

Determining Stock Market Forecasting by Means of LSTM Methodology

Dr. Ashutosh M. Kulkarni¹, Ms. Mayura V. Shelke², Mr. Vijayendra S. Gaikwad³ and Mr. Rahul B. Diwate⁴

¹Vishwakarma Institute of Technology, Pune, India

Email: ashutosh.kulkarni@vit.edu

²AISSMSIOIT, Pune, India

Email: mayura.shelke@aissmsioit.org

³SCTR's Pune Institute of Computer Technology, Pune, India

Email: vij711@gmail.com

⁴R. K. Dewan & Co., Pune, India

Email: diwate.rahul@gmail.com

Abstract—The Stock Price Prediction will help the user to detect prices as the market fluctuations have consistently been vague for investors because of very significant aspects. This paper focuses on reducing the risk of prediction significantly with some of the reliable deep learning and machine learning algorithms via neural networks. This study compares machine learning models Decision Tree, Random Forest, Support Vector Classifier, Naïve Bayes, K-Nearest Neighbors, Logistic Regression and Artificial Neural Network ANN and two powerful deep learning methods namely Long short-term memory (LSTM) and Recurrent Neural Network. Here, this paper used LSTM with ANN algorithms, where input is given in which firstly calculates the continuous data and then converting into binary data, this for the data training process. After creating LSTM and RNN models shows different outperforms different prediction models.

Index Terms— Long Short-Term Memory Model and Recurrent Neural Network, Web Application.

I. INTRODUCTION

The task of predicting Stock prices is one of the herculean tasks as the main reason behind it is to buy stocks and then to sell them that are likely to increase prices and sometimes may fall. The stock market is considered to be very complicated and dynamic in nature. The prediction of future prices precisely may lead to greater yield of profit for shareholders through share investments. According to the predictions, investors who want to invest will be able to pick the stocks that may give a higher return.[4]

Several predictions don't depend on a single factor, there are a lot of factors which are continuously influencing it like previous grades, data collected rational and irrational behaviors etc. These combined factors can make major changes making the prediction difficult especially in case of stocks. Here in this paper collected some stock prices of companies whose data was available on open platforms.[8] Then using several algorithms present in machine learning this paper tried to predict the future values of stocks, starting with linear regression technique, and averaging and Auto ARIMA [6] and Long Short-Term Memory techniques.[3][10].

II. REVIEW LITERATURE

The First focus of our literature reviews was to test different algorithms and models to determine if stock price predictions could be made on real prices.

This work has tried to look at existing plans, analyses major issues, and improve them. A search over the problems led us to LSTM. After deciding to use the LSTM neural network to make stock forecasts, time-series data are collected from stock prices and economically related variables from the year 2015 and onwards. In the following Table I discussed about challenges and discussion over all learning techniques and algorithm used for existing system.

