

Enhancing Stock Market Prediction with ARIMA and Machine Learning

Miss. Sakshi G. Gade¹, Dr Shabnam F. Sayyad²

Research Scholar, Department of Computer Engineering, AISSMS College of Engineering, Pune.

Assistant Professor, Department of Computer Engineering, AISSMS College of Engineering, Pune.

Abstract

ML and ARIMA-based stock market forecasting. Brokers, derivatives, currencies, and stocks are used in the vast and convoluted structure of the financial markets, which is always changing and evolving. Compared to the hazards of beginning a new business or the requirement for a high-paying employment, this market provides investors the potential to make money and live a happy life with a minimal initial investment. However, assessing and managing the performance of machine learning requires human-assessed risk management procedures and security precautions. For this research, it is important to predict stock prices using ARIMA and machine learning techniques. With the aid of machine learning and the ARIMA model, stock values may be forecasted with simplicity. This includes a range of work that was done on the review paper using different learning techniques. The most notable features are ARIMA and the built-in machine learning. Oblivion Gate removes data that doesn't match the algorithm, leaving only data that does. As soon as information enters the network, rules enable selection. Three gate structures combine to produce a single network structure.

Keywords: LSTM, CNN, Machine Learning , Trade open, Trade close, Trade low, Trade high

INTRODUCTION:

The topic of the paper is stock market forecasting using ML and ARIMA. On the financial markets, which are dynamic and complicated systems, people can buy and sell stocks, currencies, derivatives, and stocks through broker-supported virtual platforms. This market gives investors the possibility to generate money and lead happy lives with a little initial commitment, as opposed to new business risk or the requirement for a high-paying employment. However, using human-judged risk management techniques and security precautions is necessary for assessing and managing machine learning performance. In practical contexts, such as predicting the weather and the financial market, time series forecasting is a widely utilized technique. The introduction of stock market forecasting makes use of the material in the study. High stock market volatility and uncertainty are caused by a variety of circumstances. In this study, stock price forecasting involves the use of ARIMA and machine learning techniques. Making stock price predictions aims to help investors make better informed and precise investment decisions. This paper includes numerous works that were completed for the review paper using diverse teaching techniques. Machine learning and ARIMA will be primarily involved. There are two different stock types. You may be familiar with intraday trading through the phrase "day trading." Interday traders frequently hold securities positions for multiple days up to weeks or months, but at least from one day to the next. Because they have the capacity to store historical data, LSTMs are particularly effective at solving sequence prediction issues. This is significant in our situation since a stock's

3. Indian stock market forecast using ANN

According to Hiransha et al. (2018), many people are either directly or implicitly involved in this field. Therefore, it is crucial to understand industry trends. As a result, as the stock market has expanded, so has interest in stock price forecasting. An exchange is a digital marketplace where derivatives and company stocks can be traded at predetermined pricing [8].



Source - <https://www.slideshare.net/RAJSHEKHARPATIL2/stock-market-prediction-technique>

Some suggestions related to stock market prediction , indian stock market forecast -

Due to the dynamic character of the market and its sensitivity to rapid changes, stock price forecasting is difficult. Previous research by Stock M has demonstrated that efficient techniques for learning event representations that can effectively capture semantic and syntactic data about text corpora exist, and that these techniques can be applied in the future for tasks like script event prediction [4]. Demand and supply for stocks determine the stock market. The stock market is one of the industries that is expanding the fastest in any country. Many people are now working in this industry, either directly or indirectly. Using ARIMA, analyse stock marketing.

Dhyani, et al. (2020) : Assert that RNNs are capable of understanding temporal dependencies. Long-term forecasting abilities are further improved by cells. Knowing how the price correlations of two assets will evolve over time is essential for portfolio optimisation. ARIMA models are also used to combine linearity and nonlinearity in the model. The ARIMA model extracts the information's clear patterns and feeds the residuals to the model [1]. In our empirical investigation, the ARIMA model outperforms all other financial models in terms of predicted accuracy. ARIMA hybrid models are compared to conventional predictive financial models such complete history models, constant correlation models, single index models, and multi-group models [10]. The ARIMA model should be considered by the author when predicting correlation coefficients for portfolio optimisation, according to research.

