their limitations and not solely rely on them for investment decisions. Investors should consider fundamental analysis, market sentiment, and diversify their strategies when making investment choices.

Future Prediction



University of Hertfordshire UH

Project Team

Each task is assigned to the group members and the contribution is recorded below and regular Group meeting were conducted to understand the challenges and to finish the tasks.

Task 1 – Market selection (Praveenkumar Anthonimuthu)
Task 2 – Model selection (Samson Raj, Praveenkumar Anthonimuthu)
Task 3 – Code (Sathish Balachandran)
Task 4 – Prediction (Sathish Balachandran, Mohith Sethupandian, Shanmugapriya Ramakrishnan)
Task 5 – Presentation (Samson Raj, Mohith Sethupandian, Shanmugapriya Ramakrishnan, Sathish Balachandran)

References

- https://colah.github.io/posts/2015-08-Understanding-LSTMs/
- https://medium.com/@redeaddiscolll/a-comprehensive-guide-to-forecasting-ibm-stock-prices-using-lstm-networks-in-pytorch-639223142a9
- https://developer.nvidia.com/discover/lstm
- https://uk.mathworks.com/campaigns/offers/deep-learning-for-signal-processing-white-paper.html
- https://uk.mathworks.com/help/deeplearning/ug/long-short-term-memory-networks.html
- https://finance.yahoo.com/quote/IBM