Stock Prediction Overview and a Simple LSTM based Prediction Model

Shyam R[1], Vinayak Patil[2]

[1]*Student, Dept. of Computer Science & Engineering, The National Institute of Engineering, Mysuru, India*
[2]*Student, Dept. of Computer Science & Engineering, The National Institute of Engineering, Mysuru, India*

---***---

Abstract - *With rapid increase in computational efficiency which facilitates the functioning of deep learning algorithms, there is a surge in the number of areas practitioners tend to apply these thoughtfully crafted mathematical techniques. Applications of deep learning in the 21st century are in highly versatile and dynamic environments and one such environment is that of stock prediction. Stock prediction helps people make wise decisions about the investments they should go for, or the number of stocks to trade by providing a completely mathematical and reliable forecast of the stock attributes such as closing price, market trends etc. It is estimated that 1.4 billion shares are traded each day world-wide. Deep neural networks, mainly Recurrent Neural Networks (RNNs) when coupled with the ability to comprehend long term dependencies form Long-Short Term Memory (LSTM) networks, which are some of the widely used algorithms in the task of stock forecasting. In this paper we will discuss profusely about a proposed LSTM based implementation for prediction of closing prices of two companies with data obtained from NASDAQ. We also briefly give insights into the world of stock prediction and the sophisticated terminology and theories that exist.*

Key Words: Accuracy in predictions, Deep Neural networks Efficient Market Hypothesis, Long-Short Term Memory, Random Walk, Stock prediction.

1. INTRODUCTION

Better forecasting is the key element for better financial decision making given the increasing financial market volatility and internationalized capital flows[1].

For ages, the stock market has attracted investors, traders and economists due its complexity, uncertainty and vagueness. This also means that stock market trends cannot be assessed easily based on just a single factor or just based on historical data, as there are many intricacies involved in the fluctuation of stock prices.

For this reason, economists and experts suggest the use of various parameters based on the current public opinion or sentiment[2]. This is where twitter-sentiment analysis and other such methods play a vital role in the prediction. Although, there is no algorithm which can predict the stock prices with 100% accuracy, the objective of using deep learning is to increase the accuracy of prediction to a considerable extent with minimum data, so that traders can accordingly exercise their investment choices[3].

Statistical machine learning algorithms such as linear regression or clustering heavily rely upon historical data for the prediction. However, it is evident that a simple linear curve is not going to be an ideal prediction for the multivariate nature of the market [4], [5]. This calls for complex higher degree polynomial functions with multiple features and thus the use of an Artificial Neural Network comes into the picture. Due to the easy availability and feasibility of Distributed computing platforms and virtual hardware tools it becomes easier to train and validate deeper neural networks with multiple features.

In the remainder of this paper we will talk about the (features other than historical data) that affect the stock market trends and about the theories that govern the market. Some Terminology and Factors are also introduced briefly for the benefit of the readers who are unaware of such jargons. We then illustrate the deep learning based experiment we carried out and justify beforehand the usage of a particular algorithm i.e. Multi-variate LSTM. LSTM models are commonly used for stock prediction. We examine in detail the results and conclude by restating our findings.

2. An Overview of the Stock Prediction Scenario

Stock prediction is a major tool in helping investors to make rational decisions about the stocks they invest in. In a way they can make or break the economy. In this section we give a brief overview about the prevalent stock prediction scenarios. According to [6], Fundamental analysis and technical analysis are used to research and forecast the price trend of the stock in future.

1)Fundamental Analysis : Studies all those factors which have an impact on the stock price of the company in the future, such as financial statements, management process, industry, etc [6].
2)Technical Analysis : is a method that involves recognizing patterns and trends to forecast the price move-ments of the entity in the future. As pointed out by [7], while applying Machine Learning to Stock Data, we are more interested in doing a Technical Analysis to see if our algorithm can accurately learn the underlying patterns in the stock time series. Moreover, a hybrid implementation involving randomized techniques helps in efficient stock price forecasting and prediction. For the benefit of the reader we elicit two major definitions as an introduction to the stock prediction environment.

