Stock Prediction Model using Seq2Seq and Bi-directional LSTM

Akshat Bansal
Student,
Dept of CSE
R V College of Engineering
Bengaluru, India
akshatbansal.cs19@rvce.edu.in

Dency Narendra Patel
Student,
Dept of CSE
R V College of Engineering
Bengaluru, India
dencynarendrap.cs19@rvce.edu.in

Khetan Rishabh
Student,
Dept of CSE
R V College of Engineering
Bengaluru, India
khetanrishabhr.cs19@rvce.edu.in

Sneha M
Assistant Professor
Dept of CSE
R V College of Engineering
Bengaluru, India
sneham@rvce.edu.in

***Abstract*—The stock market is an exchange platform that provides a venue for investors, traders and institutions to buy and sell shares of listed companies. It acts as a surety between the buyer and lender for the smooth exchange of stocks on the accepted price. The stock market sets those accepted price using the mechanism of price discovery based on both fundamental and technical variables The entire idea of predicting stock prices is to gain significant profits. In this paper, a model has been put forward for the prediction of the Closing prices of the stock by using BLSTM based Seq2Seq Model. Seq2Seq model is used as it helps in the mapping of the input-sequence to the output-sequence. This proposed model was compared with various existing algorithms such as K-Nearest Neighbour, Decision tree and Linear Regression. The results shows that the Root Mean Square Error value for the proposed model was found to be least compared to other models.**

***Index Terms*—stock prediction, BLSTM, Seq2seq, Predicting**

I. Introduction

AI has a huge potential in the prediction of stock prices. Taking the past performance and behaviour of any stock and training the data available using neural networks and machine learning models can help in understanding how a stock might behave in the future. Industrially talking, the system would have huge relevance. It can be used by traders to gain an edge over others and can also be used by financial institutions for quant-vol trading.

A company that wishes to raise capital from the open market can do so by offering shares their shares on the exchange for people willing to buy. To do so it needs to offer its securities to the public through a process called Initial Public Offering (IPO) [1].

The Stock Market consists of 3 main components, namely a buyer, a seller and a trade [2]. The stock market is a composite framework where individuals can trade monetary forms, stocks, equity and subordinates over virtual stages upheld by specialists. The key fundamental factors in stock prices are the level of the earnings base (represented by measures such as EPS, cash flow per share, dividends per share), the expected growth in the earnings base, and the discount rate, which is itself a function of inflation, the perceived risk of the stock. Then there are the technical factors like inflation, substitutes, demographics, market sentiments, liquidity etc. Apart from the technical aspect, the stock market also depends on the sentimental aspects such as tweets or speeches from high Ranking officials from government or private organizations, or natural calamities or wars. Such aspects are nearly impossible to predict as they change people's opinion about businesses operating in that region leading them to sell or buy stocks without even considering its technical aspect causing sudden spikes or drops. [3].

Moreover the stock price movement depend not only on the history of individual stock movements, but also complex hidden dynamics associated with other correlated stocks [4]. As seen there are many factors that cause uncertainty in the stock market, Widely used is the automated trading system(ATS) that works with the help of computer codes which perform better and more precise readings on the stock market.

The objective of the proposed model is to predict the prices of the stocks observing the previous patterns in its prices by running is through a neural network.

II. Literature Review

Stock price prediction models can be implemented using various models (KNN, Decision tree, Linear regression etc.). In the following section, some of the esteemed work done in this field in past years has been discussed.

A model based on the ML algorithm Random forest Classifier and Support Vector Machine(SVM) was implemented by Sadia et al. [5]. It was discovered that their dataset was raw and needed to be preprocessed. They took the dataset from Kaggle. Hence, it was concluded that the dataset downloaded from Kaggle is raw and other websites needed to be explored for the dataset.

M. Asam et al. [6] presented a comparative study in which they evaluated the various algorithms after implementing Principle Component Analysis and concluded the result by comparing both solutions which resulted in KNN providing better results than Support Vector Machine and Naive Bayes.

