STOCK PRICE PREDICTION

[1]Dr. V BIKSHAM

Assiociate Professor *in Sreyas Institute of Engineering and Technology, JNTUH, India.*

[2]Mr. G VENU

B.Tech in Sreyas Institute of Engineering and Technology, JNTUH, India.

[3]Mr. B VISHAL KUMAR

B.Tech in Sreyas Institute of Engineering and Technology, JNTUH, India.

[4]Mr. C RAHUL

B.Tech in Sreyas Institute of Engineering and Technology, JNTUH, India.

[5]Mr. M BHARGAV SAI

B.Tech in Sreyas Institute of Engineering and Technology, JNTUH, India.

ABSTRACT *-*

Machine learning has many important applications in the stock price prediction. Here, we will discuss about predicting the returns on stocks. This has uncertainties and it is a very complex task. This project will be developed into two parts: First, we will learn how to predict stock price using the Long Short-Term Memory neural networks. Predicting stock market prices involves human-computer interaction. For stock market analysis, conventional batch processing methods cannot be utilized efficiently due to the correlated nature of stock prices. We suggest an algorithm that utilizes a kind of recurrent neural network (RNN) called Long Short-Term Memory (LSTM), where using stochastic gradient descent the weights are adjusted for individual data points.

1. INTRODUCTION

In stock market, there are large number of investors and traders who buy and sell stock, pushing the price up or down. The prices of stocks are made by the principles of demand and supply, and the final goal of purchasing shares is to earn money by buying stocks in companies whose value is expected to rise. Stock markets are linked with the economics, the rise and fall of prices can be traced back to some Key Performance Indicators or KPI's. The commonly used KPI's are the opening stock price, end of day price, intraday low price, intraday peak price, and total volume of stocks that are traded during the day. Stock prices and economics are mainly depends upon perceptions about the stock market. We can say that it is nearly impossible to predict stock prices, owing to the elasticity of factors that play a vital role in the movement of prices. However, it is possible to make an estimate of prices. Stock prices never vary in isolation. This type of stock price movement can be used as an important tool to predict the prices of many stocks at once. Due to the large volume of money involved and number of transactions that takes place in each minute, there comes a trade off between the accuracy and the volume of predictions made; as such, most prediction systems are implemented in a distributed, parallelized fashion.

1.1. Problem Statement

Stock market prediction is defined as trying to determine the stock value and offer an idea for everyone to know and predict the market and the stock prices. Generally, it is presented using the quarterly financial ratio using the dataset. So, relying on a single dataset is not sufficient for the prediction and can give a result which is inaccurate. Hence, we are regarding towards the study of machine learning with various integration of datasets to predict the market and the stock trends. If a better stock market prediction algorithm is not proposed, then the problem with estimating the stock price will remain a problem. Predicting how the stock market will perform is very difficult. Stock market prediction is an ability to predict the effect of recent events on the investors. Those events can be political events like a statement of a political leader, a piece of news on scam etc. These type factors make stock price prediction very difficult. Once the right data is collected, then it can be used to train a machine and to generate a predictive result.

2. LITERATURE SURVEY

Initially, the focus of our literature survey was to explore generic online learning algorithms and see if they can be adapted to our use case that is, working on real-time stock price data. This includes Online AUC Maximization, Online Transfer Learning, and Online Feature Selection. But we were not able to find any possible adaptation for stock price prediction, so we decided to look at the existing systems and analyze the major drawbacks, and see if we could improve upon them. We zeroed in on the correlation between stock data as the key issue that we wished to solve. After searching generic solutions to the above problem, it led us to RNN's and LSTM. So, after deciding to use an LSTM neural network to perform stock prediction, we searched a number of papers to study the concept of gradient descent and its various types. After this we concluded our literature survey by seeing how gradient descent can be used to tune the weights of an LSTM network and how this process can be optimized.