METHODOLOGY:

Multilayer perceptrons, convolutional neural networks, naive Bayes networks, backpropagation networks, single-layer LSTMs, support vector machines, and recurrent neural networks are a few examples of artificial intelligence approaches. Artificial intelligence-based techniques and statistical methods are the two primary categories of forecasting methodologies highlighted in the author's study (Ghosh et al., 2019). Statistical methods include models like ARCH and logistic regression.

By selectively erasing data that does not fit its criteria, the Oblivion Gate programme keeps only pertinent information. To decide which information is allowed into the system, the network uses rules. Three "gate" components make to the network's distinctive structure [6]. The LSTM unit is made up of these three parts: the forget gate, input gate, and output gate. This document's experimental data is made up of real historical documents that were downloaded from the Internet.

TABLE 3.1 Summary of the LMS Algorithm

<i>Training Sample:</i>	Input signal vector = $\mathbf{x}(n)$ Desired response = $d(n)$
<i>User-selected parameter:</i>	η
<i>Initialization.</i>	Set $\hat{\mathbf{w}}(0) = \mathbf{0}$.
<i>Computation.</i>	For $n = 1, 2, \dots$, compute
	$e(n) = d(n) - \hat{\mathbf{w}}^T(n)\mathbf{x}(n)$
	$\hat{\mathbf{w}}(n + 1) = \hat{\mathbf{w}}(n) + \eta\mathbf{x}(n)e(n)$

Because nonlinear methodologies are constantly presented on the internet, LMS is used to show that linear algorithms can be utilised to generate reliable stock market predictions [11]. Because nonlinear methodologies are constantly presented on the internet, LMS is used to show that linear algorithms can be utilised to generate reliable stock market predictions [11].

The Oblivion Gate programme preserves only relevant information by selectively eliminating data that does not meet its criteria. The network utilises rules to determine which information is permitted entry into the system. The characteristic structure of the network consists of three "gate" components. The forget gate, input gate, and output gate are the three components that make up an LSTM unit.

Actual historical papers that were retrieved from the Internet make up the experimental data in this document.

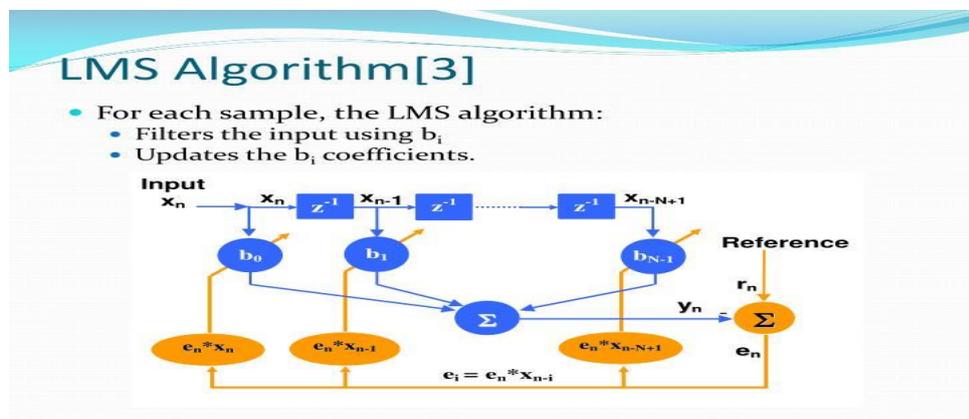


Figure-1 (Source- Slideserve)

Because nonlinear methodologies are constantly presented on the internet, LMS is used to show that linear algorithms may be utilised to generate precise stock market predictions.

Design :

The many designs that are used in this project are covered in this section, including cooperation diagrams and sequence diagrams. Some of these diagrams' designs and structures are covered in detail.

