

GAYATRI VIDYA PARISHAD COLLEGE OF ENGINEERING (AUTONOMOUS)



STOCK MARKET PRICE PREDICATION USING MACHINE LEARNING

Under the guidance of :
K. Suma Sree ,[M Tech],
Assistant Professor ,
Computer Science Dept.

Presented By :
B. Yashaswini (18135A0505)
B. Sai Kiran (18135A0507)
B. Manjunadh (18135A0508)

CONTENTS

- Abstract
- Introduction
- Tools and Technologies used
 - LSTM
- System Requirements
- Architecture Design
- Control Flow Diagram
- Methodology
 - Stage 1 – 5
- Output Screens
- Conclusions
- Future Enhancements
- References

ABSTRACT

- Predicting stock market prices is a complex task that traditionally involves extensive human-computer interaction. Due to the correlated nature of stock prices, conventional batch processing methods cannot be utilized efficiently for stock market analysis. Stock price prediction is one among the complex machine learning problems. Machine learning itself employs different models to make prediction easier and authentic. Stock prices are represented as time series data and neural networks are trained to learn the patterns from trends.
- Here we use learning algorithm that utilizes a kind of recurrent neural network (RNN) called Long Short Term Memory (LSTM) to predict stock values.

INTRODUCTION

- The stock market is a vast array of investors and traders who buy and sell stock, pushing the price up or down. The prices of stocks are governed by the principles of demand and supply, and the ultimate goal of buying shares is to make money by buying stocks in companies whose perceived value (i.e., share price) is expected to rise.
- From the rise of Machine Learning era, there has been many algorithms developed for time series forecasting problems, but LSTM beats them all.
- There are algorithms like Moving Average, Linear Regression, ARIMA, but LSTM over shines them all.

TOOLS & TECHNOLOGIES USED :



Keras : Keras is Tensorflow's High level API for building and training deep learning models. We import Sequential model, LSTM and Dense layers from keras.



Numpy : It is a library for Python, adding support for large, multi-dimensional arrays and matrices. Having large collection of high level math functions, numpy makes our work much more easier.



Matplotlib : It is a plotting library for Python. We plot our actual data and predicted data using matplotlib functions.



Scikit-learn : It is one of the machine learning libraries for Python, featuring various classification, regression and clustering algorithms.



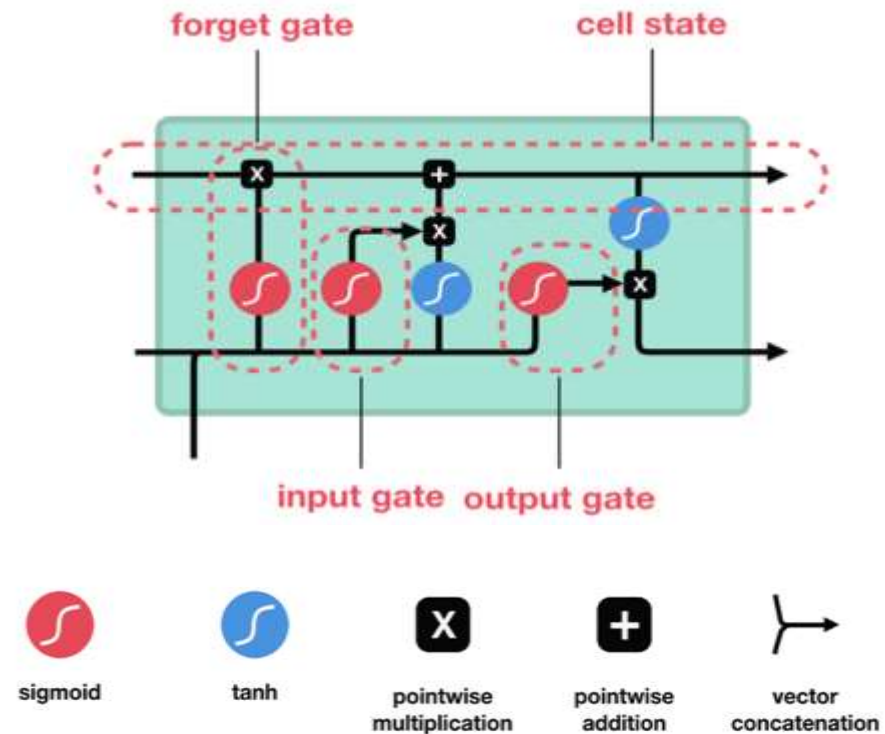
Jupyter : It is a workflow environment, where we download our datasets, plotting or results and experiment with data.



