

# Price Trend Prediction of Stock Market Using Outlier Data Mining Algorithm

Shubham Nipane<sup>1</sup>, Prashant Kale<sup>2</sup>, Payal Kapsekar<sup>3</sup>, Pooja Agrawal<sup>4</sup>, Prof. Omkar Dudbhure<sup>5</sup>

<sup>1234</sup>UG Students, MIET Bhandara (Shahapur), India, 441906  
<sup>5</sup>Assistant Professor, MIET Bhandara (Shahapur), India, 441906

**Received on:** 30 April, 2021, **Revised on:** 25 May, 2021, **Published on:** 27 May, 2021

**Abstract** – In the financial world, stock price forecasts are gaining momentum. Successful estimates of future stock costs could yield significant returns. Focuses on the use of Regression and LSTM (Short-Term Memory) machine learning to predict stock prices. The considered features are open, close, low, high and volume. The latest trend in stock market forecasting technology is the use of machine learning that makes predictions based on current stock market indicators by training at their previous prices.

## I-INTRODUCTION

Global total market capitalization started from \$ 2.5 trillion in 2010. At the end of 2019, it was \$ 68.65 trillion. Proper stock forecasting can lead to great benefits for both the trader and the trader. In our LSTM stock forecasting model, a single sequence was defined as a sequential daily database of any single stock over a specified period of time (N days). The in-depth reading method known as the Gated Recurrent Unit (GRU) contains the same type of structure as the LSTM model, except that the memory cell structure is simplified in GRU. The introduction of machine learning in stock forecasting has led to much research due to its efficient and accurate measurements.

## II- RELATED WORK

Computer technology improvements have affected the markets in many ways. For example, to introduction of the electronic trading system, previously handled manually among financial market brokers [1]. In addition, the present capacity offered by computational systems and extensive financial data significantly benefited financial time series forecasting. However, financial time series dynamics characterized by nonlinearity present challenges to the building of forecasting models [2]. So, in recent decades, numerous studies have suggested ways to predict financial time series [7]. Indeed, technical features, such as noisy, nonparametric, and a chaotic nature, occur because financial prices are affected by different reasons and many macro-economic factors.

Hence, forecasting accuracy has become a huge challenge and of great interest to investors. Traditional statistical methods commonly assume that time series are generated from a linear process and make predictions for future values [5]. On the other hand, AI techniques, such as SC, have been applied with success because they can capture nonlinear behavior among the relevant factors [3]. One of the most widespread AI techniques for predicting is the use of ANN. The first study on the applicability of this approach to financial markets was

prepared by [4]. Based on this research, several studies have emerged and broadened the horizons of price and movement forecasts for capital market assets [4], [7]. Nevertheless, ANN models have limitations that have spurred the development and application of new ML techniques to solve these issues. The model's shortcoming main suffered is over fitting, the major drawback of the risk minimization principle [11].

Deep network techniques have been applied with relative success to financial market prediction [2] as have other ML methods [6]. In addition to AI's ability to capture nonlinear relations among relevant issues without prior knowledge [3], another relevant factor is the lower computational cost of this method [10]. Among the models that have applied ML, [8]– [5], have received more attention and are related to this paper. Advanced AI models, such as DL, have attracted attention by using LSTM. RNN have shown proprieties that can learn a large quantity of temporal data [10] in hidden layers [1].

Introduced by [9], LSTM is an improvement on recurrent networks. Neural networks were created with the goal of representing the human brain mathematically. Their structure contains units called neurons, similar to the human biological system. The existing interactions between neurons, responsible for information transmission, are represented by activation functions. Traditionally, ANN have connections in a single direction RNNs, however, have the capability to backflow information [5]. When inferring about a certain situation, the human brain recurs from a preexisting memory about the context. An ANN also needs this recursive ability. However, it is a challenge for RNNs, which suffer from a large amount of data. In this case, error signals tend to miss with a short-term memory, creating the vanishing gradient effect [9].

Therefore, a novel recurrent network architecture with an appropriate gradient-based capability to handle this error back-flow was presented in [5].