Aparna et al. [7] proposed a model in which the traditional dataset is aggregated with news/tweets of a company. They implemented their model using this dataset and observed the company's stock trend on a monthly basis. They concluded that the Decision tree performs better on these types of datasets.

J. K. Hao et al. [8] used ANN algorithm to predict the stock price. They used the stock data of the Shanghai Stock Exchange of 1 month. The model was successful in predicting the stock values but only for a short span of time. S. Mohammed et al. [9] used a Bi-Directional LSTM model(BLSTM). They implemented the model for both short and long-term Stock Market Prediction. For their dataset they used SP 500 historical data and evaluated the results based on closing price. They compared the results with 3 different models - SLSTM, LSTM and MLP-ANN. Their model showed better efficiency than other models.

G. Y. Lee et al. [10] implemented a model using Sequence to Sequence model as it helps in the mapping of the input-sequence to the output-sequence in the network. They trained 3 models - Seq-seq, WaveNet and LSTM and used daily trading data and technical indicators. The seq to seq model showed good performance.

III. Proposed Model

The proposed model in the paper is Seq-2-Seq BLSTM which uses the LSTM recurrent neural network to predict the future prices based on the trends created by the the existing prices.

A. Dataset Description

In today's world, Stock prices are easily accessible on the internet. In the proposed model, historical data of Google(GOOG) from 2004 - May'20 [11] is chosen. The dataset was downloaded from Yahoo Finance. The dataset consists following features:

- Date - Day of the stock price
- Open - Price of the stock at market open
- High - Highest price reached in the day
- Low - Lowest price reached in the day
- Close - Price at the end of the day
- Volume - Number of shares traded

The four main attribute to all stocks are Open, High, Low And Close price (OHLC) [12]

B. BLSTM

BLSTM networks are a type of RNN networks that uses some additionl units with the standard units of RNN model which makes BLSTM networks more powerful. [13]

In this paper a Bi-LSTM based Encoder Decoder was implemented [14]. It is used to process the Seq-Seq predictions. Encoder and Decoder work simultaneously. The function of the encoder is to convert input of variable length to a vector of fixed length. Decoder decodes the output sequence with the help of input. Encoder is made up of BLSTM units, where each unit maps the features of stock data. These BLSTM units try to depict the link between past and future sequences. At any instance, window length sequence is given as input to Encoder. Window length is chosen in a way that minimizes the error in output prediction. The ouput vector is decoded using LSTM units. The decoder gives a single sequence as output.

C. Seq2Seq Modeling

This model is used as it maps the input sequences to the output sequences of data. This model uses encoder-decoder structure [15]. It does so by use of the recurrent neural network (RNN). Encoder reads the input sequence and summarizes the data to the context vector in case of LSTM. The results of the encoder were disposed of and just protect the inside states. This setting vector expects to epitomize the data for all information components to assist the decoder with making exact expectations.

The decoder is a LSTM whose underlying states are introduced to the last conditions of the Encoder LSTM, for example the setting vector of the encoder's last cell is a contribution to the primary cell of the decoder organization. The initial states of the decoder do generate output sequences and these sequences are taken into consideration for the further outputs. Figure 1 shows how the proposed model works.

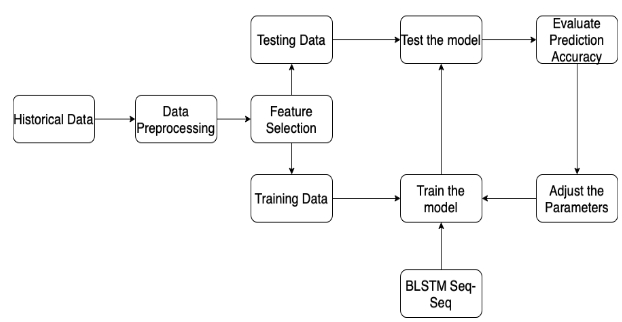

Fig. 1. Overall architecture of the proposed model

IV. IMPLEMENTATION

MacOS was used for the python implementation of the BLSTM model with the given hardware specification:

- macOS Monterey v 12.0.1
- Quad-core 10th-Generation Intel Core i5 Processor @ 2.0 GHz
- 16GB RAM, 512GB SSD Intel Iris Plus Graphics

For the experiment following libraries were used:

1) NumPy – v 1.11.0 2) TensorFlow – v 2.3.0 3)Matplotlib – v 3.1. 4)Keras – v 2.9.0 5)Pandas – v 1.4.2

A. Pre-processing Dataset

The dataset used has google Stock prices, which includes Date, Open, High, Low, Close, Adj Close and Volume of a day. The dataset was spilt into a ratio of 80:20 where 80 percent is used for training the models and 20 percent is used for testing the model. All Null values are removed from the dataset as they tend to disturb the actual trends. Using the MinMaxScaler of sklearn, the data is then normalized [16]. Normalization of data is transforming the data to appear on the same scale across all the records. For normalization first, the minimum value is subtracted from all the value in the dataset. Following this the maximum value is divided with all the value in the dataset. The minimum value in the array will always be normalized to 0 and the maximum value in the array will be normalized to 1. All the other values will be in the range between 0 and 1 [17]. MinMaxScaler is done for Open, Close, High and Low Price.

B. BLSTM Seq-seq Structure

The above pre-processed data is taken as input to the forward layer which is passed to the backward layer and further passed to the full connection layer. Predicted result is given as output.Initialization of biases and weights of full connection layer are done on the basis of normal distribution. BLSTM links LSTM in two opposite directions and extracts information from both layers simultaneously.

Fig. 2. Bi-LSTM Process

Then it makes use of seq-seq to return the output vector. A dense layer is implemented to map outputs to GOOGLE values of the output sequence. The following table shows the proposed model's layers implemented and the parameters taken into account for the training phase.

Table I is an example of LSTM architecture. Col1(Layer Type) tells about the layer used in the architecture(input layer, output layer, seq-seq layer). Col2 represent the ouput value of each layer. NULL is the placeholder for the batch size. Constraints defines how many numbers in the resulting tensor will be

C. Comparison with other models

To check the efficiency of the model, the result of the model is compared with other models i.e

1. Linear Regression [18] 2. Decision Tree [19] 3. KNN [20]

TABLE I
LSTM ARCHITECTURE

Layer(Type)	O/p Shape	Constraints
blstm(BLSTM)	(NULL, 50, 40)	10399
blstm$_1(BLSTM)$	(NULL, 50)	20201
layerDense	(NULL, 1)	1273a
layerDense1	(NULL, 1)	25

Root Mean Squared Error(RMSE) is used as a comparison factor. Root Mean Square Error is calculated by finding the the average error of the expected output to the actual output of the testing dataset and taking the root of that average value. The model with least RMSE value turns out to be the best. RMSE shows how well data is around the line of best fit, hence less RMSE value means better model. The formula for calculating RMSE is:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} i - O_i}{n}}$$

where n = total number of data-points; i = predicted output; $O_i = actualoutput.$

V. RESULTS AND ANALYSIS

The experimental result for this particular paper is a comparative study of the predicted prices obtained from the BLSTM to the other models mentioned and the actual price. After performing the series of experiments, the Root Mean Square Error is calculated [21] and the value of all the models using the parameters listed in Table II.

TABLE II
MODELS AND THEIR RMSE VALUE

Models	Constraints	RMSE
BLSTM	Adam(Optimizer) MSE(lossFunc)	22
Linear-Regression	alpha=0.2	55
Decision-Tree	Height(5)	58
KNN	kValue(13)	150

In the progression of examinations on the RMSE values of BLSTM predicted values at different epochs (iterations), it is concluded that the value of RMSE is the lowest at 200 epochs when compared to other values at various epochs as listed in Table III. Epoch values of both upper and lower end are considered to arrive at this decision and all the data will be iterated through the LSTM 200 times to provide the least Root Mean square error and hence the best-predicted value.