Pandas : It is sw library for data manipulation and analysis. It offers data structures and operations for manipulating tables and time series

LSTM (Long Short Term Memory)

- Long Short-Term memory is one of the most successful RNNs architectures.
- LSTM introduces the memory cell, a unit of computation that replaces traditional artificial neurons in the hidden layer of the network.
- With these memory cells, networks are able to effectively associate memories in time, hence suit to grasp the structure of data dynamically over time with high prediction capacity.



SYSTEM REQUIREMENTS

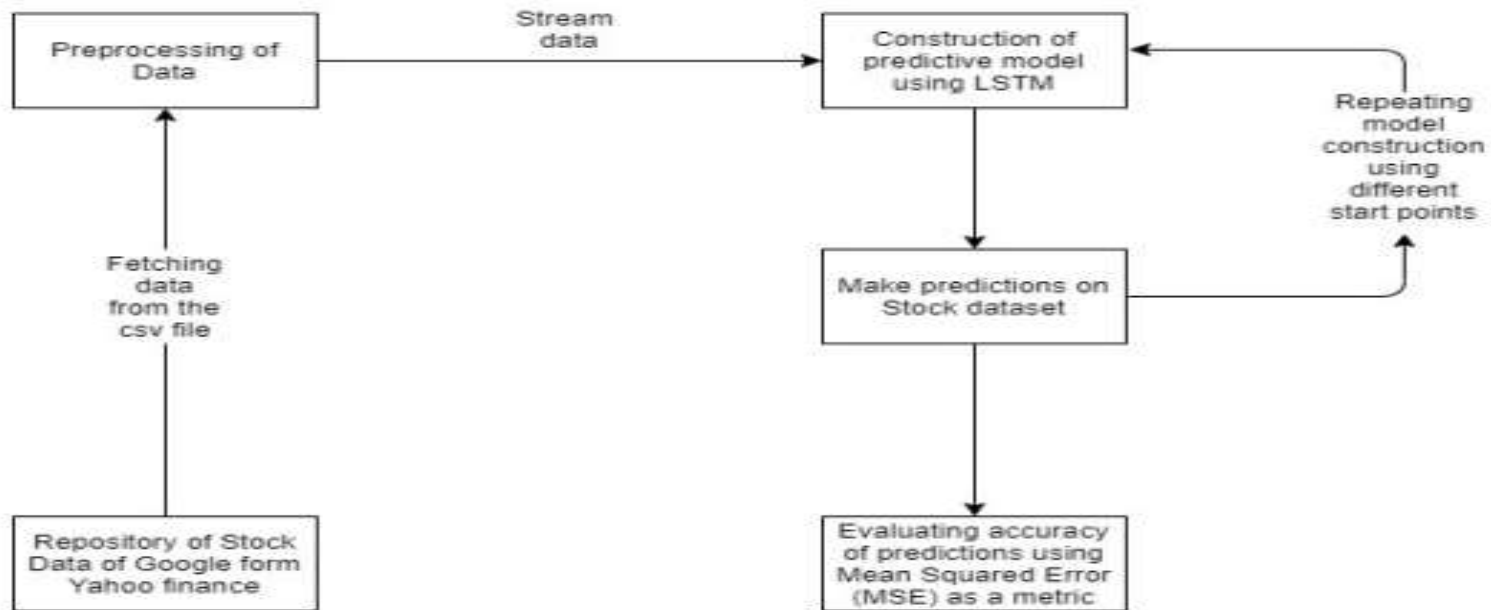
SOFTWARE REQUIREMENTS:

- Operating System : Windows / Linux
- Software : Collab / Jupyter / Any Python IDE
- Language : Python
- Libraries:
 - NumPy
 - Pandas
 - Tensorflow
 - Keras