Software engineers and business executives frequently use sequence diagrams. Developers can better understand the underlying system needs, new components, and current documentation, with the help of these diagrams. They demonstrate how a set of items interact, as shown in Figure 4. The class of integrated diagrams includes sequence diagrams.

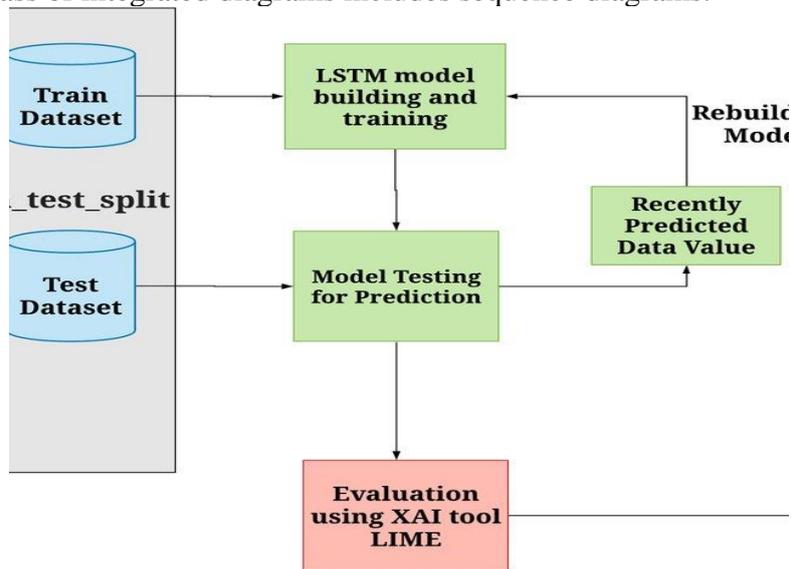


Figure 2 - Execution model (Source- Researchgate.net)

1. Collaboration flow chart -

Collaboration diagrams can be used to depict the specific unique behaviours that the Indian stock market demonstrates. Collaboration diagrams show the actions that objects take and are essential sources of knowledge for identifying roles within a class and related linkages [5]. Collaboration diagrams are very helpful when it's crucial to show relationships between objects.

Architect collaboration is used in addition to sequence diagrams to explain the importance of various factors and describe the responsibilities that control the particular event flow in the current situation. Information is represented differently in collaboration diagrams and sequence diagrams.

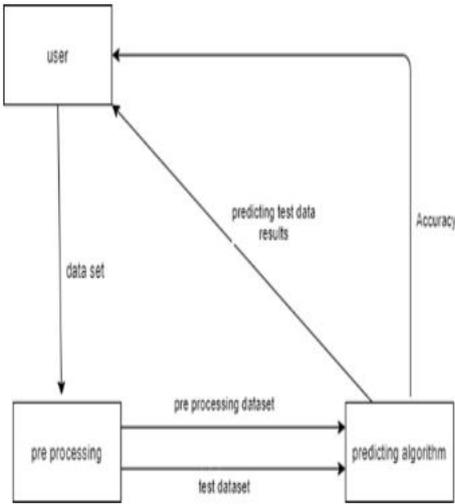


Figure 5: Data transfer between modules

RESULTS:

The testing portion of the training phase typically lasts 10 to 15 minutes [7]. On the other hand, the prediction procedure and accuracy evaluation only take a short while. The sample input and output for several datasets are shown below:

Google-

Attribute Name	Min	Max
Open	87.74	1005.49
Low	86.37	996.62
High	89.29	1008.61
Close	87.58	1004.28

Table 1: Google Dataset

Nifty50

Attribute Name	Min	Max
Open	7735.15	12932.5
Low	7511.1	12819.35
High	8036.95	12948.85
Close	7610.25	12938.25

Table 2: Nifty50 Dataset

Reliance

Attribute Name	Min	Max
Open	205.5	3298.0
Low	197.15	3141.3
High	219.5	3298.0
Close	203.2	3220.85