2.1 The Efficient Market Hypothesis (EMH)

EMH is an investment theory. Considering Risk -adjusted excess returns as we can theoretically assume that neither technical nor fundamental analysis can produce consistently. According to the EMH, stocks are always traded at their fair value on stock exchanges. Hence making it impossible for investors to either purchase undervalued stocks or sell stocks for inflated prices [8]. There are three forms of EMH:-

- Weak Form EMH
- Semi-Strong Form EMH
- Strong Form EMH

2.2 Random Walk

According to [8] Random walk theory suggests that changes in stock prices have the same distribution and are independent of each other. Random walk theory claims that stocks take a random and unpredictable path that makes all methods of predicting stock prices futile, contrary to the EMH theory.

3. Influential Factors

The unstable stock prices make equity investing risky. While the risk seekers invest eagerly in equity for achieving higher returns, the fluctuations enable real price discovery. The risk seekers only get fair returns for the higher risk. Each news or update results in fluctuation in the stock prices. Good news boosts investor sentiments. Whereas any setbacks decrease their hopes thus causing a price decline[9]. Some factors are firm-specific like a change in management, new product launch, policy change etc. Other factors are political, economic and technological that cause variation in stock prices. If the company becomes more valuable, then its stock prices increase. Any event that erodes company values causes stock prices to sink.
Here is an In-depth look into determinants of stock prices:

3.1 Company news and performance

Some company-specific factors which affect stock prices are:-
a) news releases on earnings and profits, and future estimated earnings
b) announcement of dividends
c) introducing a new product
d) accounting errors or scandals
e) trade performance

3.2 Investor sentiment

Investor sentiment or confidence will cause the market to go up or down, which might cause stock prices to rise or fall. The two terms used to describe market conditions are:

Bull market – a strong stock market where stock prices are rising and investors' confidence is growing. It's often related to economic recovery or an economic boom[7].

Bear market – a weak market where stock prices are decreasing and investor's confidence is fading. It happens When an economic recession and unemployment is high, with rising prices.

3.3 Economic Factors

3.3.1 Interest rates

If a corporation borrows cash to expand and improve its business, higher interest rates can have an effect on the value of its debt. This could scale back company profits and also the dividends it pays shareholders. As a result, its share worth might drop[10].

3.3.2 Economic outlook

The stock prices may increase, if the economy is going to expand. Investors may buy more stocks assuming, will see future profits and higher stock prices and investors may reduce their buying or start selling if the economic outlook is uncertain.

3.3.3 Inflation

It is the rise in the price of goods and services. Rising inflation has an adverse effect: input prices are higher, customers can purchase fewer goods, revenues and profits decline, and the economy slows down for a time until a measure of economic equilibrium is reached[11].

3.3.4 Deflation

Deflation is a condition where a country experiences lowering prices. This is the opposite of inflation, which is characterized by increasing prices. Deflation may help customers in terms of short-term affordability of goods and services in the market, but it has an adverse macroeconomic impact on stock markets.

3.3.5 Economic and political shocks

Changes around the world can affect both stock costs and the economy. For example, an increase in energy costs can lead to lower sales, lower profits and lower stock prices. An act of terrorism can also lead to a decrease in economic activity and a fall in stock prices.

3.3.6 Changes in policy

If a new government comes into power, it may build new policies. Sometimes these changes are good for business, and sometimes not. They'll result in changes in inflation and interest rates, that successively might have an effect on stock prices[12].

3.3.7 Return On Equity (ROE)

ROE = Net_income/Shareholders_equity
Shareholders' equity is equal to the company's assets minus its debt. ROE could be thought as return on net

assets.

3.3.8 Return on Assets (ROA)

ROA = Net_Income/Average_Total_assets
ROA is a financial ratio which gives the percentage of profit a company earns in relation to its overall resources.

4. Why LSTMs ?

Here we give a brief overview of Long-Short Term Memory(LSTM) and also justify its effectiveness in stock prediction. LSTMs are a special manifestation of Recurrent Neural Network units (or RNNs). Unlike traditional neural networks, RNNs are networks with loops in them, allowing information to persist. One of the appealing factors of RNNs is the idea that they are able to connect previous information to the present task, like using previous video frames might inform the understanding of the present frame[13].

RNNs suffer from the vanishing gradients problem. As the gap between two units widens, the weight of the gradient has exponentially lesser impact.Thus it renders them unable to resolve "Long-term dependencies". LSTMs help in resolving the Long-term dependency issue[1].