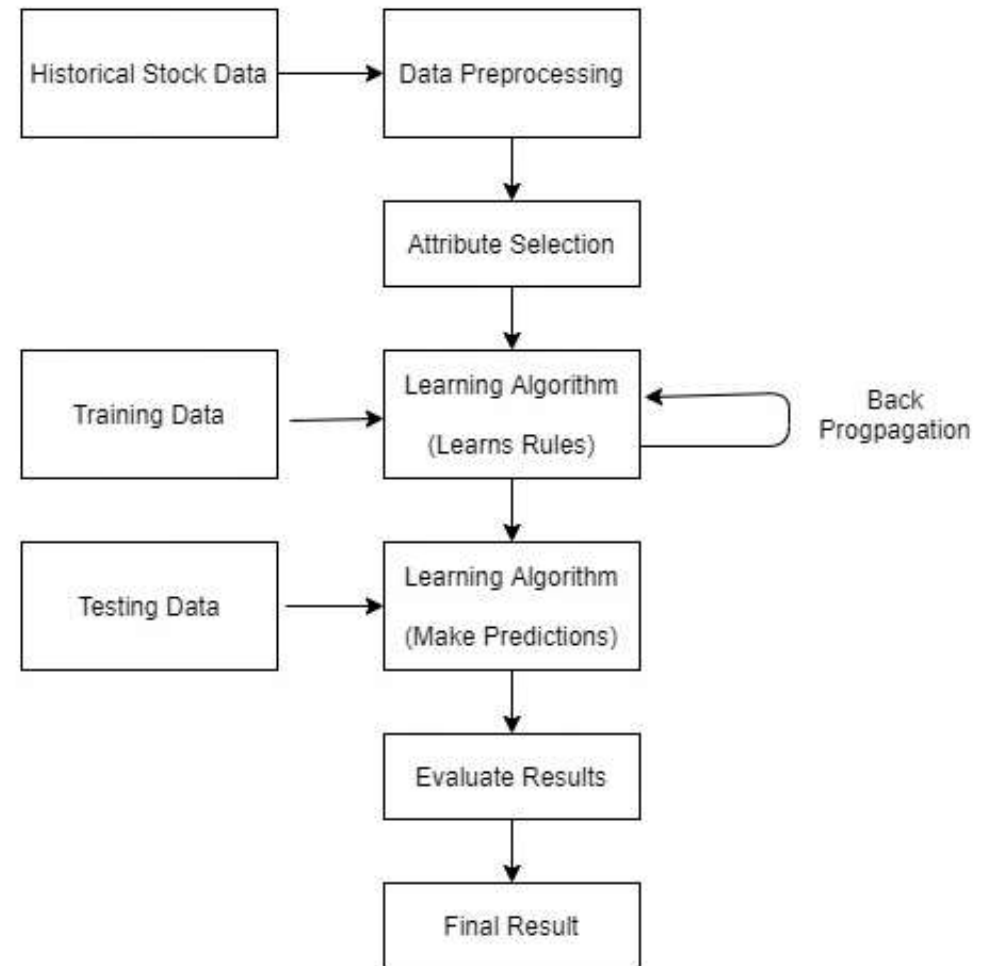
HARDWARE REQUIREMENTS:

- System with GPU
- Processor : intel I5
- Ram : 8GB

ARCHITECTURE DESIGN



CONTROL FLOW DIAGRAM



METHODOLOGY

Stage 1 : Raw Data :

- In this stage, the historical stock data of google is collected from :
 - <https://finance.yahoo.com/quote/GOOG/history?p=GOOG>
- This historical data is used for the prediction of future stock prices. We have collected data from 2012 to 2017 .

Time Period: May 01, 2012 - May 01, 2017 ▼

Show: Historical Prices ▼

Frequency: Daily ▼

Stage 2 : Data Preprocessing

- Data Preprocessing is a technique that is used to convert the raw data into a clean data set.
- The pre-processing stage involves :
 - a.) **Data cleaning:** Fill in missing values.
 - b.) **Data rescaling:**
Normalization

	Date	Open	High	Low	Close	Volume
0	01-03-2012	325.25	332.83	324.97	663.59	73,80,500
1	01-04-2012	331.27	333.87	329.08	666.45	57,49,400
2	01-05-2012	329.83	330.75	326.89	657.21	65,90,300
3	01-06-2012	328.34	328.77	323.68	648.24	54,05,900
4	01-09-2012	322.04	322.29	309.46	620.76	1,16,88,800
...
1273	1/25/2017	829.62	835.77	825.06	835.67	14,94,500
1274	1/26/2017	837.81	838.00	827.01	832.15	29,73,900
1275	1/27/2017	834.71	841.95	820.44	823.31	29,65,800
1276	1/30/2017	814.66	815.84	799.80	802.32	32,46,600
1277	1/31/2017	796.86	801.25	790.52	796.79	21,60,600

