IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 8, August 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 8, August 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1008024 |

Stock Price Prediction Using ML

Md. Irshad Hussain B[1], Rakesh S D[2]

[1]Assistant Professor, Department of Master Applications, University B.D.T College of Engineering, Davangere, Karnataka, India

[2]Department of Master Applications, University B.D.T College of Engineering, Davangere, Karnataka, India.

ABSTRACT: Market forces cause daily changes in stock values (supply and demand). One of the primary stock market bets that investors now make is based on stock price forecasting, whose popularity has increased recently.Investors utilise strategies and tactics to forecast stock market values in an effort to maximise profit and reduce risk.ML are used to predict stock prices. Stock market forecasting has modernized traditional trading strategies and entered a more technologically advanced era. As market capitalization continues to rise, stock trading has become an important investment area for many financial investors. In order to forecast stock price movements and guide investors in making wise choices, several analysts and academics have devised approaches and systems. Researchers are now able to forecast markets using unusual textual data from social media thanks to sophisticated trading models. By using modern machine learning techniques such as text data analysis and ensemble methods, the prediction accuracy has been greatly improved. He was executed in five stages as follows by Collecting data, pre-processing datasets, extracting features, predicting stock prices using various techniques, and displaying results. In the first stage, data is collected from various social websites and historical company data. The second stage involves cleaning up the pre-processing stage's mistakes, duplication, and debris. Reducing the data set and choosing relevant data constitute the third stage. Various machine learning techniques and approaches, divided into supervised and unsupervised learning techniques, are used to make predictions at the fourth stage. Accuracy is assessed using a variety of methods at the final stage.

KEYWORDS: ARIMA model

I.INTRODUCTION

Researchers in this discipline are always looking for methods to enhance current forecasting models since forecasting is still an important study topic. This is because organisations and individuals have the freedom to choose how much money to allocate to various projects as well as to create and develop effective strategy for their current and upcoming projects. One of the most difficult issues in financial forecasting is the prediction of stock prices due to the complexity of the stock market [1, 2, 3]. Many investors want to adopt a forecasting strategy that will enable them to earn from stock market investments with ease and with the least amount of risk possible. This still serves as a driving force for academics to create fresh prediction models [4].

Several models and methods have been developed recently toforecast stock values. Among these are examples of artificial neural network models (ANNs). Because it can identify patterns in data and glean conclusions from unknowable data, it is particularly well-liked. [5, 6, 7] are related papers that link ANN models to stock price prediction. By utilising their distinct advantages, hybrid techniques have also recently been used to enhance stock price forecasting models [2]. ANNs are from the viewpoint of artificial intelligence. From the viewpoint of statistical models, ARIMA models.

Prediction may generally be done from two angles, according to the literature: statistical approaches and artificial intelligence techniques [2]. In comparison to the most widely used ANN approaches, For predicting financial time series, especially for short-term forecasting, ARIMA models are seen to be more accurate and dependable.([8, 9, 10]). They are extensively utilised in the economics and finance industries. Exponential smoothing, generalised autoregressive conditional heteroscedasticity, and regression techniques are other statistical models (GARCH). His ARIMA model has been used in related investigations to make predictions, including [11, 12, 13, 14, 15, 16].

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 8, August 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1008024 |

This white paper outlines a thorough procedure for developing ARIMA models for predicting short-term stock prices. The effectiveness of his ARIMA model to give investors with short-term projections that might support the investment decision-making process was proved by results acquired from real-world data.

II.LITERATURE SURVEY

Over the last year, several researchers have employed clever ways and strategies to make decisions in the stock market.

B.P. Vijaya Kumar According to research by et al.[17], a low-cost and effective stock price forecasting model employing social media sentiment analysis is essential for leveraging machine learning skills to predict stock prices. did. The smallest discrepancy between the value predicted by the model and the daily stock value is what is most important.