TABLE I. LITERATURE REVIEW OF VARIOUS TECHNIQUES AND METHODOLOGIES FOR FINDING STOCK PREDICTION

Authors & Year of Pub.	Learning techniques	Algorithms Used	Dataset Used	Summary	Challenges and Discussions
Aureliano W. Li , Guilherme S. Bastos.(2020) [9]	Deep learning, stock market forecasting, systematic review,	Profitability metrics, risk management, technical analysis, technical indicators.	Yahoo Finance	Discussions are based on four key points of view: 1.profitability metrics 2.forecasting strategies 3.disaster/ risk management 4.trading strategies	Studies show the LSTM process used in 73.5% of cases. This contribution of work is to highlight certain constraints developed in the paper, as only 35.3% of subjects have analyzed net profit.
Sourav Sah,Abhijoy Sarkar, Chittaranjan Pradhan, Abhaya, Sahoo.(2020) [12]	Deep Learning, Neural Networks.	Sentiment analysis, RNN.	Google stock market dataset.	LSTM fails to fully mimic the design of a trend so that it works in any short-term decision.	The LSTM model has a dropout of 20% in this paper. Optimizer and Regularization followed by Dropouts was used for model construction.
J. Rao, K. Joshi, Bharathi N. (2020) [21]	Natural Language Processing (NLP), sentiment metric, kernels learning technique	Support Vector Machine, Naive Bayes classifier, Polarity Detection and Random Forest.	1)mining based framework 2) financial forecasting 3) Twitter feeds about airlines	Dictionaries were used for positive and negative names, where pre-processing of text data was a challenging task. They then compiled their own dictionary that included specific financial terminals. 3 Different classification models were then compiled on that data.	Random Forest has an average of 90% of accuracy, following SVM with 86%, and finally Naive Bayes with 83%.
Sidra Mehtab, Jaydip Sen.(2019) [6]	Classification, Regression , Deep Learning	Convolution Neural Network (CNN), Multivariate time series	NSE (National Stock Exchange) Nifty50	This paper first works on multiple algorithms and after they work CNN based regression it gives the best result among all tested models. They work on Different CNN model depends on training data	Machine Learning predictive models are not as good as CNN predictive models. Univariate analysis is less accurate than multivariate analysis.
Mahla Nikou,Gholamr eza Mansourfar,Jam shid Bagherzadeh. (2019)	Deep machine Learning.	Support Vector Regression (SVM) and Random Forest.	Research data from Yahoo finance site. The data used use iShares UKfund(NYSEA RCA: EWU)	In this they made a model using deep learning and then it compared with the machine learning algorithms. Followed by SVR method and	Paper recommends to use combined models for predicting price and reducing accuracy

[7]				random forest is with a smaller number of errors.	
Chun Yuan Lai , Rung-Ching Chen ² , C. Y. Lai, R. Eko Caraka.(2019) [11]	Long Short-Term Memory (LSTM) with RNN (Recurrent Neural Network).	Stochastic oscillator, On Balance Volume Moving Average Divergence and Convergence and RSI.	Quanta Computer incorporated, Taiwan Cement Corporate, Taiwan Semiconductor Manufacturing Company of Taiwan, Formosa Petrochemical Corp., Foxconn.	A comparison between LSTM and traditional Arima.	RNNs were successful in sequence prediction tasks, due to the explosion / disappearance of the gradient, which was problematic in dependence for long term.
Gholamreza Mansourfar, J. Bagherzadeh.(2019) [19]	Machine Learning	SVM, ANN, vector regression.	iShares, MSCI United Kingdom (Jan 2015 - Jun 2018).	Deep learning did most perfect prediction then vector regression method and lastly random forest.	Using vector prediction was challenging. different techniques are compared for prediction and comparing accuracy.
Rung-Ching Chen, Chun Yuan Lai,, Rezzy Eko Caraka. (2019) [22]	Long Short-Term Memory (LSTM) and Recurrent Neural Network (RNN).	Moving Average Divergence/Convergence, On Balance Volume and Relative Strength Index.	The London, New York and Karachi Stock Exchange	Using LR shows that accuracy is greatly improved when using PCA.RF shows high accuracy in binary split modeling	Selection of the algorithm, also depends largely on the type and volume of data in which predictions will be made.
Soman K.P., Gopalakrishnan E.A., Vinayakumar R., Vijay K.M., Sreelekshmy S.(2018) [3]	Neural Networks,deep learning architectures , ARIMA.	CNN ,LSTM, RNN	It uses a dataset of stocks listed on NSE, and it holds minute wise details and data of organization.	In this paper they have used three different types of deep learning architectures in real time using minute-wise data-the sliding window approach for future values prediction based on a short-term basis.	This paper shows CNN has more accuracy than RNN and LSTM deep learning models. ARIMA model perform worse than the deep learning model its error percentage is very high compared to deep learning models
Kunal Jadhav, Brij k. Vihangam, P. Gabda, G. Navale.(2016) [8]	Artificial Neural network (ANN) and Data Mining.	Auto Regressive Moving Average (ARMA algorithm)	-	For dealing with this type of big data for prediction of stocks they use data mining technique using ARMA algorithm and for predicting they state the ANN is good.	This model also shows good accuracy but still there is need to improve the accuracy and performance.
Ganapati Panda, Babita Mr.Venkata Pagolu and Kamal Reddy.(2016) [2]	Natural Language Processing (NLP), Machine Learning, N-gram	N-gram Representation, Word2vec Representation	2,50,000 tweets from AUG 2015 to AUG 2016 from Twitter	This project is about Stocks prediction on sentiments expressed by the public based on Twitter using N-gram and Word2vec representation.	The accuracy of this model is approximate 69% When we use logistic regression. This percent depends on the amount of training dataset
Dr. P. K. Sahoo, Mr. Krishna Charlapally. (2015) [19]	Auto regressive, Moore and Penrose technique	Artificial Neural Network back propagation algorithm, Markov chain	One day stock of Hyundai Merchant Marine.	Auto regressive model was seen predicting the stock prices very accurately,	No major challenges. Implementation was done using java and other three modules and normal classes and regression was done.