Stage 3 : Feature Extraction

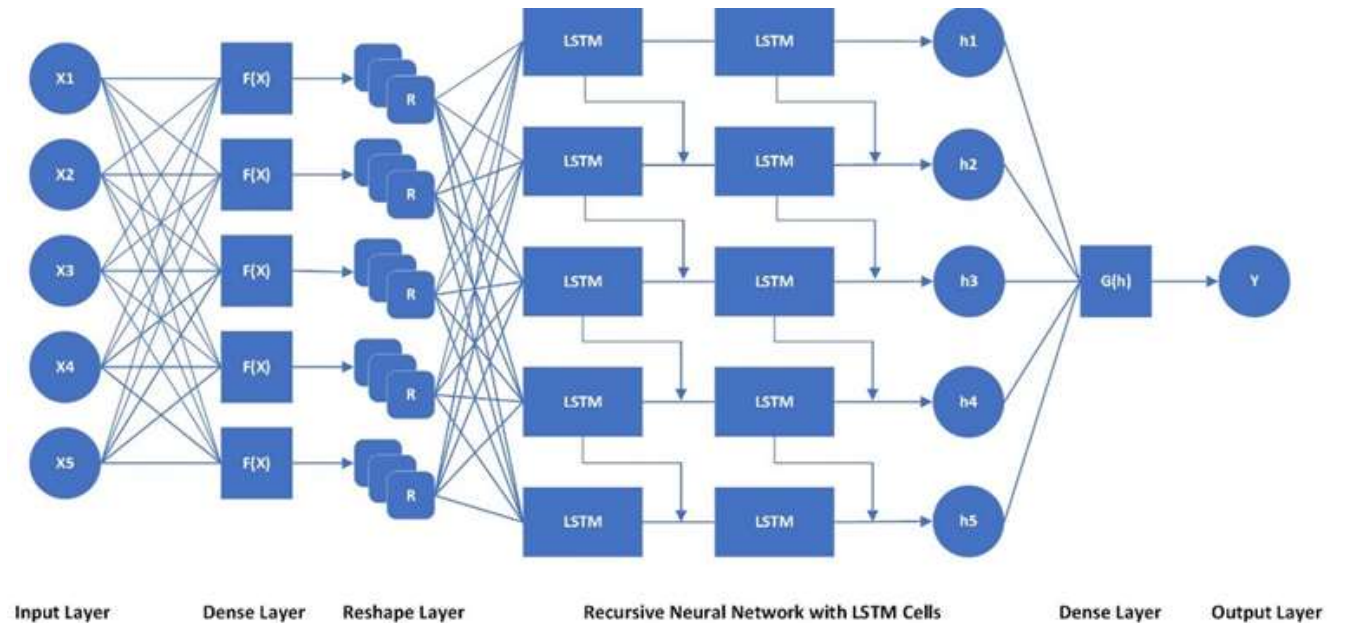
- In this layer, only the features which are to be fed to the neural network are chosen. We will choose the feature from Date, open, high, low, close, and volume.
- We chose “Open” feature to work with LSTM’s

```
➜      Open
0      325.25
1      331.27
2      329.83
3      328.34
4      322.04
...
1017   688.61
1018   702.18
1019   723.60
1020   723.58
1021   713.85
```

```
[1022 rows x 1 columns]
```