All RNNs have the form of a chain of repeating modules of neural networks. LSTMs also have this chain like structure, but the repeating modules have a different structure. Instead of having a single neural network layer, there are four, which are mainly composed of multiplexed sigmoid & hyperbolic tangent (tanh) functions interacting in a very special way. The idea is to let every step of an RNN pick information to look at from some larger collection of information[15].

The key to LSTMs is the cell state. Remembering information for long periods of time is practically their default behavior.

The fact that LSTMs have cell states which depend on previous data help them predict trends much more effectively than say a simple RNN or algorithms like Linear Regression and so on. It also helps bring context into consideration which help in factoring in several properties that affect stocks as listed in Section 3. Thus the LSTM approach is an ideal candidate to experiment with for Stock Prediction. In the next section we discuss more about the experimental details of our implementation of LSTMs for stock prediction. Note that we don't argue this to be the best implementation. The sole purpose of the experiment is to highlight the pros and cons of using LSTMs in the stock prediction realm.

5. LSTMs as a Prediction Model

In this section, we discuss how the Long-Short Term Memory algorithm or LSTM is used for the task of stock prediction. We used the data scraped from NASDAQ to predict the closing prices of two stocks Amazon (AMZN) and Google (GOOG). (The code for our experiments can be found at *https://gitlab.com/shyam.9201.08).*

In order to depict the advantages and shortcoming of an LSTM implementation we ran the AMZN stock prediction with a severely low data set that is about 1500 days' closing price data to predict the prices for the next 30 days and for the GOOG closing price predictions we used a more extensive 14000 days' data set to predict the next 1400 days' stock values. Our model ideally depicts the problems an RNN faces with scarcity of data and how it considerably overcomes this when approximately 10 times more data is provided. We have omitted considering the variety of influential factors now for the sake of simplicity.

5.1 For Amazon closing price (AMZN)

Preprocessing:
For initial data segregation, we alienated the Closing price from all the other available stock prices such as Opening price, High, Low and Volume. Next we converted and sorted the data into a time series, with the timestamp being the days sorted by the date. The data was then split into training and validation sets in the ratio of 9:1 (90% was used to train the model due to lack of data)

LSTM model:
To keep this a simple experiment, we chose 5 LSTM units with the output layer having the sigmoid activation function and all the inner layers use the Rectified Linear Unit (or ReLU) as activation with dropout regularization to reduce variance (preventing overfitting). RMSprop was used as the optimizer in our example for reasons as stated in [16].

Training & Results:
We chose a batch size of 10 (The network updates its weights after processing 10 days of stock price) and a time step of75 (i.e. the network looks back 75 days before predicting the value of the next day).

The model loss obtained is depicted in the following figure. A minimized loss of 0.00085 was what we obtained as the best result of the experiment.

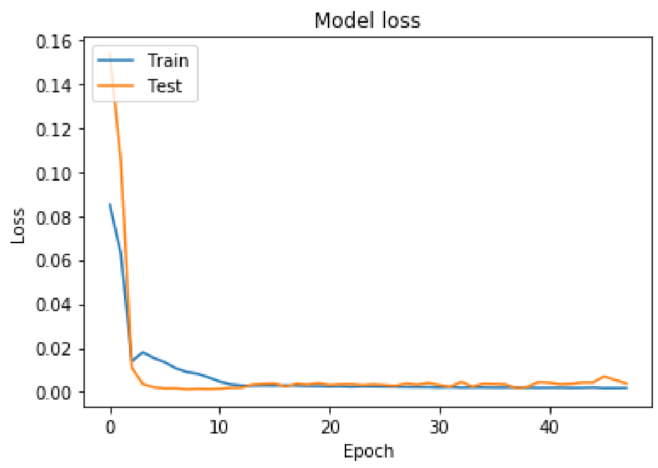

Figure -1: Model Loss for AMZN

It turns out that even with considerable hyper-parameter tuning the LSTM performance was poor and severely underfitting in predicting the next 30 days' closing prices.