The model architecture has LSTM and the capability to learn time intervals over 1,000 steps, even in cases of noisy and incompressible input sequences. This new approach, called memory cell, passes information through gate units, and a constant error fflow avoids vanishing gradient between the steps. In addition to the input and output gates, the forget gate unit is responsible for retraining or forgetting the necessary information about the current stage, see Figure 1

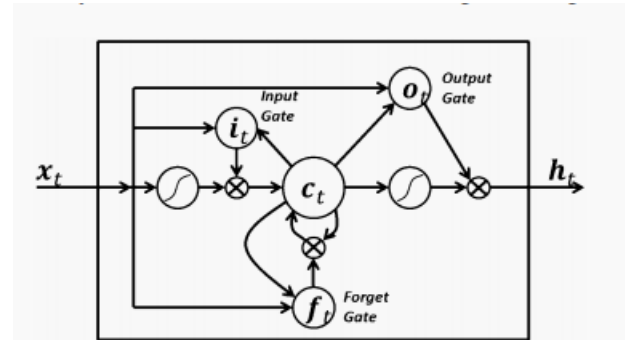


Fig. 1: Architecture of long short-term memory [9].

In practical terms, the forecasting model constructed in [10] is based on ML and LSTM techniques for day trading where the position is finished in the same trading day. The evaluation model was compared with ML traditional algorithms and other well-established investment strategies. Thus, the experimental results showed that the proposed model achieved good accuracy and created profit. The proposed LSTM in [8] contributed to an automated algorithm of the decision-making process. The model executes sequential decisions using 1 minute candles from financial time series. To avoid over fighting, the most popular regularization dropout technique was chosen. The experimental results in all three simulation systems exhibited the capability to yield profits.

In addition to the benefits of this technique, other studies, such as [7], [8] combined the use of LSTM with technical indicators to improve the results. The indicators added in [5] model reduced the influence of noise in the market, characteristic of this type of temporal series. The results reveal the significant contribution of the proposed method. An automated trading system model based on LSTM is proposed to predict price movements in stock index futures.

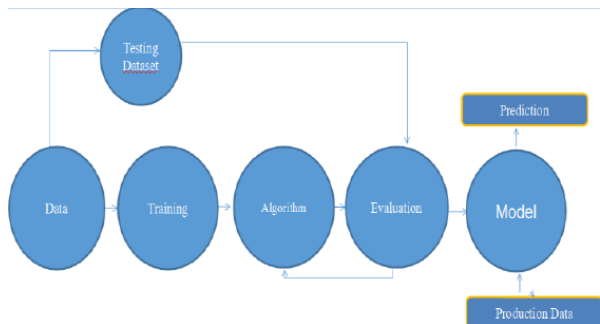
Additionally, trading strategies such as risk management (RM), take profit (TP), and stop loss (SL) are implemented to avoid erroneous entries. This component represents the most significant contribution of this study because trading strategy is rarely explored by researchers while, in practice, it is the most important factor for investors making trading decisions. If this evaluation is not applied, it is difficult to know that the proposed model works in the real world. As suggested by [6], perhaps, many failed models could exist in the literature because the models are not evaluated in the real market.

The next section explains the methodology applied in this paper.

### III- METHODOLOGY

Stock market prediction seems a complex problem because there are many factors that have yet to be addressed and it doesn't seem statistical at first. But by proper use of machine learning techniques, one can relate previous data to the current data and train the machine to learn from it and make appropriate assumptions.

LSTM was first proposed by Hochreiter and Schmidhuber in 1997, and later became very popular especially to address time series prediction problems. Being a modified RNN method, LSTM works well on a large variety of problems, and is widely used now. LSTM is intended to maintain a problem having long term dependency.



Although machine learning as such has many models but this paper focuses on two of the most important amongst them and made the predictions.

