PREDICTING STOCK PRICES USING LSTM

1)Summary:

Stock market price forecasting is a challenging process that frequently requires intensive human-computer interaction. Traditional batch processing techniques cannot be used effectively for stock market analysis due to the linked nature of stock prices. We suggest an online learning approach that makes use of the Long Short Term Memory (LSTM) kind of recurrent neural network (RNN). When compared to current stock price prediction systems, this will produce more accurate findings. With varied data sizes, the network is trained and assessed for accuracy, with the results being tabulated.

1.1)Introduction:

A major issue in the economic world has always been the stock price's tendency to fluctuate. Many internal and external elements, including the domestic and international economic environment, global situation, industry prospect, financial information of listed businesses, and stock market operation, have an impact on stock prices. As a result, the forecasting approach has a different emphasis.
The rise and fall of share prices can be attributed to a few Key Performance Indicators (KPIs), which demonstrate how closely stock markets and the realm of economics are related. The opening stock price, closing stock price, intraday low price, intraday peak price, and volume of all stocks traded during the day are the five KPIs that are most frequently employed.

2)Problem Statement:

Norcom Inc's main source of revenue is commission earned on the stocks traded through our app. Due to increasing competition in the market they are losing a significant number of active investors on the app. The goal is to come up with a stock predicting model with high accuracy which would enable the investors to take high profit trades.

2.1) EXISTING SYSTEMS AND THEIR DRAWBACKS;

Traditional methods for analyzing the stock market and predicting stock prices include fundamental analysis, which considers a stock's performance in the past and the overall credibility of the company, and statistical analysis, which is only interested in crunching numbers and finding patterns in stock price variation. The latter is frequently accomplished using Artificial Neural Networks (ANNs) or Genetic Algorithms (GA), however these algorithms

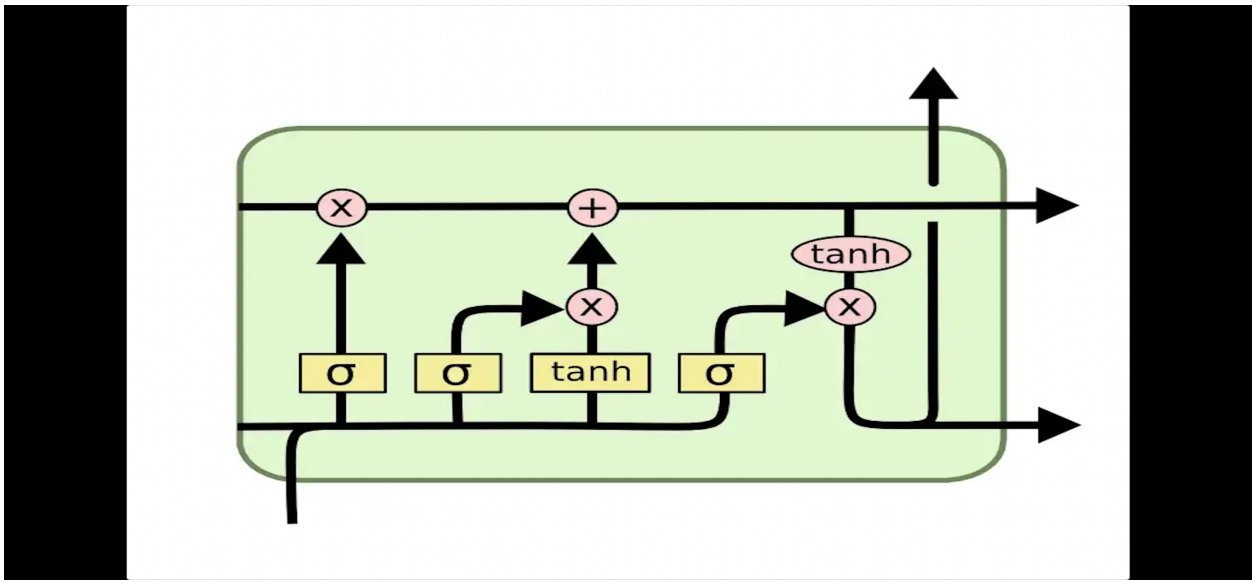
are unable to capture correlation between stock prices in the form of long-term temporal dependence.

3)Proposed Technique:

The proposed Idea of this project is to predict stock prices using just one KPI (end-of-day-price) of stocks using Long Short Term Memory which is a type of Recurrent Neural Network.

3.1)Rationale behind choosing LSTM

LSTMs are an improved version of recurrent neural networks (RNNs). RNN can only connect recent previous information and cannot connect information as the time gap grows. This is where LSTMs come into play; LSTMs are a type of RNN that remember information over a long period of time, making them better suited for predicting stock prices.



Source - https://colah.github.io/posts/2015-08-Understanding-LSTMs/

3.2)Explanation on working of LSTM

1. When certain parts of the cell state should be replaced with more recent information is determined by the forget gate. It produces values that are near to 1 for cell state components that should be kept and zero for values that should be ignored.

