Designing an Interface to Forecast Stock Close Price Using ARIMA

Shivdutt Dixit, Manoj Kumar Ahirwar, Damam Saketnath, Navneet Kaur

School of Computer Science and Engineering, Lovely Professional University, Punjab

Corresponding author: Shivdutt Dixit, Email: dxt_shiva@hotmail.com

Stock market prediction involves a great effect of the current market trend on the price. It is thus necessary to create a model that efficiently identifies the market trend and predicts the values with high consistency and accuracy. This paper focuses on designing a system that uses the ARIMA model to predict the values. The python flask API has been used to design the web application and tkinter API has been used to design the GUI interface. To fetch the live market data nse2r library of R and nsetools library of python has been used to fetch the past market data for a period of six years.

Keywords: ARIMA, stock prediction, NSE, flask, tkinter, financial forecasting, stock analysis, regression analysis, system design, short-term prediction

2023. In Saroj Hiranwal & Garima Mathur (eds.), *Artificial Intelligence and Communication Technologies*, 879–884. Computing & Intelligent Systems, SCRS, India. https://doi.org/10.52458/978-81-955020-5-9-83

1. Introduction

Stock market is one of the most hotshot topics under discussion today. With the increasing amount of investment being made, the impulse to maximize the yield with minimum investment risks has also increased drastically. To achieve this, various theoretical and practical approaches to predict the movement of the market have been made from time to time but the method that has gained most popularity in recent times is through extensive machine learning algorithms. Recurrent Neural Network (RNN), Convolutional Neural Network (CNN), Artificial Neural Network (ANN), Autoregressive Integrated Moving Average (ARIMA), Long Short-Term Memory (LSTM), and Multilayer Perceptron (MLP) have been some of the most widely used machine learning algorithms used to predict the stock prices with utmost accuracy. Whereas the outcome through neural networks has been a bit contentious mostly because the number of neurons and layers needs to be adjusted manually which can direly variate the results and due to the limitations in terms of time taken that they can pose, ARIMA has been more prominent and efficient in forecasting financial data and has shown the efficient capability for generating short-term predictions.

Developed by Box and Jenkins, it is an elicitation of Autoregressive Moving Average (ARMA) model and is applied on the data that has some nonlinearity in it. The AR part of the model uses regression and takes preceding output as input for the proceeding time step for the regression equation that it generates. The MA part of the model calculates mean for a given window which is then used to calculate the average for the next window. To overcome the non-linearity of the data, the I part of the model uses differentiation until a stationary value is not obtained. Each one of these attributes is aimed to support the model fit the data as precisely as possible. The functionality of ARIMA model can be expressed using the formula mentioned in equation 1 [2][5][6]:

$$y_t{}' = I + \alpha_1 y'_{t-1} + \alpha_2 y'_{t-2} + \cdots + \alpha_p y'_{t-p} + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_{t-q} \quad \text{-1}$$

where I is the data value substituted by d, α is the coefficient that is learnt from the model, θ is the weight, p is the lag order, and q is the moving average window. This paper is an effort to design a system by implementing ARIMA model on historical dataset and use live data to predict the Close value for the stock. The system for this paper has been well designed using the combined features and libraries from Python and R. All the stocks being traded on National Stock Exchange (NSE) and their respective stock codes have also been listed in order to make the search and retrieval easy.

2. Dataset Description

The data has been divided into two parts: one being the past data of the stock used for learning and the other being live data that will be used for predicting the values. Data for the training of the model is taken for past six years and then it is scaled and fitted into the model to be used for prediction. The past data for the selected stock is fetched from NSE using the nsetools library of python whereas the live data is fetched using nse2r library of R which is implemented in python using rpy2 library. The past data is filtered out initially, only to retain Close price of the stock. The live is continuously added to the past data and model is retrained every time to identify the trend for the day and increase the accuracy.

3. Literature Review

In prior research conducted by us to determine the best suited model for designing the interface for stock prediction [25], it was found that ARIMA model had the best accuracy when compared with MLP and LSTM. For the research conducted a MLP model with 1987 hidden layers with the maximum iteration of 300 and three dense layers having 2200, 2570, and 2800 neurons, respectively were used [23] and similarly and LSTM model with two simple input layers with 270 and 170 neurons, another layer of 155 neurons, a dense layer of 25 neurons and an output layer of single neuron was used [25]. Also, the ARIMA model used had a moving average window of 15, lag order of 0 and degree of differencing 1 [25]. There have been many efforts to using various models to predict the values for various entities most of them being stocks and indices. One such research has explained in detail various techniques that are being used for these predictions [19]. A group of researchers who applied different models on markets of different countries concluded that ARIMA was best for Indian market [1]. In another research best model equations were derived for stocks and indices and then were used to predict the prices of them [2]. In research that attempted to design a system for prediction using ARIMA [4] and another that tried to increase the efficiency of the model [3], found out that the model had the potential to predict prices with great accuracy [20]. A group of researchers also applied the model on data for COVID19 and designed various algorithms to predict the spread of virus [5]. It was concluded that ARIMA had slightly upper hand