Richard MacKinnon, Carson Kai-Sang Leung, Yang Wang. (2014) [20]	Data mining and Machine Learning	Structural support vector machines (SSVMs)	All the dataset are taken in from yahoo Finance, Adobe Systems Incorporated (ADBE) dataset	This proposes about business intelligence application using machine learning They connect collaborating companies in the I.T. sector with graph form and use SSVM to forecast negative or positive movements in values of stocks.	The model shows low accuracy because they don't have enough training data. this is because of quality of test sample is lower than the training data5
Adebiyi M. O., Otokiti S. O., Ayo C. K., Adebiyi A. A. (2012) [4]	Artificial Neural Network (ANN)	backpropagation network (BPN), hybridized and technical analysis, backpropagation algorithm.	Random Sample of Daily Stock Price from CIS journal.	A high level of accuracy was found in hybrid approaches that have good results than technical analysis. Multilayer perceptron model was done.	Determining the critical impact of certain variables of the fundamental analysis on the quality of stock price forecasts was challenging.

III. METHODOLOGY

A. Technology

Following Technology used for the implementation of proposed work.

- Python Language
- Machine Learning
- Streamlit is the framework for creating the app for the machine learning and data science models.
- Plotly.

B. Dataset

Research Dataset collected from Yahoo finance website. This site provides all types of data about all stock markets present in the world such as New York Stock Exchange (NYSE), National Stock Exchange (NSE), Bombay Stock Exchange (BSE), London Stock Exchange (LSE) and many more. Data of Listed companies on this stock exchange is present on this yahoo finance. NSE listed companies are present as companies code name and end with .NS and BSE companies with .BO. This dataset is imported from python library “yahoo finance” [3][4][5].

Data has information in table format in each time period. First it shows the date and it shows the opening price of stock on that day. The higher value shows the highest value of stock on that day. Then the Close column shows the closing price of stock on a particular day. The Low column shows the lowest price of stock on a given date.[7][9][13]

C. Algorithm

LSTM is one of the main types of RNN (Recurrent Neural Network), it's also capable of storing data from past stages of algorithm and use it for predictions, also called as an ANN.

ANN consists of three layers which are input, hidden layers, and output layer. To build the mode here LSTM is used. For the purpose of training the model proposed work use 80% of data and for testing purposes proposed work use 20% of remaining data. Mean Squared errors (MSE) are used here to enhance the model and also have different epochs are used for different layers of LSTM training data epochs such as 14 epochs, 15 epochs, 40 epochs and 75 epochs. Before processing data and preparing data, the first important step is to collect data from the market. The next step is to pre-process data where the data with mistakes will be cleared in the dataset.[5]

The preparation of the dataset, for the execution of the model using prepared data. The final step is to draw the data using visualization techniques that helps to show variation of the data in the output. Here, a web app for presenting and predicting stock prices for desired companies has been made. Web app is made with help of streamlit library where the raw data, plot for raw data and the predicted output for stocks in the form of data is presented.[5][8][9][14]

D. Execution Steps

- Step 1: - Collecting the required data.