TABLE III
MODELS AND THEIR RMSE VALUE

Epochs	RMSE
50	26.94
100	23.86
200	21.83
300	31.37
400	49.21
500	25.21

A graph has been plotted comparing the predicted values obtained from all the different implemented models and the actual closing price recorded on a particular day, as seen in Figure 3.

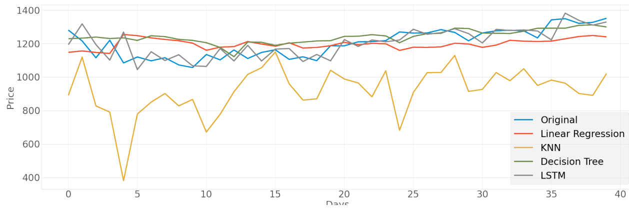

Fig. 3. Comparison with other algorithms

It is evident from Figure 3 that BLSTM has much a much higher accuracy as compared to the other models. The detailed comparison of the actual and predicted prices has been listed in Figure 4

Date	Actual Price	LSTM	Linear Reg.	Decision Tree	K-NN
2020-04-15	1262.469971	1249.831055	1178.453431	1144.758281	889.541354
2020-04-16	1263.469971	1234.953613	1178.701617	1198.114363	999.325331
2020-04-17	1283.250000	1235.863892	1180.725238	1198.114363	993.635690
2020-04-20	1266.609985	1256.876099	1202.507832	1198.114363	1081.082257
2020-04-21	1216.339966	1236.862427	1199.284338	1198.114363	906.231156
2020-04-22	1263.209961	1186.866699	1178.426143	1144.758281	841.254103
2020-04-23	1276.310059	1244.269531	1191.902488	1198.114363	1002.173795
2020-04-24	1279.310059	1252.695312	1219.459388	1144.758281	971.730004
2020-04-27	1275.880005	1252.447876	1216.085958	1198.114363	1063.193852
2020-04-28	1233.670044	1248.345337	1213.198386	1144.758281	893.822223

Fig. 4. Comparison with other algorithms

VI. CONCLUSION

As seen from the experimental results, the comparison of BLSTM prediction is plotted against the prediction of regression models such as Decision trees, K-Nearest Neighbor and linear Regression. The accuracy of the BLSTM is proved to be the highest consistently compared to the output of the other methods. The accuracy OF BLSTM is found to be 97.3% compared to others which stand at 93.6% for decision tree, 78% for KNN and 95% for Linear Regression. The project has a lot of scope in future and the topic and related research will always be in demand because the application of this has the power to control the flow of money which will keep the interest boosted. Hybrid models could be developed to get better accuracy and eliminate flaws that a single model produced. Apart from that, Natural Language Processing can also be used to view things from a sentimental analysis point of view.

REFERENCES

[1] S. V. Shenoy and K. Srinivasan, “Relationship of IPO Issue Price and Listing Day Returns with IPO Pricing Parameters,” International Journal of Management Studies, vol. V, no. 4(1), p. 11, Oct. 2018.

[2] C. Pop et al., “Decentralizing the Stock Exchange using Blockchain An Ethereum-based implementation of the Bucharest Stock Exchange,” in 2018 IEEE 14th International Conference on Intelligent Computer Communication and Processing (ICCP), pp. 459-466, 2018.

[3] M. C. Joshi, “Factors Affecting Indian Stock Market,” SSRN Electronic Journal, 2013.

[4] C. Li, D. Song, and D. Tao, “Multi-task Recurrent Neural Networks and Higher-order Markov Random Fields for Stock Price Movement Prediction,” in Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery Data Mining, pp. 1141–1151, 2019.

[5] Sadia, K. Hiba, Aditya Sharma, Adarrsh Paul, Sarmistha Padhi, and Saurav Sanyal. ”Stock Market Prediction Using Machine Learning Algorithms.” IJEAT (2019).