when predicting for short-term. Various group of researchers applied various model on different stocks and indices and concluded that ARIMA predicted the price with utmost accuracy [6]. In research that used genetic algorithm with ARIMA, suggested the use of other techniques in combination with it [7]. One of the research projects also tried to combine SVR model with ARIMA to predict the prices [18]. Another research implemented the ANN model along with ARIMA to predict the electricity consumption. Various researchers have also used other models such as DNN, ANN, RNN, LSTM, Random Forest, MLP, Support Vector Regression etc. to predict the movement of market with great accuracy [10] [11] [13] [14] [16] [17] [21] [22][25][26]. Some researchers have also tried to predict the value of various entities by varying the p, d, and q values of the model [8] [9] [12] [15] [23] [24]. It has been seen that varying the value of these parameters shows significant variation in the prediction. In one research it was also discovered that employing a model methodology to choose a neural network model that would precisely forecast the value of stocks yielded better results than using the ARIMA model [10]. It is evident from various observations that suitable neural network selection for different stocks can vary significantly.

4. Methodology

The most important part of the design included training the model to analyse the past market trend and then predict the close value based on current market trend. To implement this, live market data was added to the past data and then the model was retrained every time and the predictions were updated simultaneously accordingly. The flow of control is defined as mentioned in Fig.1:

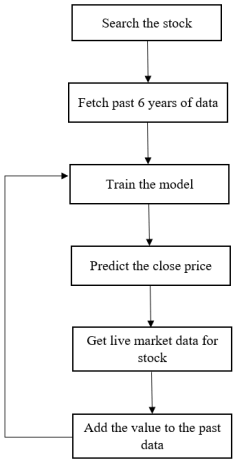

Fig 1: Flow chart for the interface

The stocks names and their respective stock codes of listed companies were fetched using the nsetools library. Once the user selected a stock the same library was used to fetch the past six years of data for that stock. The model was then trained on the entire dataset with the p value 0, d value 1 and q value 15 [25]. Before going to live predictions, an initial prediction was made based on the past trend. Then using nse2r library, live market data was fetched at an interval of eight seconds. This data was then added to the past data, the model was retrained, and the subsequent predictions were made accordingly. To facilitate the users, a graph of the data for past six years was also added to the interface. Also, the updated current price of the stock was continuously shown on the interface.

