_____

Stock Market Analysis and Prediction Using ML

Abhishek Panda[1], Anusha S[2], Aisha Aslam[3], Annay Maity[4], Yerriswamy T[5]

[1,2,3,4,5]*School of Computing and Information Technology, UG Student Reva University, Bengaluru Karnataka, India*

Email: pupun123abhishek5@gmail.com[1], anushasahadev11@gmail.com[2], aisha.aslam1505@gmail.com[3], annaymaity.cse2017@gmail.com[4], yerriswamy.t@reva.edu.in[5]

*Corresponding author's E-mail: pupun123abhishek5@gmail.com

Article History	Abstract
Received: 06 June 2023 Revised: 05 Sept 2023 Accepted: 30 Nov 2023	*Seventy-Five percentage of Business pioneers state 'development' as the key wellspring of significant worth from investigation yet Just sixty percentage of these pioneers have prescient exam abilities. The solution to this is applying the correct arrangements of applications, which can work on huge bits of information. Machine Learning Calculations are major application which makes it possible. The place of this task is to break down through different expecting techniques to anticipate future stock benefits dependent on past data returns to build up a game plan of various stocks to know the varieties in the outcomes. We are doing this by applying machine learning strategy to determine stock value and provide the users with a better prediction and assist them with proper stock investments on web.*
CC License CC-BY-NC-SA 4.0	**Keywords:** *Machine Learning, Stock price analysis, Future Prediction, Web Application, Python*

1. Introduction

Nowadays, as the relationship between generally economies are fixed by globalization, external irritations to the financial business areas are not, now local. With propelling capital business areas, progressively more data is being made day by day.

The inalienable assessment of an association's stock is the value directed by surveying the ordinary future wages of stock and restricting them to the present, which is known as the book esteem. This is unquestionable from the market estimate of the stock, which is obliged by the association's stock cost. This market assessment of a stock can wanders from the characteristic worth as a result of reasons insignificant to the association's fundamental exercises, for instance, market assessment. The two fundamental stock exchanges India are NSE (National Stock Exchange) and BSE (Bombay Stock Exchange) which are composed by the Securities and Exchange Board of India (SEBI). The record of NSE is NIFTY – 50 and that of BSE is SENSEX. The stock developments and vacillations of the NSE stock value variety of an organization throughout a specific time. Stock markets are closely connected with the universe of economics — the rise and fall of share prices can be followed back to some Key Performance Indicators (KPI's). The five most normally utilized KPI's are the initial stock value ('Open'), finish of-day value ('Close'), intraday low value ('Low'), intra-day top value ('High'), and complete volume of stocks exchanged during the day ('Volume').

The difference of the protections trade is savage and there are many frustrated financial pointers. A few people with wide experience and data can appreciate the meaning of the markers and use them to make extraordinary assumptions to get a fortune. Many people rely solely on karma to make money in the stock market. Notwithstanding, the progression in innovation gives a chance to acquire a consistent fortune from the securities exchange and furthermore can assist specialists with discovering the most enlightening markers to improve expectations. Financial exchange forecast is an intense and testing action since the stock costs are straightforwardly affected by various variables. The segments fuse political and government decisions, budgetary news, public security events, etc. Public security events may be man-made events like dread put together oppressor practices or with respect to the next hand calamitous occasions or it very well may be a result of some pestilence. Lately, the scene of overall pandemic Novel Coronavirus (COVID-19) has exceptionally affected the protections trade wherever on the world. In India, the trading development had been ended for around 45 minutes in a remarkable

manner. In this way, it is important to foresee these vacillations as right on time as conceivable with the goal that the financial backers can settle on smarter choices. The estimate of the market regard is of principal importance to help in intensifying the advantage of venture opportunity purchase while keeping the danger low.