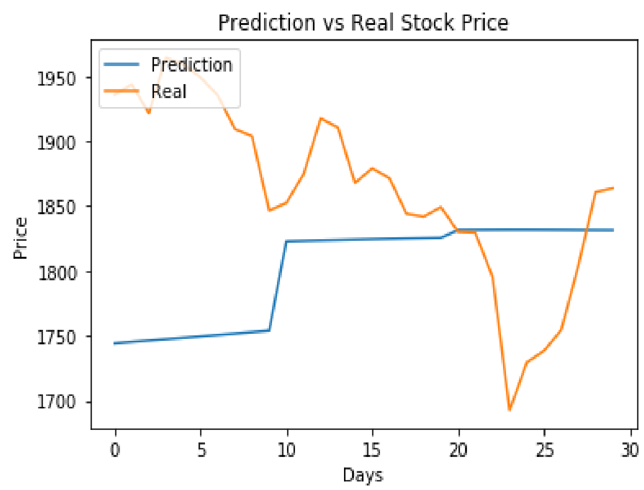

Figure -2: Predicted vs Real Stock Price for AMZN

The reader is welcome to implement their own version and maybe remove some of the dropout layers (we used two such layers with a dropout of 0.4 and 0.2) or increase the number of training epochs to hopefully find better results but we found that this was sufficient for portraying our ideas.

5.2 For Google closing price (GOOG)

Preprocessing:
The preprocessing done was basically the same as that given in section 5.1. However, we had about 10 times more data to process than that of the AMZN experiment.

LSTM model:
We used the same LSTM model again as that of AMZN and this was done so that the only variation with this experiment and that is the number of training examples which will effectively convey the performance dependence of LSTMs on amount data. But here we used 60% data for training and 40% for Validation.

Training & Results:
Batch size used here was 20, the Time step being 60. The Optimizer again being RMSProp with Dropout Regularization recommended by [17].

Model Loss was obtained as:

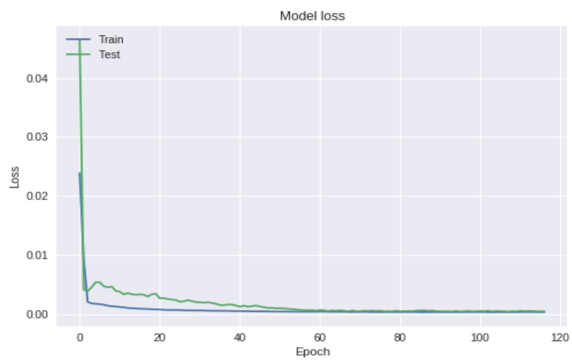

Figure -3: Model Loss for GOOG

Here the minimum loss obtained was 0.0018, relatively higher than that of the AMZN experiment but we find that the predictions for the next 1355 days relative to the actually obtained prices were:

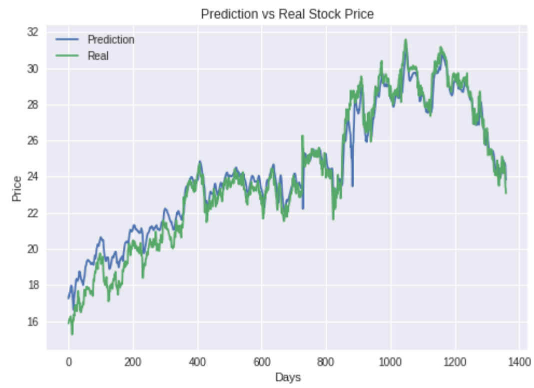

Figure -2: Predicted vs Real Stock Price for GOOG
Thus, with larger data the same LSTM performs much better, which is precisely our point.

5.3 Inferences

This is also an indicative of their high reliability on data and ability to fit complex nonlinear functions with relative ease. We observed that, with 10 times more data the LSTM performs a very good job of eliminating high bias. High data dependency is also a downside for LSTM because it might be difficult to obtain close to 1TB of data and securely storing them, also considering the disk access time needed for such huge amounts of data, it severely jeopardizes the performance of an LSTM. Using sentiment analysis (what we also omitted here) coupled with several influential factors in parallel can help create better and more accurate multi-variate LSTMs. Interested readers can find more complex experiments in [18] and [19]. We also verified that our results are consistent with the findings made by [20].

We summarize our results in a table below for simplicity.

Table -1: Performance of Experimental LSTM

Stock	Batch Size	Time step	Optimizer	Dropout	Min Loss	Overall Accuracy
AMZN	10	75	RMSProp	Two Layers (0.4,0.2)	0.00085	52.23%
GOOG	20	60	RMSProp	Three Layers (0.4,0.4,0.4)	0.0018	89.44%

3. CONCLUSIONS

Our intent in this paper was to provide a simple yet comprehensive view into the stock prediction world. We first discussed the various Analysis techniques and theories that view stock prediction from different angles. We further introduced some terminology and listed an array of factors important to consider for stock prediction to familiarize the reader into the field.