3. EXISTING SYSTEM

The traditional approaches to stock price prediction and stock market analysis includes fundamental analysis, which looks at past performance of a stock and the general credibility of the company and statistical analysis, which is solely concerned with number of identifying patterns in stock price variation. It is commonly achieved with the help of Genetic Algorithms (GA) or Artificial Neural Networks (ANN's), but these will fail to capture correlation between stock prices in the form of long-term temporal dependencies. Another issue with using simple ANNs for stock prediction is the phenomenon of (exploding / vanishing) gradient, where the weights of a large network either become too large or small, exceptionally slowing their convergence to the optimal value. This is caused by two factors: weights that are randomly initialized, and the weights closer to the end of the network that also tend to change more than those at the beginning. An alternative approach to stock market analysis is to reduce the dimensionality of input data and apply feature selection algorithms to shortlist a set of features that have biggest impact on stock prices across markets. But this method does not consider long term trading strategies as it fails to take the entire history of trends into account; and there is no provision for outlier detection.

4. PROPOSED SYSTEM

In proposed system, we propose an algorithm for predicting the end of day price of a given stock with the help of Long Short Term Memory (LSTM), a type of Recurrent Neural Network (RNN). LSTM's are subset of RNN's which can capture context specific temporal dependencies for long periods of time. Each LSTM neuron is a memory cell that can store other information, that means it maintains its own cell state. While neurons in normal RNN's merely take in their previous hidden state and current input to output a new hidden state, an LSTM neuron also takes in its old cell state and outputs its new cell state. An LSTM memory cell has three components, or gates:

1. Forget gate: it decides when specific portions of the cell state are to be replaced with most recent information. It outputs values close to one for parts of the cell state that should be retained, and zero for values that should be neglected.
2. Input gate : based on the input, this section of the network learns the conditions under which any information should be stored in the cell state.
3. Output gate: depending on the input and cell state, this portion decides what information is propagated forward to the next node in the network.

So, we can say that LSTM networks are perfect for seeing how variation in price of a stock can affect the prices of several other stocks over a long period of time.

5. RESULTS AND DISCUSSIONS

```
df=pd.read_csv("NSE-TATA.csv")
df.head()
```

	Date	Open	High	Low	Last	Close	Total Trade Quantity	Turnover (Lacs)
0	2018-10-08	208.00	222.25	206.85	216.00	215.15	4642146.0	10062.83
1	2018-10-05	217.00	218.60	205.90	210.25	209.20	3519515.0	7407.06
2	2018-10-04	223.50	227.80	216.15	217.25	218.20	1728786.0	3815.79
3	2018-10-03	230.00	237.50	225.75	226.45	227.60	1708590.0	3960.27
4	2018-10-01	234.55	234.60	221.05	230.30	230.90	1534749.0	3486.05

```
df=pd.read_csv("INFY.csv")
df.head()
```

	Date	Symbol	Series	Prev Close	Open	High	Low	Last	Close	VWAP	Volume	Turnover	Trades	Deliverable Volume	%Deliverble
0	2000-01-03	INFOSYSTCH	EQ	14467.75	15625.00	15625.20	15625.00	15625.20	15625.20	15625.18	5137	8.026657e+12	NaN	NaN	NaN
1	2000-01-04	INFOSYSTCH	EQ	15625.20	16800.00	16875.25	16253.00	16875.25	16855.90	16646.38	56186	9.352937e+13	NaN	NaN	NaN
2	2000-01-05	INFOSYSTCH	EQ	16855.90	15701.00	16250.00	15507.45	15507.45	15507.45	15786.38	164605	2.598516e+14	NaN	NaN	NaN
3	2000-01-06	INFOSYSTCH	EQ	15507.45	15256.65	15300.00	14266.85	14266.85	14266.85	14462.82	81997	1.185908e+14	NaN	NaN	NaN
4	2000-01-07	INFOSYSTCH	EQ	14266.85	13125.50	13125.50	13125.50	13125.50	13125.50	13125.50	7589	9.960942e+12	NaN	NaN	NaN