This artificial neural network was thought to be useful for issues like arithmetic, research, and similarity discovery by Sunil Kumar Khatriet al. [18]. In this article, Twitter data is used to assess emotional evaluations. The performance and market values of his top five IT departments were utilised to construct an artificial neural network in this study utilising sentiment analysis data from Yahoo. The future return on investment is important to businesses.

Sensex and Nifty were identified by Aditya Bhardwajet al. [19] as his two primary indicators for forecasting Indian stock market conditions. By frequently examining and cross-referencing Senex Point and Nifty projections, shareholders and investors may have a sense of the situation of the stock market right now. To learn about the stock market sentiment and market circumstances, this article gathers live server data from the Sensex and Nifty at different time intervals. Researchers employ the Python programming a rapid implementation environment and a language to helpin deciding which stocks to buy.

According to Eyeman E. Kedle et al. [20], investors regard the stock market as a rich mine of information. The author employs two different forms of news material in this strategy. Financial reports written by financial experts via the publishers of market news, business news, and stocks.

According to Professor G.S. Mate and colleagues[21], there may be a correlation between stock price predictions and financial news. For this article, stock indices and news data were taken through his NY Times Archive API.

Sentiment analysis, or natural language processing-based identification of emotions and actions, is the most popular use of the Python Natural Language Toolkit module. Because you can determine that this news effect is favourable for the stock price, the stock price is more probable if the news is positive. Bad news may also cause a drop in stock values. In order to forecast stock prices, machine learning models are given the sentiment analysis data.

Based on time series forecasting techniques like ARIMA, RNN, and Facebook Prophet, Saloni Mohan and others [22] created forecasting models. In this study, we forecast stock values using time series models, neural networks, and news sources. Regular stock prices of S&P 500 firms are included in the dataset in this document overfive years and press stories on the businesses. With RN, authors have seen improved results.

A fresh approach to forecasting stock market trends was put out by Jing Zhang et al. [23]. It may forecast intervals of growth (or drop) rates in addition to financial performance throughout a predetermined forecast period. The model, according to the report, divides the raw user data from each stock is divided into a number of clips from a particular time period and classified as up, down, flat, or unknown heuristic Organize into four groups. Several current methods are outperformed by the suggested model.

III.METHODOLOGY

Inventory forecasting approach

A time series is a collection of measurements that are recorded at regular intervals. Time series may be categorised into the following groups according to frequency:

Every year, quarter, month, day, hour, minute, and second.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 8, August 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1008024 |

ARIMA Algorithm:

There are three components to the ARIMA algorithm. These are consolidation, moving average, and autoregression. These elements' main function is to create a satisfactory match with the given data. The data are unsuitable for modelling or prediction since the model is utilised for non-stationary data (means, variances, and covariances might fluctuate over time). Additionally, it results in inaccurate data being shown (which suggests that there may or may not be correlations between variables) [24]. Typically, ARIMA may be represented by the three parameters "p," "d," and "q." Non-negative integers are required for these parameters [23].

The order of the autoregressive model is denoted by "p."

Degree of Differentiation is denoted by "d,"

while the order of the moving average model is denoted by "q."

A popular notation is ARIMA (p, d, q), which allows references to be made based on the non-null parameters even if two of the three parameters are null. The generalisation of the ARIMA model is as follows: for drift values $\delta/(1 - \Sigma\varphi i)$

$$\left(1-\sum_{i=1}^{p}\phi_i L^i\right)(1-L)^d X_t = \delta + \left(1+\sum_{i=1}^{q}\theta_i L^i\right)\varepsilon_t. \quad \text{..... [24]}$$

Calculating the difference between successive observations is how differentiation is accomplished.

Structure of the System

When predicting stock market variance using the ARIMA algorithm, data is first processed before the results are presented in table and chart style. Three key modules make up the system. These include the modules for historical data, current market analysis, and outcomes and projections. The data collecting and pre-processing module of the historical data module, the sentiment analysis module of the current market analysis, and the data training and forecasting module of the results and forecasting module are further split [24].