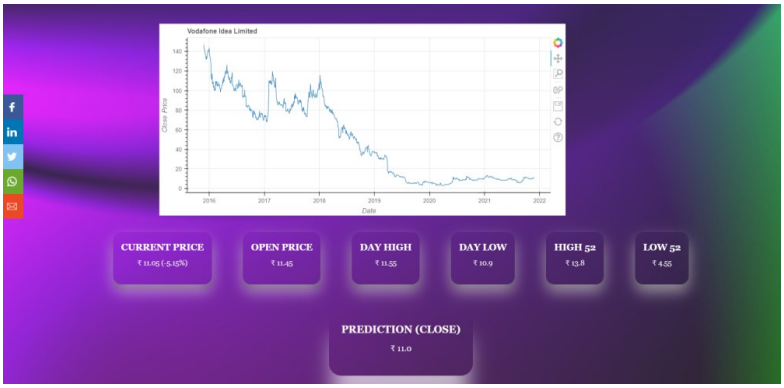

Fig 2: Web app interface

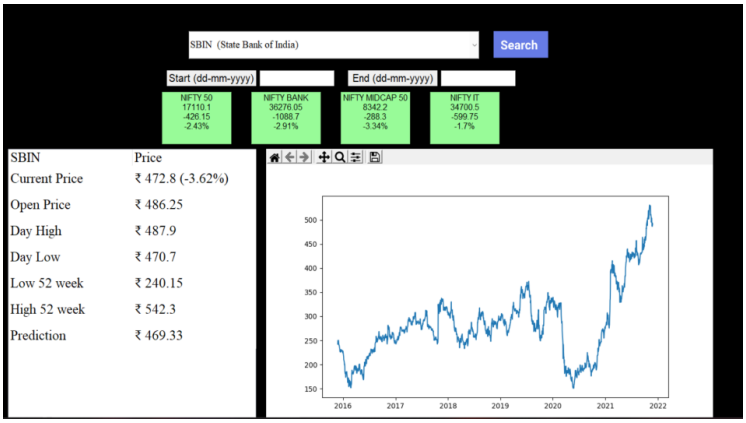

Fig 3: tkinter GUI interface

5. Experiments and Results

The model was tested on various stocks to check how much accurately it was able to identify the trend. Since the data that NSE APIs provide are snapshots of the data at certain interval and not the real-time updates, the trend identification took a bit of time. To smoothen the process of current trend identification, the interface was left running for three minutes initially and then the predicted close values were observed and compared with the next live data. It was observed that if the values were taken into consideration immediately, there were certain fluctuations in the predictions. Once enough data was collected, the fluctuations started to decrease and subsequently the predicted close values became more coherent with the next live data. Also, the stocks with great variation in the current market trend were hard to predict with more precision compared to the one's that had average variation in the live market trend. It was also observed for the stocks that had high price value in market, the deviation was a bit higher, and the time taken also increased for smoothing the trends.

6. Conclusion

This paper is an effort to design an efficient system for close price prediction of live data for a stock. The system design should be easy to use and fast to predict the values by identifying the current market trend. We observed that once ample data was available the close prices were much coherent

with the live market prices depicting that the model was effectively identifying the underlying trend in the data to predict the future values. Also, for the stocks that had high fluctuations in the price, the initial time taken to train the model had to be increased from three to around six-seven minutes to have efficient prediction. The main reason for this being the limitations that these libraries pose, which being the non-availability of data for the day. These libraries can only fetch the current price and not the entire price variation for the stock from the time market has opened to the time the stock was searched. Also, for stocks with high market price, the data required should be more extensive in order to identify the trends more easily and predict the prices with a larger accuracy and more quickly.

7. Future Scope

The system design does not include the data from the beginning of the trading hour of the stock being searched. Instead, it uses live data and creates its data set from the time a stock is searched to the time a new stock is searched. The model can work more efficiently if the data is also trained for the live market data from the trading hour till the time the value is searched. Also, ARIMA combined with other machine learning algorithms that have proven to work very efficiently on stock analysis can be implemented together as a new model for more precise prediction. This project focuses mainly on quantitative analysis of the trend in market but the fluctuation in values also depends on the fact that is there a positive or negative news running for a stock in market. If sentimental analysis is combined with the quantitative analysis of the stock, prediction can be made more accurate and concise.

References

[1] Ahmar, A. S., Singh, P. K., Thanh, N. V., Tinh, N. V., Hieu, V. M. (2022). Prediction of BRIC Stock Price Using ARIMA, SutteARIMA, and Holt-Winters. CMC-Computers, Materials & Continua, 70(1), 523–534.

[2] B, Uma & D, Sundar & P, Alli. (2013). An Effective Time Series Analysis for Stock Trend Prediction Using ARIMA Model for Nifty Midcap-50. International Journal of Data Mining & Knowledge Management Process. 3. 65-78.

[3] Goyal, Anil & Kundu, Amit. (2020). ARIMA and Indian Stock Market Forecasting. Xi'an Jianzhu Keji Daxue Xuebao/Journal of Xi'an University of Architecture & Technology. XII. 60-70.

[4] Ayodele Adebiyi, Aderemi Adewumi, Charles Ayo, "Stock price prediction using the ARIMA model", 2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation, UKSim, pp. 105-111.

[5] Markevičiūtė, J., Bernatavičienė, J., Levulienė, R., Medvedev, V., Treigys, P. et al. (2022). Attention-Based and Time Series Models for Short-Term Forecasting of COVID-19 Spread. CMC-Computers, Materials & Continua, 70(1), 695–714.

[6] Rangel González, Javier & Frausto-Solis, Juan & González Barbosa, Juan Javier & Pazos-Rangel, Rodolfo & Fraire-Huacuja, Héctor. (2018). Comparative Study of ARIMA Methods for Forecasting Time Series of the Mexican Stock Exchange.

[7] Vantuch T., Zelinka I. (2015) Evolutionary Based ARIMA Models for Stock Price Forecasting. In: Sanayei A., E. Rössler O., Zelinka I. (eds) ISCS 2014: Interdisciplinary Symposium on Complex Systems. Emergence, Complexity and Computation, vol 14. Springer, Cham.

