



LSTM and Regression Methods for Stock Market Prediction

Nishanth Vaidya, Nikhil Bharadwaj

Student, Student

Department of CSE,

Sambhram Institute of Technology, Bangalore, India

Abstract-By using stock market prediction, the aim is to predict the values of stocks in the future. There has been a growing interest in stock market prediction technologies via the use of machine learning. This is done by taking the current values of the market after taking previous stock values as the training data. Various Machine learning techniques can be used to make prediction simpler in general. The proposed paper works on the working of 2 different algorithms, the Regression and LSVM based algorithm.

Keywords - Least Square Support Vector Machine, Particle Swarm Optimization, regression, and volume.

1. INTRODUCTION

Stock market prediction has been at the forefront for a long time due to the benefits it can reap. But the prediction of the stock market is not a simple task, mainly as a consequence of the random-walk behaviour of a stock time series. But in recent times, it is brought out that prediction is chaotic rather than being classified as random, meaning that it can be predicted by analyzing the history of the stock market. Initially fundamental and technical analyses were the two methods used for stock market prediction. But nowadays, the most commonly used method is Artificial Neural Networks (ANN).

One of the most important part for any of the prediction methods used is the dataset. This model that has been proposed has been applied and evaluated on fourteen different datasets such as the Yahoo Finance Dataset. Open, close, high and low are different bid prices for the stock at distinct time periods with almost direct names. The volume is given by the number of shares that passed on from one holder to another. The model is tested on the data after that for both the methods. In the first method [1], Regression and LSTM models are engaged separately. Regression involves the process of minimizing error and LSTM [3] [4] contributes to remembering the data and results indefinitely. Finally, the graphs for the variation of prices with the dates and between the predicted and actual price are plotted.

2. MACHINE LEARNING METHODS

Machine learning is one of the hottest topics in research and industry, with new methodologies being developed all the time. The speed and complexity of the field makes keeping up with new techniques difficult for experts and beginners alike.

A) Supervised Learning - Supervised learning is a type of machine learning task which learns a function that maps an input to an output based on example or training input-output pairs. It includes algorithms such as Support Vector Machines (SVM), k-Nearest Neighbours (k-NN) and Artificial Neural Networks (ANN), Genetic algorithms and Decision Trees (DT).

- **Naïve Bayes (NB)** - These algorithms entail probabilistic classifiers that make the prior assumption that the features of the input data are independent of each other. They are scalable and only require small to medium training datasets to produce appreciable results.

- **Support Vector Machines (SVM)** - **Support Vector Machines** or SVM for short, are models with associated learning algorithms which analyze the given datasets for both, classification as well as regression. Consider examples consisting of training data, with each example being marked as belonging to either of two categories. For these examples, the SVM algorithm builds a model that separates new examples to either category. Doing so makes it a non-probabilistic classifier. An SVM represents its examples as varying points in space. They are mapped so that the examples of the separate categories are clearly distinct.

B) Unsupervised Learning – Unsupervised Learning is a type of machine learning algorithm that makes decisions from datasets consisting of input data without any labelled outputs. Unsupervised Learning consists of two tasks being Association and Clustering.

- **Clustering** – Clustering consists of grouping data points that present alike characteristics. Well known approaches typically include algorithms such as k-means and hierarchical clustering. Clustering methods are scalable but the scalability is limited in a sense, but they represent a feasible solution that is used as a phase before adopting a supervised algorithm or for other detection purposes.
- **Association.** The aim of association is to identify unknown patterns across data, making them suitable for the purpose of prediction. However, they tend to produce an large output of sometimes invalid rules, hence they require a human overseer to function.

3. LITERATURE SURVEY

- Ishita Parmar, Navanshu Agarwal, Sheirsh Saxena “Stock Market Prediction Using Machine Learning” 2018 First International Conference on Secure Cyber Computing and Communication(ICSCCC), 2018
- Osman Hegazy 1 , Omar S. Soliman 2 and Mustafa Abdul Salam, " A Machine Learning Model for Stock Market Prediction," International Journal of Computer Science and Telecommunications December 2013
- H. L. Siew and M. J. Nordin, "Regression techniques for the prediction of stock price trend," 2012 International Conference on Statistics in Science, Business and Engineering (ICSSBE), Langkawi, 2012, pp. 1- 5.
- Khalil A.S.: “An Investigation into Optimization Strategies of Genetic Algorithms and Swarm Intelligence”. Artificial Life (2001).

4. PROPOSED METHODS FOR STOCK MARKET PREDICTION

A) LSTM and Regression - The dataset being analyzed was picked up from Yahoo Finance. The dataset consists of close to 900,000 records of the required stock prices and other values that prove relevant as well. The data successfully reflected the stock prices at certain daily intervals for the year. It consisted of various sections namely date, symbol, etc.

I) Regression Model

In general, the Regression based Model is used for predicting continuous values through some given independent values [5]. Regression uses a linear function previously given for predicting continuous values:

$$V = a + bk + \text{error}$$

Where, V is not a discrete value; K represents independent values currently; and, a, b are the coefficients.