- Step 2: - Pre-processing of the data collected from the market.
- Step 3: -Importing datasets and libraries and reading the datasets.
- Step 4: -Feature extraction of the data so that data can vary in 0 and 1.
- Step 5: -Creating algorithms for data processing.
- Step 6: -Building LSTM model.
- Step 7: -Adding LSTM layer and some dropout command for removing unwanted data from the dataset.
- Step 8: -Adding the output layer.
- Step 9: -Making predictions and visualizing the different output.
- Step 10: -Developing a web app for displaying the output of models predicted graphs.

IV. RESULTS & DISCUSSION

The LSTM model which is proposed is using python for its execution which forecasts the price of share in the future based on its previous data. The graph of prediction is illustrated in the data visualization below. This graph represents the prediction of the stock price of a share in our paper using the LSTM algorithm. An analysis of Figure 1 shows its results by comparing them to the trained model by means of an algorithm that is defined in section previous. A share price is shown on the "x" axis. The "y" axis represents the year.



Fig 1: Graph for Real value stock vs Predicted value stock for HCL Company using LSTM



Fig 2: Graph for Real value stock vs Predicted value stock for MRF Company using LSTM

There is a choice for which company's stock value you want to predict. It also shows the predicted price for tomorrow.



Fig 3: Graph for Real value stock vs Predicted value stock for Dr.Reddy Company using LSTM

V. LIMITATIONS

LSTM's work very good for some issue, but some of the issue disadvantages are:

- LSTM's take a longer time to train
- To train LSTM's requires more memory.
- LSTM's overfit easily.
- In LSTM's dropout is difficult to implement
- LSTM's are delicate to disparate random weight initializations

VI. CONCLUSION

This paper used a Deep Learning-based finalization for predicting the stock price of a particular company. In this paper can see that Deep Neural Network architectures are competent in finding unseen dynamics and are capable of making predictions. The model was trained using the stock prices data of companies and was able to predict the share price of that company. Hence, proposed work is competent in identifying some interrelation within the data.

Furthermore, in this work of prediction of many shares, there perhaps be some opportunity for a distinct business review. The disparate pattern of the stock price of company sectors can be studied and it can also be analyses with a plot of various time periods to bring consistency and to find good accuracy. The stock market study and prediction of growth of different sectors of company in various time period will be helped by this foundation. Incorporating other components (e.g. twitter sentiment analysis, election result, global politics) that are not directly agree with the closing stock price may accurate the prediction accuracy of stock price

FUTURE SCOPE

In future enhancements of our proposed work, includes sentiment analysis from social media -- Twitter and Facebook API -- to capture what the market is thinking about changes in share prices. Facebook, after all, is a huge social media platform that has a large amount of trend information posted by users.

REFERENCES

- [1] Khedr, Ayman E., and Nagwa Yaseen. "Predicting stock market behavior using data mining technique and news sentiment analysis." *International Journal of Intelligent Systems and Applications* 9.7 (2017): 22.
- [2] Pagolu, Venkata Sasank, et al. "Sentiment analysis of Twitter data for predicting stock market movements." 2016 international conference on signal processing, communication, power and embedded system (SCOPEs). IEEE, 2016.
- [3] Selvin, Sreelekshmy, et al. "Stock price prediction using LSTM, RNN and CNN-sliding window model." 2017 international conference on advances in computing, communications and informatics (icacci). IEEE, 2017.
- [4] Adebiyi, Ayodele A., et al. "Stock price prediction using neural network with hybridized market indicators." *Journal of Emerging Trends in Computing and Information Sciences* 3.1 (2012).