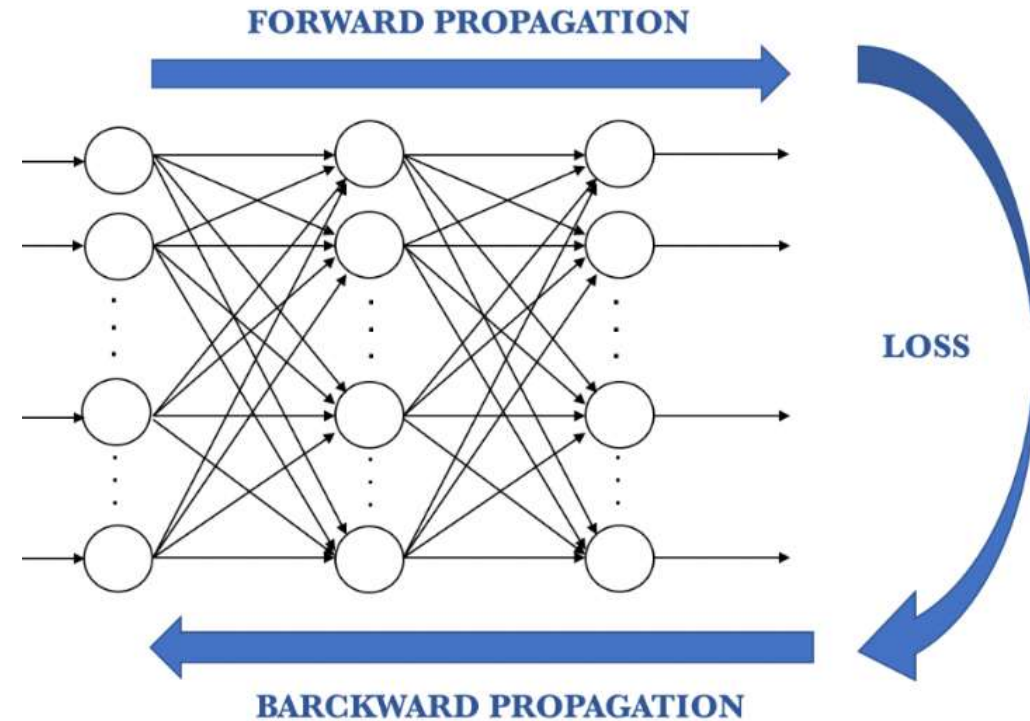
Stage 4 : Training Neural Network

- In this stage, the data is fed to the neural network and trained for prediction assigning random biases and weights. Our LSTM model is composed of a sequential input layer followed by 10 LSTM layers and a dense layer with sigmoid activation and then finally a dense output layer with linear activation function.



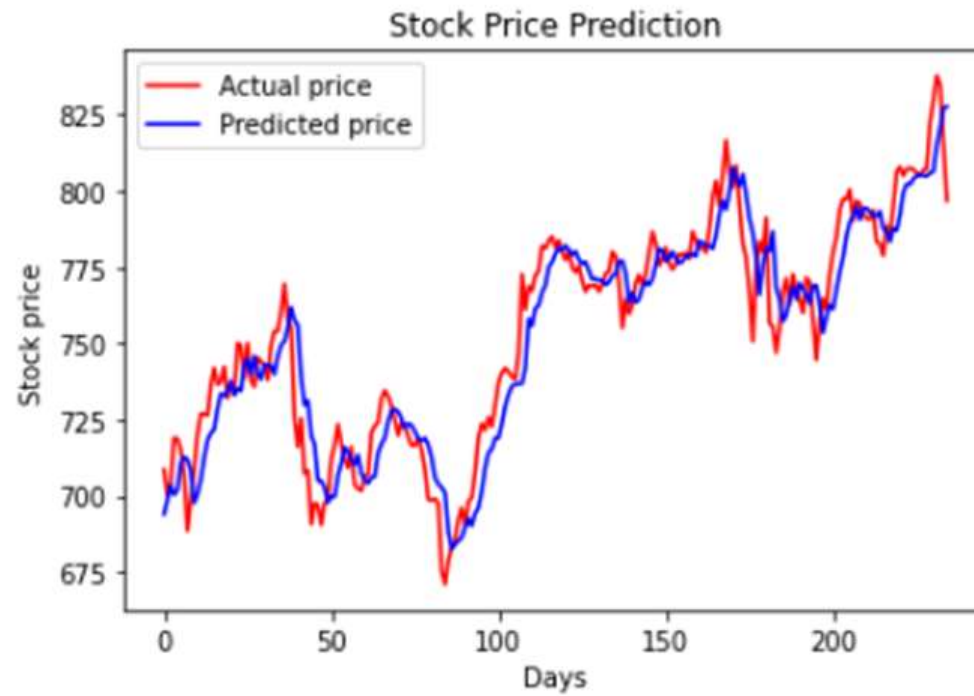
Stage 5 : Output Generation

- In this layer, the output value generated by the output layer of the RNN is compared with the target value. The error or the difference between the target and the obtained output value is minimized by using back propagation algorithm which adjusts the weights and the biases of the network.



OUTPUT SCREENS

Visualization of TEST DATA by LSTM



MSE Value obtained by LSTM : 9.975059700162928

CONCLUSION

- The popularity of stock market trading is growing rapidly, which is encouraging researchers to find out new methods for the prediction using new techniques. The forecasting technique is not only helping the researchers but it also helps investors and any person dealing with the stock market. In order to help predict the stock indices, a forecasting model with good accuracy is required.
- In this work, we have used one of the most precise forecasting technology using Recurrent Neural Network called Long Short-Term Memory unit which helps any person interested in investing in the stock market by providing them a good knowledge of the future situation of the stock market.

FUTURE ENHANCEMENT

- Several future enhancements can be made, by increasing the richness of the data and the size of the dataset we might be able to predict the stock prices at a greater precision. There is also good scope for new algorithms for predicting time series forecasting which may help in predicting stock prices. Using bigger dataset can improve the output quality of predicted stock price.

REFERENCES

- IRJET paper of “Stock price prediction using Long Short Term Memory” by Raghav Nanda Kumar, Uttamraj K, Vishal R, Y V Lokeshwari
- International Journal on “Stock Market Forecasting Techniques” by Vivek Rajput, Sarika Bobde
- <https://towardsdatascience.com/predicting-stock-price-with-lstm-13af86a74944>
- https://www.researchgate.net/publication/328930285_Stock_Market_Prediction_Using_Machine_Learning



THANK YOU