International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 05 | May 2021 www.irjet.net p-ISSN: 2395-0072

STOCK MARKET PREDICTION USING MACHINE LEARNING

Abhinit Davane[1], Shailesh Bhagat[2], Mehul Bhoi[3], Prajaktee Rane[4]

[1,2,3] *Student, Department of Computer Engineering, Theem College of Engineering, Boisar, Maharashtra, India*
[4] *Professor, Department of Computer Engineering, Theem College of Engineering, Boisar, Maharashtra, India*

Abstract - *In the finance international stock trading is one of the maximum crucial activities, our goal is to predict the future value of the stocks of a company. The latest trend in stock market prediction technologies is the use of machine learning which makes predictions primarily based totally at the values of present-day stock market indices by training on their previous values. Machine learning gaining knowledge to make prediction less complicated and authentic. The programming language is used to predict the stock market using machine learning is Python. The paper makes a specialty of the use of Decision Tree and LSTM primarily based totally on Machine learning to predict stock values. Factors taken into consideration are open, close, low, high & volume.*

1.INTRODUCTION

Predicting the Stock Market has been the bane and intention of buyers since its existence. A accurate prediction of stocks will cause immense profits for the vendor and therefore the broker. Everyday billions of dollars are traded at the exchange, and behind each dollar is an investor hoping to profit in a way or another. Frequently, it is brought out that prediction is chaotic instead of random, which suggests, it can be foretold by rigorously analysing the history of the various stock market. Machine learning is an efficient way to represent such processes. It predicts a value near to the tangible value, thereby increasing the accuracy. The introduction of machine learning the realm of stock prediction has appealed to several researchers due to its efficient and accurate measurements.

The important a part of machine learning is the dataset used. The dataset ought to be as concrete as attainable as a result of a little change within the knowledge will carry on huge changes in the outcome. In this project, supervised machine learning is used on a dataset obtained from Tiingo or Yahoo. This dataset contains of following 5 variables: open, close, low, high, and volume. Open, close, low, and high are totally different bid costs for the stock at separate times with nearly direct names. The amount is that the range of shares that passed from one owner to a different during the amount time. The model is then tested on the test data from Tiingo or Yahoo.

Long Short-Term memory is one of the most successful RNNs architectures. LSTM introduces the memory cell, a unit of computation that replaces traditional artificial neurons within the hidden layer of the network. With these memory cells, networks are able to effectively associate recollections and input remote in time, thus suit to know the structure of data dynamically over time with high prediction capacity.

2. ALGORITHM

The stock market prediction looks a complex problem as a result of several factors and it doesn't appear applied math at first. however, by correct use of machine learning techniques, one will relate previous data to the present data and train the machine to learn from it and build acceptable assumptions. Machine learning as such has several models however paper focuses on the 2 models' predictions.

2.1 Decision Tree Model

Decision Tree analysis is one among predictive modelling approaches employed in Statistics, Data Processing and Machine Learning.

A Decision Tree will be wont to visually and expressly represent selections and decision making. In data mining, a decision tree describes data however the resulting classification tree are often associate input for decision making.

Work is finished on CSV format of data through panda library and calculated the parameter that is to be predicted, the value of the stocks regarding time. The data is split into completely different train sets for cross-validation to avoid overfitting. The test set is mostly kept 20% of the whole dataset. Linear regression as given by the above equation is performed on top of the data and so predictions are made, that are plotted to indicate the results of the stock market prices vs time.