•   Related Work

"Application of Hidden Markov Models in Stock Trading" [1] Proposed a methodology which inspects the forecasting of the stock exchange list with hidden Markov models that consider hidden states inside the stock exchange file and the conventional ARIMA model. Hidden Markov Model considers the rearmost probabilities of various hidden states utilizing the Assumption Maximization calculation. The HMM model considers the end cost of the file when computing change states and rearmost probabilities at 2, 3, 4, and 5 hidden states. BIC (Bayesian Information Criterion) and the Akaike Information Criteria (AIC) was utilized to distinguish the HMM states.

"Hybrid ARIMA-BPNN model for time series prediction of the Chinese stock market"[2] This paper proposes another mixture ARIMA-BPNN model that incorporates specialized measurements to anticipate four individual stocks that incorporate both main board and expanding enterprise market in the product and data administrations industry. Tests show that the recommended arrangement outflanks the ARIMA, BPNN, and Khashei and Bijari crossover models for every arrangement regarding one-stride ahead determining accuracy.

"A Prediction Approach for Stock Market Volatility Based on Time Series Data" [3] Since financial business sectors are exceptionally competitive and vulnerable to sensational change, an essential point of value pattern research is to grow new imaginative methodologies for estimating stocks that produce high profit. This examination intends to inspect the stock exchange's time series data and build up a mathematical model that can precisely gauge future stock costs.

"Survey of stock market prediction using machine learning approach" [4] Used a system utilizing the notable productive relapse approach at estimating stock exchange costs utilizing stock exchange information. Regression analysis should be more profitable than conventional systems. Numerous regression models' outcomes could be improved later on as more factors are utilized.

"Stock Market Prediction with Historical Time Series Data and Sentimental Analysis of Social Media Data" [5] The proposed research expects to build up another strategy by joining sentiment analysis and traditional stock exchange forecast from time-series information utilizing deep learning procedures. It accumulates feelings from news occasions, online media pages, particularly Twitter, and consolidates extremity to improve forecast precision. The analysis and evaluation uncovered that applying the feeling extremity scores to the proposed arrangement improved forecasting precision.

"Stock market prediction using hybrid approach" [6] Proposed a model for foreseeing market valuation development utilizing an opinion mining and clustering approach. They utilized a market-specific approach to predict the best-promoted stocks from every space. This technique varies from past strategies in that it thinks about broad perspectives or feelings, and sentiments are joined into the stock prediction model of the specific objects of the organization or corporation. Utilizing the proposed approach, issues and related investor conclusions are subsequently eliminated from the pieces on a message board, in addition to isolating clusters of comparable kinds of stocks from those using clustering algorithms. This methodology would give two arrangements of results, one from sentiment analysis and the other from forecast dependent on clustering comparable to cutting edge stock exchange boundaries. By contrasting the entirety of the outcomes, a right estimate can be made. This report considers stocks with the biggest market capitalization in all significant enterprises for near purposes.

2. Materials And Methods
•   *System Design and Analysis*

•   Flow of work

•   *Software used*

The requirements can be broken down into 3 major categories namely functional, hardware and software requirements.

•   Functional Requirements:

It should provide the user with various companies to predict future returns on their stocks.

After performing prediction, it should provide the most positive result.

The user should be able to register on the website, so he/she can log in to his/her account anytime.

•   Hardware Requirements:

The hardware requirement is minimal, and the software can run with minimal requirements. The basic requirements are as enlisted below:

Processor: Intel Core2Duo processor or a processor with higher specifications

Processor speed: 1.5GHz or above.

RAM: 1GB or above

Storage space: 1GB or above

Monitor resolution: A color monitor with a minimum resolution of 640*480

•   Software Requirements:

•   An MS-DOS-based working framework like Windows 98/2000/XP/Vista/7/8/10/, Linux, macOS.

•   Python 3: A deciphered high-level generaly useful programming language.

•   XAMPP Server: It is a shortening for cross-stage, Apache, MySQL, PHP, and Perl, and it permits you to fabricate a WordPress webpage offline, on a nearby web worker on your PC.

•   Notepad++

•   *Software implementation*

As mentioned earlier, we look at a bunch of stock costs and utilize the likelihood strategy for Markov chains to anticipate the upsides of the stock expenses temporarily. Accordingly, to lead this work, a data set of such costs was first gathered, analyzed, and afterward the probability technique was applied. We have elected years value information for 5 totally different stocks and have applied Markoff chain calculations on this data so as to form the predictions.