### IV- DESIGN

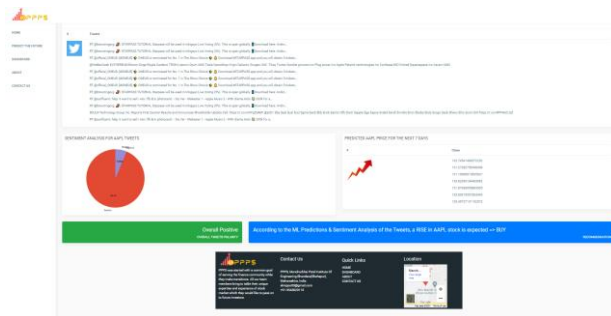
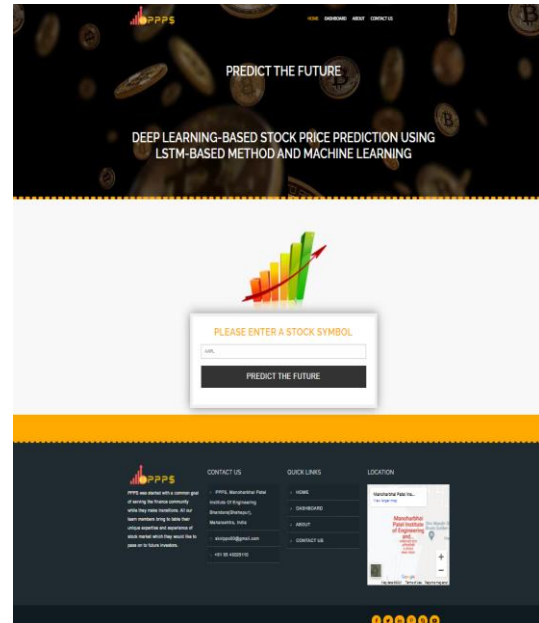


Fig 2- Output According to The ML Prediction and Algorithms



### V- CONCLUSION

This project was an attempt to determine the future prices of the stocks of a company with greater accuracy and reliability using machine learning and LSTM techniques. Both the techniques have shown an improvement in the accuracy of predictions, thereby yielding positive results with the LSTM model proving to be more efficient. In future, more functionalities and indicators will be integrated into the system. Further data analysis or data science components will be emphasized and added.

### REFERENCES

- [1] *How to Use Stock Data for Data Science Education: "A Simulated Trading Platform in Classroom"*, Xinxiang, China June. 12-14, 2020
- [2] *"Stock Market Prediction Using Machine Learning"*, 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC).
- [3] *"Deep Learning-Based Stock Price Prediction Using LSTM and Bi-Directional LSTM Model"* Proceedings of NILES2020: 2nd Novel Intelligent and Leading Emerging Sciences Conference.
- [4] *"Prediction of Stock Price Based on LSTM Neural Network"* 2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM).
- [5] *"Stock Market Behavior Prediction using Stacked LSTM Networks"*, Samuel Olusegun Ojo, Pius Adewale Owolawi, Maredi Mphahlele and Juliana Adeola Adisa\$ Department of Computer Systems Engineering.

- [6] "Market capitalization of listed domestic companies-world," accessed: 2020-06-17. [Online]. Available: <https://data.worldbank.org/indicator/CM.MKT.LCAP.CD?locations=IW>
- [7] A. G. Salman, Y. Heryadi, E. Abdurahman, and W. Suparta, "Single layer & multi-layer long short-term memory (lstm) model with intermediate variables for weather forecasting," *Procedia Computer Science*, vol. 135, pp. 89–98, 2018.
- [8] U. Singh, S. Chauhan, A. Krishnamachari, and L. Vig, "Ensemble of deep long short term memory networks for labelling origin of replication sequences," in *2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. IEEE, 2015, pp. 1–7.
- [9] R. M. I. Kusuma, T.-T. Ho, W.-C. Kao, Y.-Y. Ou, and K.-L. Hua, "Using deep learning neural networks and candlestick chart representation to predict stock market," *arXiv preprint arXiv:1903.12258*, 2019.
- [10] Colah, *Understanding LSTM Networks*, 2015, [online] Available: <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>.
- [11] Krollner, Bjoern, Bruce Vanstone, and Gavin Finnie. *Financial time series forecasting with machine learning techniques: A survey.* (2010)