We demonstrated through a simple experiment how a deep learning algorithm i.e. LSTM performs in the scenario of stock prediction and we also outline the importance of the Long term dependency issues (important in the case of stock predictions) that are addressed by LSTMs. An efficient implementation of this algorithm requires extensive tuning and tests.

The experiment also highlighted the high data-dependence of RNNs and LSTMs and we justified our argument by first providing severely low data and next by providing almost 10 times as much data, and we saw a remarkable improvement in accuracy.

We reiterate that our aim through this paper was to provide insights into stock prediction, highlight it's current scenario and also study a simple implementation of a deep learning algorithm in this regard. Readers are free to add on to our model and experiment with different optimization algorithms, time-steps or batch-sizes. Our intent was just to show a simplified model for easy understanding.

Stock Prediction has multi-fold applications. Almost every smart investment involves a considerable extent of trend monitoring and predictions. Hence, experimental analysis of several techniques and algorithmic research in this regard will continue to exist. For future work, We would like to incorporate market simulation, consider several other major stock influential factors for reliable and effective predictions.

REFERENCES

[1] E. Guresen, G. Kayakutlu, and T.U. Daim, "Using artificial neural network models in stock market index prediction," Expert Systems with Applications, vol. 38, no. 8, pp. 10389–10397, 2011.

[2] G. Gidofalvi and C. Elkan, "Using news articles to predict stock price movements," Department of Computer Science and Engineering, University of California, San Diego, 2001.

[3] A. Barbu and N. Lay, "An introduction to artificial prediction markets for classification," Journal of Machine Learning Research, vol. 13, no. Jul, pp. 2177–2204, 2012.

[4] J. Agrawal, V. Chourasia, and A. Mittra, "State-of-the-art in stock prediction techniques," International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, vol. 2, no. 4, pp. 1360–1366, 2013.

[5] J. Siegel, Stocks For The Long Run/Jeremy J. Siegel, vol. 37. New York: McGraw-Hill, 1998.

[6] V. H. Shah, "Machine learning techniques for stock prediction," Foundations of Machine Learning| Spring, vol. 1, no. 1, pp. 6–12, 2007.

[7] "https://keydifferences.com/difference-between-fundamental-and-technical-analysis.html."

[8] "https://www.investopedia.com/terms/e/efficientmarkethypothesis.asp."

[9] X. Zhang, Y. Hu, K. Xie, S. Wang, E. Ngai, and M. Liu, "A causal feature selection algorithm for stock prediction modeling," Neurocomputing, vol. 142, pp. 48–59, 2014.

[10] A.Cowles, "Stock market forecasting," Econometrica, Journal of the Econometric Society, pp. 206–214, 1944.

[11] J. M. Dalton and J. Dalton, How the stock market works. New York Institute of Finance, 1988.

[12] C. S. Eun and S. Shim, "International transmission of stock market movements," Journal of financial and quantitative Analysis, vol. 24, no. 2, pp. 241–256, 1989.

[13] "http://colah.github.io/posts/2015-08-understanding-lstms/."

[14] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[15] "http://karpathy.github.io/2015/05/21/rnn-effectiveness/."

[16] M. C. Mukkamala and M. Hein, "Variants of rmsprop and adagrad with logarithmic regret bounds," in Proceedings of the 34th International Conference on Machine Learning-Volume 70, pp. 2545–2553, JMLR. org, 2017.

[17] W. Zaremba, I. Sutskever, and O. Vinyals, "Recurrent neural network regularization," arXiv preprint arXiv:1409.2329, 2014.

[18] R. Akita, A. Yoshihara, T. Matsubara, and K. Uehara, "Deep learning for stock prediction using numerical and textual information," in 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS), pp. 1–6, IEEE, 2016.

[19] X. Ding, Y. Zhang, T. Liu, and J. Duan, "Deep learning for event-driven stock prediction," in Twenty-fourth international joint conference on artificial intelligence, 2015.

[20] R. Singh and S. Srivastava, "Stock prediction using deep learning," Multimedia Tools and Applications, vol. 76, no. 18, pp. 18569–18584, 2017.