Fig 2- Data Sets for NSE TATA and INFY

Fig 3- Closing Prices for NSE TATA DataSet

Fig 4- Closing Prices for INFY DataSet

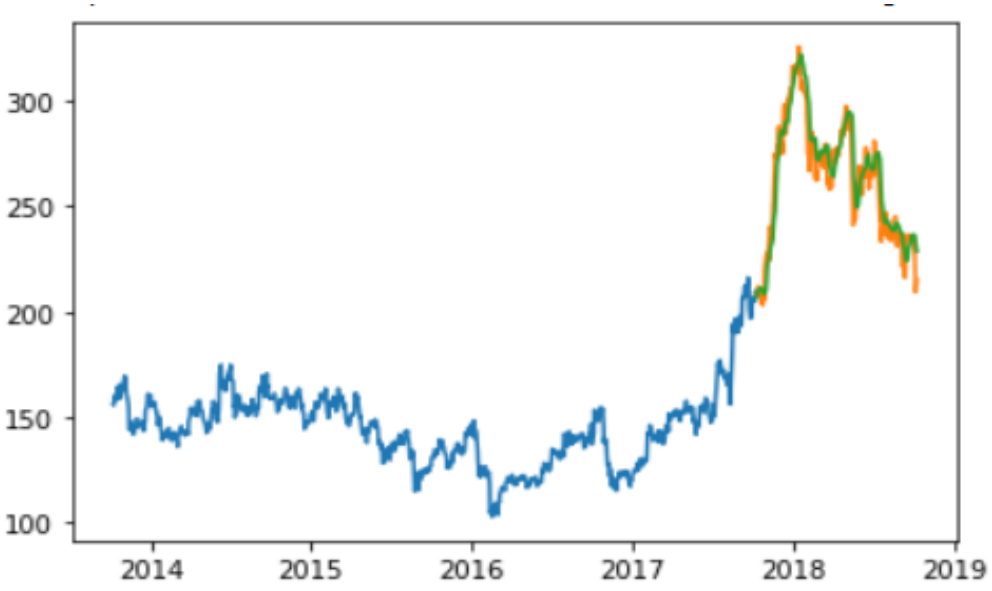

Fig 5- Prediction Prices for NSE TATA DataSet

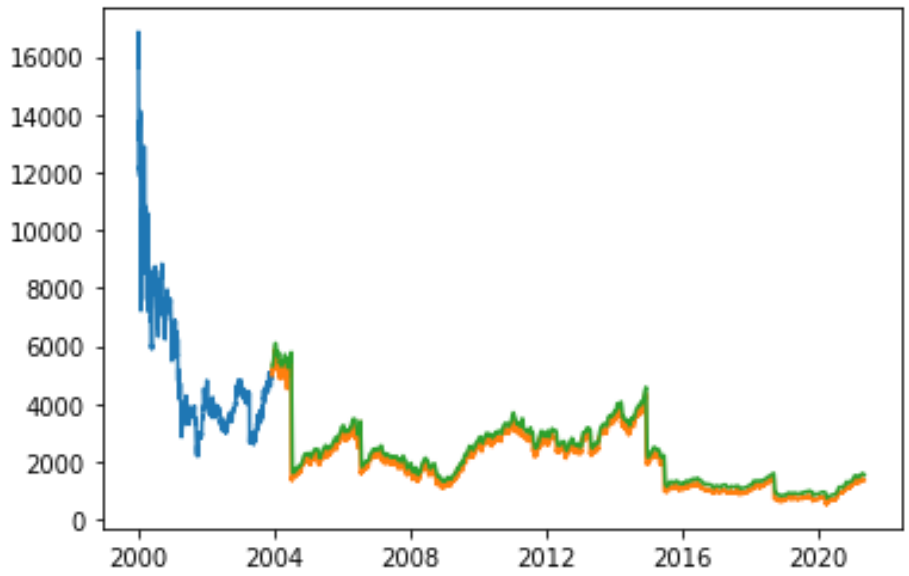

Fig 6- Prediction Prices for INFY DataSet

We can observe that LSTM has predicted stocks almost similar to actual stocks. LSTM is very useful in the future predictions of stocks. LSTM's are strong in problems like sequence prediction because they're able to store past information. In our case this is important because the previous price of a stock is important in predicting its future price. A stock market is a platform for trading of a company's stocks at an agreed price. Supply and demand of shares runs the stock market. In any country we can say that the stock market is one of the most emerging sectors. Now a days, most people are indirectly or directly related to this sector. Therefore, it is very important to know about market trends. So, with the development of the stock market, many people are interested in forecasting stock prices. But, due to dynamic nature and likely for quick changes in stock price, prediction of the stock price becomes a challenging task.