- [5] Sharma, Ashish, Dinesh Bhuriya, and Upendra Singh. "Survey of stock market prediction using machine learning approach." 2017 international conference of electronics, communication and aerospace technology (ICECA). Vol. 2. IEEE, 2017.
- [6] Mehtab, Sidra, and Jaydip Sen. "Stock price prediction using convolutional neural networks on a multivariate timeseries." arXiv preprint arXiv:2001.09769 (2020).
- [7] Nikou, Mahla, Gholamreza Mansourfar, and Jamshid Bagherzadeh. "Stock price prediction using DEEP learning algorithm and its comparison with machine learning algorithms." *Intelligent Systems in Accounting, Finance and Management* 26.4 (2019): 164-174.
- [8] Navale, G. S., et al. "Prediction of stock market using data mining and artificial intelligence." *International Journal of Computer Applications* 134.12 (2016): 9-11.
- [9] Li, Audeliano Wolian, and Guilherme Sousa Bastos. "Stock market forecasting using deep learning and technical analysis: a systematic review." *IEEE Access* 8 (2020): 185232-185242.
- [10] Misra, Meghna, Ajay Prakash Yadav, and Harkiran Kaur. "Stock Market Prediction using Machine Learning Algorithms: A Classification Study." 2018 International Conference on Recent Innovations in Electrical, Electronics & Communication Engineering (ICRIEECE). IEEE, 2018.
- [11] Lai, Chun Yuan, Rung-Ching Chen, and Rezzy Eko Caraka. "Prediction stock price based on different index factors using LSTM." 2019 International Conference on Machine Learning and Cybernetics (ICMLC). IEEE, 2019.
- [12] Sarkar, Abhijoy, et al. "LSTMSA: A Novel Approach for Stock Market Prediction Using LSTM and Sentiment Analysis." 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA). IEEE, 2020.
- [13] Attanasio, Giuseppe, et al. "Combining news sentiment and technical analysis to predict stock trend reversal." 2019 International Conference on Data Mining Workshops (ICDMW). IEEE, 2019.
- [14] Li, Audeliano Wolian, and Guilherme Sousa Bastos. "Stock market forecasting using deep learning and technical analysis: a systematic review." *IEEE Access* 8 (2020): 185232-185242.
- [15] Xu, Yichuan, and Vlado Keselj. "Stock prediction using deep learning and sentiment analysis." 2019 IEEE International Conference on Big Data (Big Data). IEEE, 2019.
- [16] T Liu, Tong, et al. "Recurrent neural networks based on LSTM for predicting geomagnetic field." 2018 IEEE International Conference on Aerospace Electronics and Remote Sensing Technology (ICARES). IEEE, 2018.
- [17] Rezk, Nesma M., et al. "Recurrent neural networks: an embedded computing perspective." *IEEE Access* 8 (2020): 57967-57996.
- [18] Truong, Nhan Duy, et al. "Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram." *Neural Networks* 105 (2018): 104-111.
- [19] Sahoo, P. K., and Krishna Charlapally. "Stock price prediction using regression analysis." *International Journal of Scientific & Engineering Research* 6.3 (2015): 1655-1659.
- [20] Leung, Carson Kai-Sang, Richard Kyle MacKinnon, and Yang Wang. "A machine learning approach for stock price prediction." *Proceedings of the 18th International Database Engineering & Applications Symposium*. 2014.
- [21] Kalyani, Joshi, Prof Bharathi, and Prof Jyothi. "Stock trend prediction using news sentiment analysis." arXiv preprint arXiv:1607.01958 (2016).
- [22] LAI, CHUN YUAN, RUNG-CHING CHEN, AND R. E. CARAKA. "PREDICTION AVERAGE STOCK PRICE MARKET USING LSTM." (2019): 1-6.