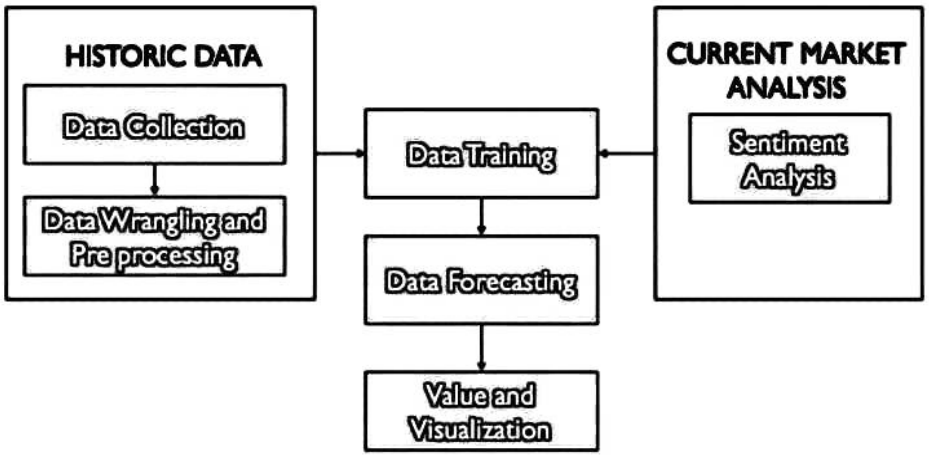

Fig 1: System Architecture

IV.IMPLEMENTATION

The entire project's backend is built using SQLite, NumPy, and Pandas, and it displays stock charts using the OpenCV library.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 8, August 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1008024 |

Live data is necessary for ML models, including news, volume, etc. For the programme to provide real-time stock charts, data is also required [24].

ML Models: The acronym ARIMA is defined by three order parameters, (p, d, and q) [24].

A regression model that makes advantage of the correlation between present and past period observations is called an autoregression, or AR(p). The time series regression equation's autoregressive (AR(p)) component refers to the usage of historical values [24].

I(d) Integral: Using the differential observation, make the time series stationary (subtracting the observation from the previous time step observation). When differentiating, d is repeatedly subtracted from the current value of the series [24].

Using the dependency between data and the residual error from a moving average model applied to staggered observations, the MA(q) Moving Average model is used. The model's error is displayed as a mixture of earlier error terms in the moving average component. The model's [24] term count is represented by the order q.

Types of ARIMA models

ARIMA: Autoregressive Integrated Moving Average (nonseasonal)

SARIMA: Temporary ARIMA

Seasonal ARIMA with exogenous variables is SARIMAX.

V.RESULTS AND COMPARISONS OF WORKS CARRIED OUT

The required modules and libraries have been imported. I imported the data andcreated many graphics.I have drawn the original series plot, the first derivative, the second derivative along with the autocorrelation plot. As can be observed, the time series reaches steady state in two different orders. However, looking at the autocorrelation plot of the second derivative, the lag very quickly enters the negative region, indicating that the series may be over-differentiated.

Data	Result
2016 January	Stock:Positive
2016 February	Stock:Positive
2016 March	Stock:Positive
2016 April	Stock: Negative
2016 May	Stock: Negative
2016 June	Stock: Negative
2016 July	Stock:Positive

Table 1: Data and Results of Stock Price

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 8, August 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1008024 |