[8] Jamal Fattah, Latifa Ezzine, Zineb Aman, Haj Moussami, Abdeslam Lachhab, "Forecasting of demand using ARIMA model", International Journal of Engineering Business Management, 2018.

[9] T. Alghamdi, K. Elgazzar, M. Bayoumi, T. Sharaf and S. Shah, "Forecasting Traffic Congestion Using ARIMA Modeling," *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*, 2019, pp. 1227-1232.

[10] Ankit Barai, Pooja Jain, Tapan Kumar, "NSE Stock Prediction: The Deep Learning Way", Soft Computing: Theories and Applications. Advances in Intelligent Systems and Computing, vol 1154, Springer, Singapore, 2020, pp. 978-981.

[11] Larsen, Jan. (2010). Predicting Stock Prices Using Technical Analysis and Machine Learning.

[12] Dadhich, Manish & Pahwa, Manvinder & Jain, Vipin & Doshi, Ruchi. (2021). Predictive Models for Stock Market Index Using Stochastic Time Series ARIMA Modeling in Emerging Economy.

[13] Mohamed, M. A., El-Henawy, I. M., Salah, A. (2022). Price Prediction of Seasonal Items Using Machine Learning and Statistical Methods. CMC-Computers, Materials & Continua, 70(2), 3473–3489.

[14] B. B. P. Maurya, A. Ray, A. Upadhyay, B. Gour and A. U. Khan, "Recursive Stock Price Prediction With Machine Learning And Web Scrapping For Specified Time Period," 2019 Sixteenth

International Conference on Wireless and Optical Communication Networks (WOCN), Bhopal, India, 2019, pp. 1-3.

[15] KumarMahto, Anil & Biswas, Ranjit & Alam, Afshar. (2019). Short Term Forecasting of Agriculture Commodity Price by Using ARIMA: Based on Indian Market.

[16] Z. Peng, "Stocks Analysis and Prediction Using Big Data Analytics", 2019 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS), Changsha, China, 2019, pp. 309-312.

[17] Mehar Vijh, Deeksha Chandola, Vinay Anand Tikkiwal, Arun Kumar, "Stock Closing Price Prediction using Machine Learning Techniques", Procedia Computer Science, Volume 167, 2020, pp 599-606.

[18] T. Ye, "Stock forecasting method based on wavelet analysis and ARIMA-SVR model," *2017 3rd International Conference on Information Management (ICIM)*, 2017, pp. 102-106.

[19] Shah, Dev & Isah, Haruna & Zulkernine, Farhana. (2019). Stock Market Analysis: A Review and Taxonomy of Prediction Techniques. International Journal of Financial Studies. 7

[20] Bijesh Dhyani, Manish Kumar, Poonam Verma, Abhisheik Jain, "Stock Market Forecasting Technique using Arima Model", International Journal of Recent Technology and Engineering (IJRTE), Volume-8 Issue-6, 2020, pp. 2694-2697.

[21] Ishita Parmar, Navanshu Agarwal, Sheirsh Saxena, Ridam Arora, Shikhin Gupta, Himanshu Dhiman, Lokesh Chouhan, "Stock Market Prediction Using Machine Learning," 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC), Jalandhar, India, 2018, pp. 574-576.

[22] S. Tiwari, A. Bharadwaj, S. Gupta, "Stock price prediction using data analytics," 2017 International Conference on Advances in Computing, Communication and Control (ICAC3), Mumbai, India, 2017, pp. 1-5.

[23] Mondal, Prapanna & Shit, Labani & Goswami, Saptarsi. (2014). Study of Effectiveness of Time Series Modeling (Arima) in Forecasting Stock Prices. International Journal of Computer Science, Engineering and Applications. 4, pp 13-29.

[24] Wibowo W., Dwijantari S., Hartati A. (2017) Time Series Machine Learning: Implementing ARIMA and Hybrid ARIMA-ANN for Electricity Forecasting Modeling. In: Mohamed A., Berry M., Yap B. (eds) Soft Computing in Data Science. SCDS 2017. Communications in Computer and Information Science, vol 788. Springer, Singapore.

[25] Isha, S. Dixit, M. K. Ahirwar, D. Sakethnath and M. Rakha, "Stock Prediction by Analyzing the Past Market Trend," *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, 2021, pp. 1-4.

[26] D. Shah, W. Campbell, F. H. Zulkernine, "A Comparative Study of LSTM and DNN for Stock Market Forecasting," 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 2018, pp. 4148-4155.