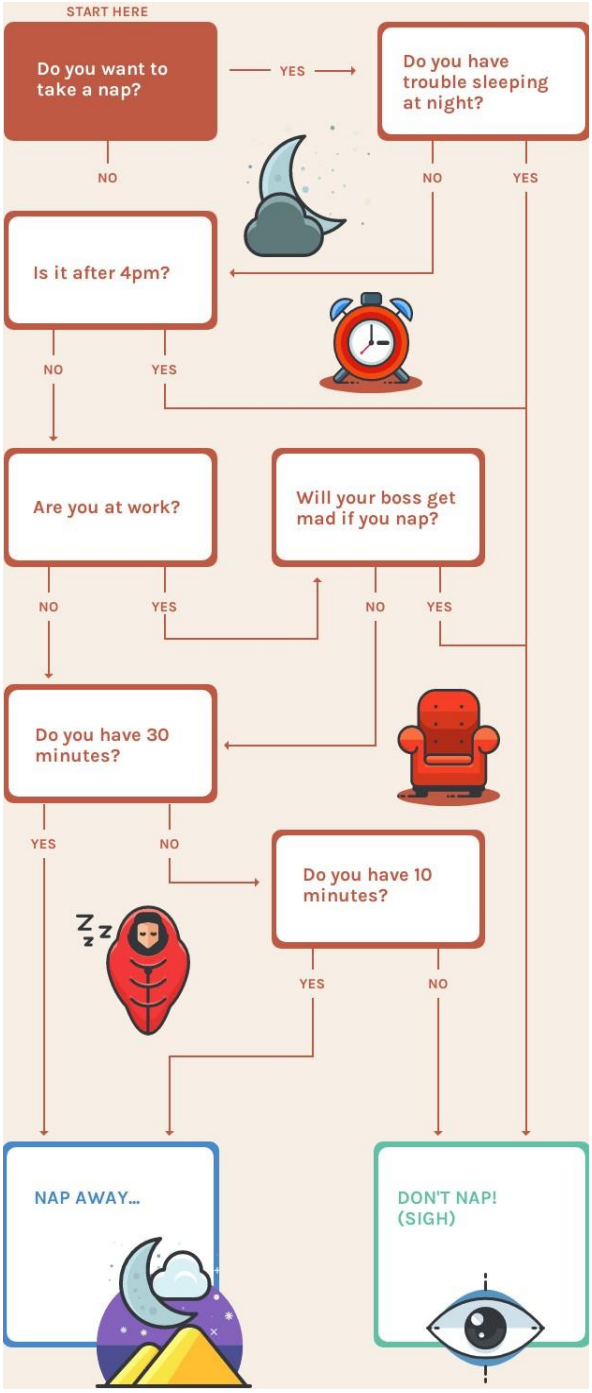

Fig -1: Decision Tree Examples

2.2 Long Short-Term Memory (LSTM) Network Based Model

Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture

The main reason behind exploitation this model in stock market prediction is that the predictions rely upon giant amounts of data and are generally dependent on the long-term history of the market. offer aid to the RNNs through retentive information for older stages creating the prediction additional accurate.

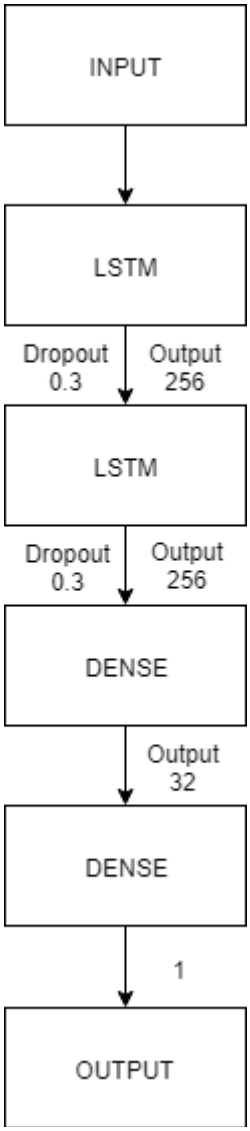

Fig -2: LSTM Layers

Since the stock market have the processing of giant data, the gradients regarding the weight matrix may become terribly tiny and will degrade the training rate of the system. This corresponds to the matter of Vanishing Gradient. LSTM prevents this from occurrence. The LSTM consists of a memory cell, input gate, output gate and a forget gate. The cell remembers the worth for long propagation and therefore the gates regulate them.

Work is finished on CSV format of data through panda library and calculated the parameter that is to be predicted, the value of the stocks regarding time. The data is split into completely different train sets for cross-validation to avoid overfitting. The test set is mostly kept 20% of the whole dataset. Linear regression as given by the above equation is performed on

top of the data and so predictions are made, that are plotted to indicate the results of the stock market prices vs time.

3. EXPERIMENT AND RESULT

The proposed system is trained and tested over the dataset taken from Tiingo or Yahoo. It is split into 2 (training and testing) sets respectively and yields the results upon passing through the different models:

3.1 Decision Tree Based Model Results

The plot in Chart -1 is the result of application of Decision Tree algorithm on the dataset to predict varying prices with respect to the time.

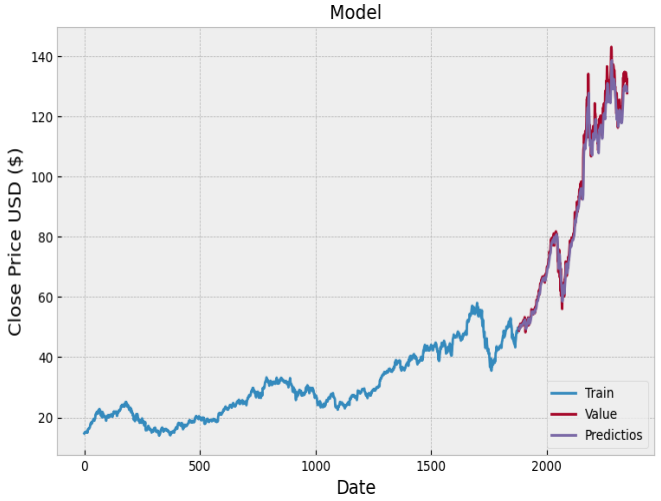

Chart -1: Plot between Price and Date Using Decision Tree

3.2 LSTM Based Model Results

The plot in Chart -1 is the result of application of LSTM algorithm on the dataset to predict varying prices with respect to the time.