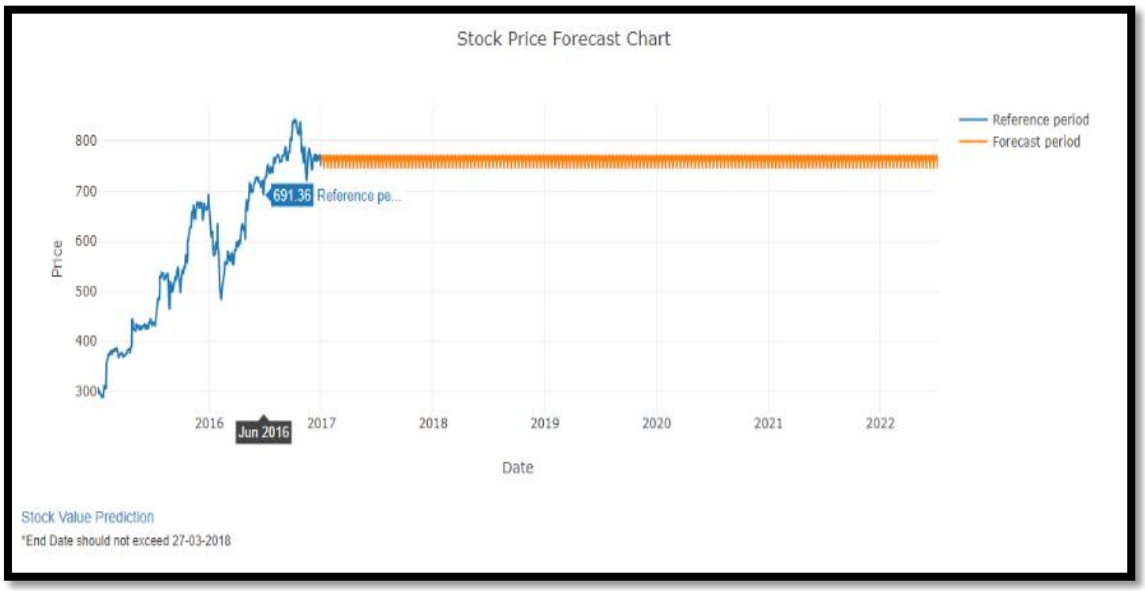

Fig 2. Price Forecast Chart

VI.CONCLUSION

A platform called Stock Prediction uses a variety of ML models, such as text-based and series-based models, to forecast changes in stock prices. Machine learning models' inventory forecasting accuracy is sufficient for use in real-world applications. News Vector data shows stock volatility so that users can invest effectively. Machine learning stock price prediction accuracy is currently 90%. Since most of the systems are open source, small traders don't have to spend time learning how affordable this tool can be for stocks.

Stock market research is strongly motivated by the idea that investors can make big profits by spotting stock movements in a split second. The financial industry, especially market forecasting models, has recently attracted a lot of interest as a result of the impressive development, success, and success of Natural language processing using machine learning and deep learning and sentiment analysis of Internet news.This study features recent research on text mining-based market forecasting systems, along with detailed explanations of the major shortcomings of the prototypes and recommendations for next upgrades.

REFERENCES

[1] P. Pai and C. Lin, "Hybrid ARIMA and Support Vector Machine Models in Stock Prediction," Omega Vol. 33, pp. 497-505, 2005

[2] JJ Wang, J.Z. Wang, Z.G. Zhang and S.P. Guo, “Stock Index Prediction Based on Hybrid Models,” Omega Vol. 40, pp. 758-766, 2012.

[3] L.Y. Wei, “ANFIS-based Hybrid Model and Adaptive Genetic Expectation Algorithm for Predicting TAIEX,” Economic Modeling vol. 33 pp. 893-899, 2013.

[4] G.S. Asarakis, E.M. Dimitrakakis. and CD Zopounidis, "Elliot Wave Theory and Neuro-Fuzzy Systems, Stock Market Forecasting: WASP Systems," Expert Systems with Applications, vol. 38, pp. 9196-9206, 2011. [5] S.K. Mitra, “Optimal Combination of Trading Rules Using Neural Networks,” International Business Research, vol. 2, no. 1, pp. 86-99, 2009. [6] G.S. Atsalakis and P.V. Kimon, "Predicting Short-Term Stock Market Trends Using Neuro-Fuzzy Methodology," Expert Systems with Applications, vol. 36, No. 7, pp. 10696-10707, 2009.