Table 3: Reliance Dataset

Sample Input:

	Trade High	Trade Low	Trade Open	Trade Volume	Trade Count
0	214.23	214.14	214.15	1022241	2274
1	214.38	214.14	214.15	582984	1902
2	214.37	214.18	214.37	705964	1943

Table 4: Sample input

Time series forecasting is a commonly utilised technique in many real-world applications, such as predicting weather and financial market movements. It entails making predictions about the future by using reliable historical data collected over an ambiguous period of time [2]. Multiple variables contribute to the volatility and unpredictability of the stock market. Although humans still have the ability to place and transmit orders to the market, automated trading systems powered by computer programmes can outperform humans in order execution. For this project to successfully anticipate stock prices, the use of ARIMA and machine learning techniques is essential [1]. Making stock price predictions is made easier by machine learning.

CONCLUSION:

The analysis and experimentation processes are made simpler by the incorporation of machine learning and ARIMA, giving useful tools for the review paper. The analysis is more effective when these techniques are used. The foundational ideas of stock pricing are covered in this review essay. The application of machine learning and ARIMA will improve the pricing analysis of the stock market, it can be concluded after looking at the significant and influential components of this research.

FUTURE SCOPE:

The goal is to assess the relative significance of more recent and older data and identify the factors that have the greatest impact on prices on the "current" or "next" day. Each market feature is given a weight by the machine learning model, which also decides how much historical data should be considered in order to forecast future stock values. Natural language processing (NLP) and image processing are two examples of machine learning approaches that can be used to speed up data collecting and format data. This can improve the ability to spot clinical patterns and help with more accurate forecasts.

REFERENCES:

- [1] Raj, J.S. and Ananthi, J.V., 2019. Recurrent neural networks and nonlinear prediction in support vector machines. *Journal of Soft Computing Paradigm (JSCP)*, 1(01), pp.33-40.
- [2] Ali, S., Shahzad, S.J.H., Raza, N. and Al-Yahyaee, K.H., 2018. Stock market efficiency: A comparative analysis of Islamic and conventional stock markets. *Physica A: Statistical Mechanics and Its Applications*, 503, pp.139-153.
- [3] Sudha, G., Shiny, T.V., Reddy, C.R., Irfan, S., Dharan, K.N. and Kumar, J.V., 2022. Stock Price Prediction Using Machine Learning. *JOURNAL OF ALGEBRAIC STATISTICS*, 13(3), pp.2080-2086.
- [4] Huy, D.T.N., Loan, B.T.T. and Pham, T.A., 2020. Impact of selected factors on stock price: a case study of Vietcombank in Vietnam. *Entrepreneurship and Sustainability Issues*, 7(4), p.2715.
- [5] Tursoy, T. and Faisal, F., 2018. The impact of gold and crude oil prices on stock market in Turkey: Empirical evidences from ARDL bounds test and combined cointegration. *Resources Policy*, 55, pp.49-54.
- [6] Jain, S. and Gupta, M., A 2018 Back Propagation Based Deep Neural Network for Stock Market Prediction. *Journal homepage: www.ijrpr.com ISSN*, 2582, p.7421.
- [7] Moghar, A. and Hamiche, M., 2020. Stock market prediction using LSTM recurrent neural network. *Procedia Computer Science*, 170, pp.1168-1173.
- [8] Poongodi, M., Vijayakumar, V. and Chilamkurti, N., 2020. Bitcoin price prediction using the ARIMA model. *International Journal of Internet Technology and Secured Transactions*, 10(4), pp.396-406.
- [9] Ghosh, A., Bose, S., Maji, G., Debnath, N. and Sen, S., 2019, September. Stock price prediction using LSTM on Indian Share Market. In *Proceedings of 32nd international conference on* (Vol. 63, pp. 101-110).
- [10] Hiransha, M., Gopalakrishnan, E.A., Menon, V.K. and Soman, K.P., 2018. NSE stock market prediction using deep-learning models. *Procedia computer science*, 132, pp.1351-1362.