[6] Khan, W., M. A. Ghazanfar, M. Asam, A. Iqbal, S. Ahmad, and Javed Ali Khan. ”PREDICTING TREND IN STOCK MARKET EXCHANGE USING MACHINE LEARNING CLASSIFIERS.” Science International 28, no. 2 (2016).

[7] Nayak, Aparna, MM Manohara Pai, and Radhika M. Pai. ”Prediction models for Indian stock market.” Procedia Computer Science 89 (2016): 441-449.

[8] Y. Bing, J. K. Hao, and S. C. Zhang, “Stock Market Prediction Using Artificial Neural Networks,” Advanced Engineering Forum, vol. 6–7, pp. 1055–1060, Sep. 2012.

[9] K. A. Althelaya, E.-S. M. El-Alfy, and S. Mohammed, “Evaluation of BLSTM for short-and long-term stock market prediction,” in 2018 9th International Conference on Information and Communication Systems (ICICS), pp. 151-156, 2018.

[10] C.-H. Cho, G.-Y. Lee, Y.-L. Tsai, and K.-C. Lan, “Toward Stock Price Prediction using Deep Learning,” in Proceedings of the 12th IEEE/ACM International Conference on Utility and Cloud Computing Companion - UCC ’19 Companion, pp. 133–135, 2019. .

[11] GOOG(Google) Stock prices from Yahoo! Finance[Online] Available: https://finance.yahoo.com/quote/GOOG/history?p=GOOG

[12] D. Wei, “Prediction of Stock Price Based on LSTM Neural Network,” in 2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM), pp. 544-547, 2019

[13] Huicheng Liu Department of Electrical and Computer Engineering Queen’s University, Canada - Leveraging Financial News for Stock Trend Prediction with Attention-Based Recurrent Neural Network.

[14] Shengdong Du, Tianrui Li, Yan Yang, Shi-Jinn Horng, Multivariate time series forecasting via attention-based encoder–decoder framework,Neurocomputing,Volume 388,2020,Pages 269-279,ISSN 0925-2312.

[15] K. Palasundram, N. Mohd Sharef, N. A. Nasharuddin, K. A. Kasmiran, and A. Azman, “Sequence to Sequence Model Performance for Education Chatbot,” International Journal of Emerging Technologies in Learning (iJET), vol. 14, no. 24, p. 56, Dec. 2019.

[16] “sklearn.preprocessing.MinMaxScaler — scikit-learn 0.22.1 documentation,” Scikit-learn.org, 2019. [Online]. Available: https://scikit.learn.org/stable/modules/generated/sklearn.preprocessing.MinMaxScaler.html.

[17] Mariam Moukalled Wassim El-Hajj Mohamad Jaber Computer Science Department American University of Beirut - Automated Stock Price Prediction Using Machine Learning.

[18] D. Bhuriya, G. Kaushal, A. Sharma and U. Singh, ”Stock market predication using a linear regression,” 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA), 2017, pp. 510-513, doi: 10.1109/ICECA.2017.8212716.

[19] Miró-Julià, M., Fiol-Roig, G., Isern-Deyà, A.P. (2010). Decision Trees in Stock Market Analysis: Construction and Validation. In: García-Pedrajas, N., Herrera, F., Fyfe, C., Benítez, J.M., Ali, M. (eds) Trends in Applied Intelligent Systems. IEA/AIE 2010. Lecture Notes in Computer Science(), vol 6096. Springer, Berlin, Heidelberg.

[20] J. Tanuwijaya and S. Hansun, “LQ45 Stock Index Prediction using kNearest Neighbors Regression,” International Journal of Recent Technology and Engineering, vol. 8, no. 3, pp. 2388–2391, Sep. 2019.

[21] Siami-Namini, Sima, and Akbar Siami Namin. ”Forecasting economics and financial time series: ARIMA vs. LSTM.” arXiv preprint arXiv:1803.06386 (2018).