[7] M.M. Mohamed, “Predicting Stock Market Movements Using Neural Networks: Empirical Evidence from Kuwait,” Expert Systems with Applications, vol. 27, no. 9、S. 6302–6309、2010

[8] L.C. Kyungjoo、Y. Sehwanおよび J. John、「Neural Network Model vs. SARIMA Model In Forecasting Korean Stock Price Index (KOSPI)」、Issues in Information System、vol. 8-2, pp. 372-378, 2007.

[9] N. Mel, V.P. Saxena and K.R. Pardasani, “Comparison of his approach of ANN and ARIMA hybrids for predicting stock trends in India,” Journal of Business Intelligence, vol. 3, number. 2, S. 23-43, 2010.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 8, August 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1008024 |

[10] J. Sterba und Hilovska, "Implementing a Hybrid ARIMA Neural Network Forecasting Model for Total Water Consumption Forecasting," Aplimat-Journal of Applied Mathematics, BD. 3, number. 3, S. 123–131, 2010.
[11] C. Javier, E. Rosario, J.N. Francisco and JC Antonio, "ARIMA Models for Predicting the Next Price of Electricity", IEEE Transactions on Power Systems vol. 18 No. 3, pp. 1014-1020, 2003.
[12] N. Langan and N. Titida, "ARIMA Models for Predicting Oil Palm Prices", Proceedings of the 2nd IMT-GT Regional Conference on Mathematics, Statistics and Applications, Universiti Sains Malaysia, 2006.
[13] M. Khasel, M. Bijari und G.A.R Ardali, "Improving AutoRegressive Integrated Moving Average Models Using Fuzzy Logic," pp. 956-967, 2009. [14] C. Lee, C. Ho, "Short-Term Load Forecasting Using Lifting Methods and ARIMA Models," Expert System with Applications, Vol. 38, No. 5, pp. 5902-5911, 2011.
[15] M. Kashei, M. Bijari, G.A.R. Ardal, "Hybridization of Autoregressive Integrated Moving Averages (ARIMA) with Stochastic Neural Networks," Computers and Industrial Engineering, vol. 63, No. 1, pp. 37-45, 2012.
[16] C. Wang, "A Comparative Study of Fuzzy Time Series Models and ARIMA Models for Forecasting Taiwan's Exports," Expert System with Applications, Vol. 38, No. 8, pp. 9296-9304, 2011.
[17] N.N. Reddy, V. K. B. P. und N. E., "Predicting stock prices using sentimental analysis with Twitter data," UTC von IEEE Xplore, 2020.
[18] S.K. Khatri und A. Srivastava, "Using Sentimental Analysis in Prediction of Stock Market Investment," International Conference on Reliability, Infocom Technologies and Optimization, S. 566–563, 2016.
[19] A. Bhardwaj, Y. Narayan, P. and V, "Sentiment Analysis of Indian Stock Market Forecasts Using Sensex and Nifty," Procedia Computer Science, vol. 70, pp.85-91, 2015.
[20] A. E. Khedr and N. Yaseen, "Predicting Stock Market Trends Using Data Mining Techniques and News Sentiment Analysis," Intelligent Systems and Applications, vol. 7, pp. 22-30, 2017.
[21] GS Amidwar, R. Kulkarni und M. Muthya, "Stock Prediction with News Sentiment Analysis," Journal of Architecture & Technology, vol. 9, S. 36-40, 2019.
[22] S. Mohan, S. Mullapudi, S. Sameta,. V. Parag und D. C. Anastasiu, "Stock Price Prediction Using News Sentiment Analysis," International Conference on Big Data Computing Services and Applications, S. 205-208, 2019.
[23] J. Zhang, S. Cui, Y. Xu, Q. Li and T. Li, "A New Data-Driven System for Predicting Stock Trends," Expert Systems With Applications, pp. 60-69, 2018.
[24] Mukhopadhyay C.K. Kundu Amit (2008), His ARIMA Forecasts on Currency Exchange Rates of South Asian Countries, University Press, University of North Bengal, India.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.165

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details