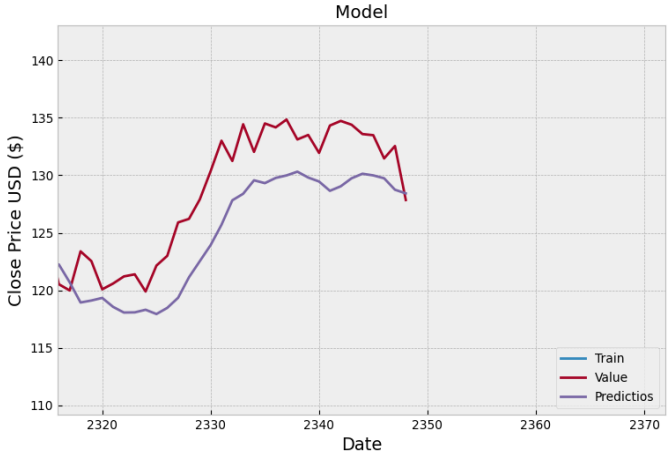

Chart -2 Plot between Actual and Predicted Trend of LSTM

The on top of graph in figure 3 and 4 is a plot over the information having batch size 512 and 90 epochs. The prediction is shown by a purple line and also the actual trend is shown by red. The proximity of these 2 lines tells, however economical the LSTM primarily based model is. The prediction approximates the real trend once a substantial quantity of time has passed. The more the system is trained the bigger the accuracy which is able to be attained.

4. CONCLUSIONS

This article uses two methods: LSTM and Decision Tree, on the Tiingo or Yahoo dataset. Each method showed an improvement in forecast accuracy and provided positive results. The use of recently introduced machine learning techniques to predict stocks has shown encouraging results, marking their use in profitable trading systems. The conclusion drawn from this is that machine learning techniques can be used to predict the stock market very accurately and effectively. The prediction system can be further enhanced by using a data set that is much larger than the currently used data set. This makes it easier to improve the accuracy of our forecasting model. In addition, you can even check alternative machine learning models to understand their accuracy.

REFERENCES

[1] Gupta, A. :"Stock market prediction using Hidden Markov Models", IEEE Engineering and Systems (SCES), 2012 Students Conference on, pp.1-4, 2012.

[2] Carlos A. Coello , Gary B. Lamont , David A. van Veldhuizen : "Evolutionary Algorithms for Solving Multi-Objective Problems",Springer, 2007

[3] Fazel Z.,Esmaeil H., Turksen B.: "A hybrid fuzzy intelligent agent-based system for stock price prediction", International Journal of Intelligent Systems, Volume 27, Issue 11, pages 947–969, 2012.

[4] Chun C, Qinghua M, Shuqiang L.: "Research on Support Vector Regression in the Stock Market Forecasting" ©Springer, Advances in Intelligent and Soft Computing Volume 148, , pp 607-612, 2012.

[5] Jui Y.: "Computational Intelligence Approaches for Stock Price Forecasting", IEEE International Symposium on Computer, Consumer and Control (IS3C), pp. 52 - 55, 2012.

[6] Olivier C., Blaise Pascal University: "Neural network modeling for stock movement prediction, state of art". 2007

[7] Suykens, J. A. K., Gestel, V. T., Brabanter, J. D., Moor, B.D and Vandewalle, J. "Least squares support vector machines", World Scientific, 2002.

[8] Nayak, S.C. :"Index prediction with neuro-genetic hybrid network: A comparative analysis of performance", IEEE International Conference on Computing, Communication and Applications (ICCCA), pp. 1-6, 2012.

[9] Kai Chen, Yi Zhou and FangyanDai —A LSTM-based method for stock returns prediction: A case study of

China stock market,|| IEEE International Conference on Big Data,2015.

[10] DhirajMundada, GauravChhaparwal, SachinChaudhari, and TruptiBhamare— Stock Value Prediction System,|| International Journal on Recent and Innovation Trends in Computing and Communication, April 2015

[11] Prashant S. Chavan, Prof. Dr. Shrishail. T. Patil —Parameters for Stock Market Prediction,|| Prashant S Chavan et al, Int.J.Computer Technology & Applications, Vol 4 (2),337-340.

[12] Zhen Hu, Jibe Zhu, and Ken Tse "Stocks Market Prediction Using Support Vector Machine", 6th International Conference on Information Management, Innovation Management and Industrial Engineering, 2013.M.

[13] Debashish Das and Mohammad shorif uddin data mining and neural network techniques in stock market prediction: a methodological review, international journal of artificial intelligence & applications, vol.4, no.1, January 2013

[14] K. jae Kim, "Financial time series forecasting using support vector machines," Neurocomputing, vol. 55, 2003.

[15] N. Ancona, Classification Properties of Support Vector Machines for Regression, Technical Report, RIIESI/CNR-Nr. 02/99.