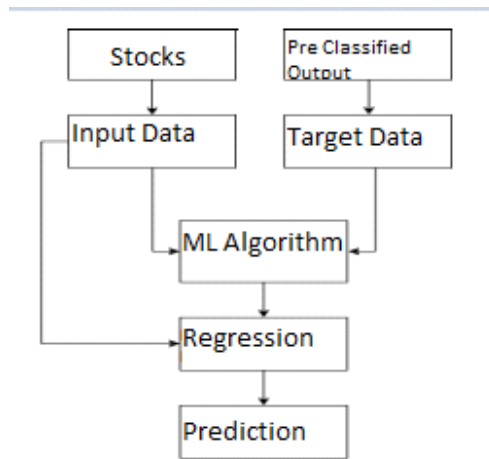


Figure 1 – Regression Based Model

Linear Regression [6] as given by the above equation has been performed on the data and then the relevant predictions are done. The factors considered for the regression were low, open, etc. The R-square confidence test was done to determine the confidence score. The predictions were plotted to show the results of the stock market prices vs time.

II) Long Short Term Memory (LSTM) Model - LSTM is the an advance version of Recurrent-Neural Networks (RNN). Here the information belonging to previous state persists. But LSTM are different from RNNs as they involve persistent dependencies while RNNs works on finding the relationship between the recent and the current information.

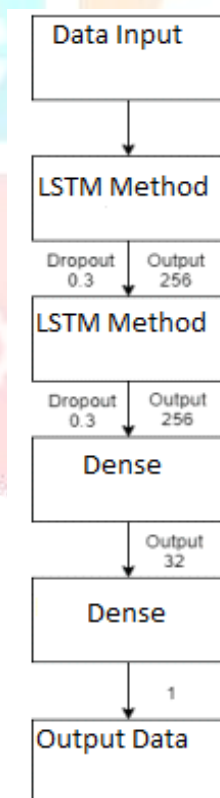


Figure 2 - LSTM method od Prediction

This indicates that the interval of information is relatively smaller compared to LSTM. The main purpose behind using this model in stock market prediction is due to that the predictions depends on large amounts of data and are generally dependent on the history of the stock market [6]. Therefore, LSTM regulates error by giving helping the RNNs by retaining information for older stages making the prediction increase in accuracy [7]. Thus, proving itself as much more reliable compared to other methods that were used. Since stock market involves processing of huge amount of data, the gradients with respect to the weight matrix may become minute and may degrade the learning rate.[8]. The R-square confidence test resulted in a confidence score of 0.86625.

5. CONCLUSIONS

This paper was created as an attempt to determine the future prices of the stocks with better accuracy and reliability using machine learning techniques such as regression and LSTM. The primary contribution of the researchers being the uses of the novel LSTM Model as a means for determining the variation of prices in the stock market. Both techniques have shown to be an improvement in the accuracy of predictions, thereby yielding better results with the LSTM model proving to be more efficient than regression. The results are promising and this has led to the conclusion that it is definitely possible to predict stock market with more accuracy and efficiency using machine learning techniques.

6. REFERENCES

- [1] Ishita Parmar, Navanshu Agarwal, Sheirsh Saxena "Stock Market Prediction Using Machine Learning" 2018 First International Conference on Secure Cyber Computing and Communication(ICSCCC), 2018
- [2] Osman Hegazy 1 , Omar S. Soliman 2 and Mustafa Abdul Salam, " A Machine Learning Model for Stock Market Prediction," International Journal of Computer Science and Telecommunications December 2013
- [3] H. Gunduz, Z. Cataltepe and Y. Yaslan, "Stock market direction prediction using deep neural networks," 2017 25th Signal Processing and Communications Applications Conference (SIU), Antalya, 2017, pp. 1-4.
- [4] M. Billah, S. Waheed and A. Hanifa, "Stock market prediction using an improved training algorithm of neural network," 2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), Rajshahi, 2016, pp. 1-4.
- [5] H. L. Siew and M. J. Nordin, "Regression techniques for the prediction of stock price trend," 2012 International Conference on Statistics in Science, Business and Engineering (ICSSBE), Langkawi, 2012, pp. 1- 5.
- [6] K. V. Sujatha and S. M. Sundaram, "Stock index prediction using regression and neural network models under non normal conditions," INTERACT-2010, Chennai, 2010, pp. 59-63.
- [7] S. Liu, G. Liao and Y. Ding, "Stock transaction prediction modelling and analysis based on LSTM," 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), Wuhan, 2018, pp. 2787-2790.
- [8] T. Gao, Y. Chai and Y. Liu, "Applying long short term memory neural networks for predicting stock closing price," 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, 2017, pp. 575-578.
- [9] K. A. Althelaya, E. M. El-Alfy and S. Mohammed, "Evaluation of bidirectional LSTM for short-and long-term stock market prediction," 2018 9th International Conference on Information and Communication Systems (ICICS), Irbid, 2018
- [9] Khalil A.S.: "An Investigation into Optimization Strategies of Genetic Algorithms and Swarm Intelligence". Artificial Life (2001).
- [10] Kennedy J., Spears W.M.: "Matching Algorithms to Problems", An Experimental Test of the Particle Swarm and Some Genetic Algorithms on the Multimodal Problem Generator. Processes. Current as of December 15th, 2003.