6. CONCLUSION AND FUTURE SCOPE

The results of comparison between Long Short-Term Memory (LSTM) and Artificial Neural Network (ANN) shows that LSTM has better prediction accuracy than ANN. Stock markets are hard to control and require lots of context when we are trying to interpret the movement and predict prices. In ANN, each hidden node is a node with a single activation function, while in LSTM, each node is a memory cell, that can store contextual information. LSTM's perform better as they are able to keep track of context specific temporal dependencies between stock prices for a longer period of time while performing predictions. An analysis of these results also indicates that both models give better accuracy when dataset size increases. With more data, more patterns can be developed by the model, and the weights of the layers can be adjusted. At its core, we can say that the stock market is simply a reflection of human emotions. Pure number processing and analysis have their own limitations; a possible extension of this stock prediction system is to increase it with a news feed analysis from social media platforms, where emotions are measured from the articles. This sentiment analysis can be linked with LSTM to better train weights and improve accuracy.

6.1. Future Scope

Stock market prediction means finding the future scope of market. A system which is essential to be built will work with maximum accuracy and it should consider all important factors that could affect the result. Predicting the performance of a stock market is hard as it takes into account to various factors. In future, this project involves adding more parameters and factors like the multiple instances, financial ratios, etc. If more parameters are considered then accuracy will be more. The algorithms can be applied for analyzing the contents of public comments and thus determine patterns or relationships between the customer and the corporate employee as well.

7. REFERENCES

[1] *https://www.kaggle.com/faressayah/stock-market-analysis-prediction-using-lstm*

[2] *https://www.analyticsvidhya.com/blog/2018/10/predicting-stock-price -machine-learningnd-deep-learning-techniques-python/*

[3] *https://machinelearningmastery.com/gentle-introduction-long-short-te rm-memory-networks-experts/*

[4] *https://data.nasdaq.com/data/NSE/TATAGLOBAL-tata-global-beverages-limited*

[5] *Suresh, Harini, et al. "Clinical Intervention Prediction and Understanding using Deep Networks."*

[6] *Pascanu, Razvan, Tomas Mikolov, and Yoshua Bengio. "On the difficulty of training recurrent neural networks."International Conference on Machine Learning.*

[7] *Zhu, Maohua, et al. "Training Long Short-Term Memory With Sparsified Stochastic Gradient Descent."*

[8] *Ruder, Sebastian. "An overview of gradient descent optimization algorithms.".*

[9] *Recht, Benjamin, et al. "Hogwild: A lock-free approach to parallelizing stochastic gradient descent." Advances in neural information processing systems.*

[10] *Ding, Y., Zhao, P., Hoi, S. C., Ong, Y. S. "An Adaptive Gradient Method for Online AUC Maximization"In AAAI.*

[11] *Altinbas, H., Biskin, O. T. "Selecting macroeconomic influencers on stock markets by using feature selection algorithms". Procedia Economics and Finance.*

[12] *A. AdebiyiAyodele, K. Ayo Charles., O. Adebiyi Marion., and O. Otokiti Sunday, "Stock Price Prediction using Neural Network with Hybridized Market Indicators," Journal of Emerging Trends in Computing and Information Sciences.*

[13] *T. AkinwaleAdio, O.T. Arogundade and F Adekoya Adebayo. "Translated Nigeria Stock Market Price Using Artificial Neural Network for Effective Prediction," Journal of Theoretical and Applied Information Technology.*

[14] *N. C. Ashioba, E.O. Nwachukwu, and O. Owolobi. "Finding the Optimal Solution for a Transportation Problem Using Neural Network. Microwave," International Journal of Science and Technology.*

[15] *Kyoung-Jae Kimand and Ingoo Han, " Genetic Algorithms Approach to Feature Discretization in Artificial Neural Networks for the Prediction of Stock Price Index," Institute of Science and Technology".*