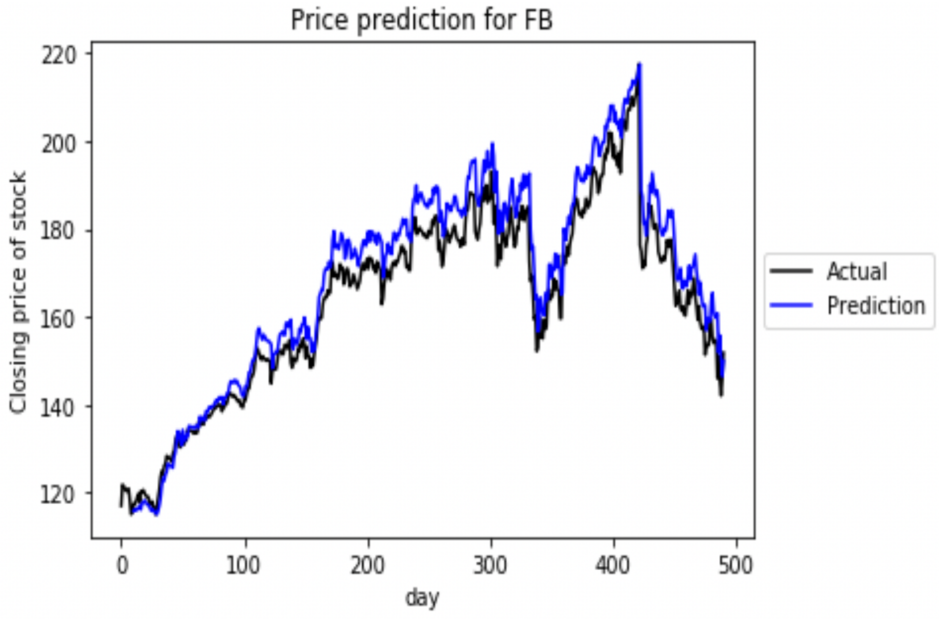
2. Input gate: Using the input (prior output o(t-1), input x(t), and previous cell state c(t-1), this part of the network learns the circumstances under which any information should be saved (or changed) in the cell state.
3. Output gate: Depending on the input and cell state, this part determines what information is forwarded to the following node in the network (i.e., output o(t) and cell state c(t)).

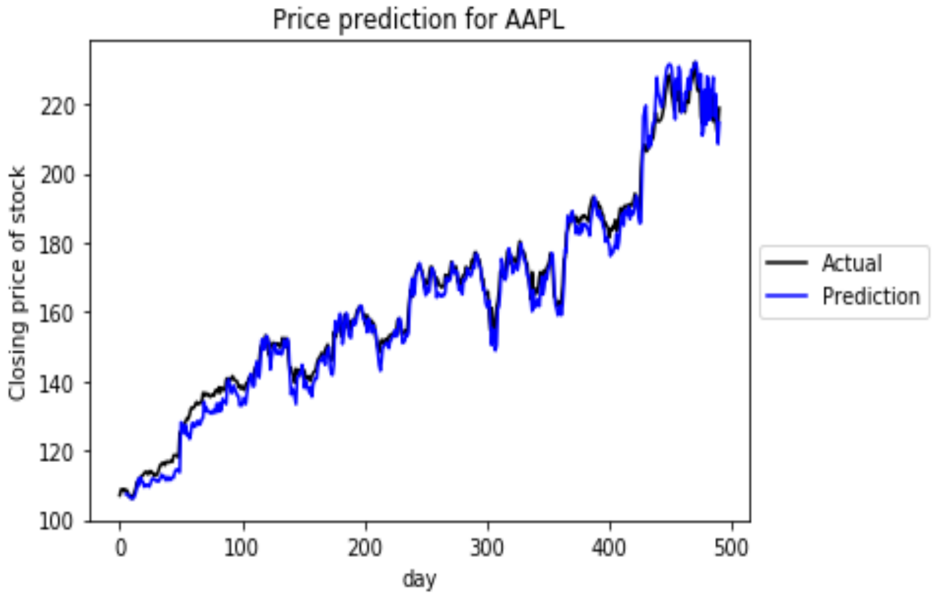
Let's take an example: Suppose the Stock price of NASDAQ changes every minute. So LSTM will save every minute's data on its nodes, using 10 previous nodes i.e last 10 minutes data of data will be used to predict the stock price.

4) Brief about the model

The project is focused on predicting the price of 32 stocks. The dataset is of 2 years for all the companies. For my model I am focusing on only one KPI that is end-of-day close price. Predicting the closing price provides useful information and would help the investor make the right decision.
I have considered the window size of 5, 7 and 10 in my model and got the best result with the window size of 10.
Window size 5 Training error – 0.0035, Testing Error 0.0037
Window size 7 Training error – 0.0042 Testing error – 0.0038
Window size 10 Training error – 0.0031 Testing error – 0.0038
To get the best results for each company we have further selected the model with the best window size which gave the least training error pertaining to that stock.

5)Graphical Presentation:





6) Conclusion & Contribution: The model is predicting end of the day stock price using LSTM networks with less than 2% accuracy. Although this model would help the investors to make high gains in the stock market due the high prediction rate. But at its core the stock markets are based on the sentiment of humans. Only crunching numbers and technical analysis will always have limitations. One possible extension of this model would be introduction of sentiment analysis with LSTM to further increase the accuracy and make the model full proof.

Appendix;

Source Code;
https://github.com/YashBhanushali9/Advance-Machine-Learning/tree/main/Final%20Project

References;

https://cs230.stanford.edu/projects_winter_2020/reports/32066186.pdf

https://www.sciencedirect.com/science/article/pii/S2666827022000378

https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0227222