Firstly, we make database entry. All the organizations years' worth of data is collected and entered in the database. We use MySQL in XAMPP as database for the project. the data are added in tabular form with the parameters of id, date, open, open_mov_av, open_diff, close, close_mov_av, close_diff, high, low. after adding the stock price of all the companies, we also add 2 more tables to the database, one to collect the login credentials and the other one to store message which is posted by the user.

When the stock costs are added, we need to apply Markov chains to the data set which will at long last leave us with the computation of moving avgs. The future stock prices are found with the help of moving avgs. The distinction between the forecasted costs and the real costs is utilized to make the future analysis for the value of stocks that will be in future. We consider the data at an interval of 3 days with which we calculate moving avgs first with opening cost and then with closing cost.

After the moving avgs were found for the arrangement of stock costs, the distinction between each real cost and moving avg cost of every day was determined which are later used to predict the future stock prices. When every day's value difference was determined, we then focused on binning every one of the distinction costs into four stretches set inside the bigger span from the most reduced difference cost to the most noteworthy difference cost. The bins were determined utilizing the accompanying equation:

K = sqrt(N)

where K is no. of all the bins present and N stands for total number of readings taken into consideration. The breadth of every bin is calculated mistreatment the subsequent formula,

w = (max value –min value) / K

where maximum difference and minimum difference are represented by 'max value' and 'min value' respectively. Quartile calculations are used to determine the intervals like as N/4, 2N/4 and 3N/4 and then those are intervals are marked as P1, P2, P3, P4 . When we finish setting up the intervals for all the data set that are considered at first, we calculate the difference in the cost and then they are named based on the interval in which they fall in.

When every subtracted cost was named with its comparing interval, the quantity of changes at every individual contrast cost interval to the following subtracted value stretch was counted by the python code. For instance, if Day 115's subtracted value had a place in interval P2, and Day 116's subtracted value had a place in interval P3, at that point a one value was added to the change from P2 to P3 (named

        Available online at: https://jazindia.com

for comfort as P23). Each such change from every interval, or state, was then counted by the system and recorded in the database. The quantity of points having a place with every interval was additionally recorded.

When every data above was recorded in system, a one-step progress matrix was fit to be made as per analysis. Every information of the matrix should be the probability of the data points moving from, or advancing from, one condition of information to another, with the states relating to the suitable rows and columns. To ascertain every passage of the matrix, the upsides of pij were separated by all out number of subtracted costs in the interval pi, which relates to the previously mentioned Pi.

When the transition matrix is calculated and build for every interval on which the moving avgs were made, the steady state probabilities can be studies. The probabilities of Steady state are discovered after calculating the linear systems with the multiplication of transition matrix with the πj vector. The steady states demonstrate the likelihood that the subtraction of the costs will be inside the previously mentioned intervals. After that it gives the value of where future subtracted costs may fall, and consequently giving us an good prediction of where the value of stocks may lie in future. The transition matrix looks like following form:

- Transition matrix

In this transition matrix each column represents the probability of being in a particular interval (calculated based upon the bin). For example, the first column shows the probability of the price lying in interval P1, the second column depicts the probability of the price lying in interval P2 and so on.

Consider Q as transition matrix. We found this technique after having our analysis done on various different methods to calculate steady state probabilities. Once we know the steady state we compute the values of Q2, Q3, Q4……., Q8. We stop at Q8 because we see that the probabilities take a constant value after 5 or 6 iterations. We have still considered till 8 iterations for issues of accuracy.

Then we have thought about the consistent state likelihood of being in a particular state as 1, that is, to find the probability or possibility of the stock price we form four steady state probability matrices that are situated at a particular range. The matrices are of the structure:

- Steady state matrices

When Q2, Q3, ……, Q8 are iteratively multiplied with Matrix A, we finally get the probability of the stock price lying in interval P1. Essentially, we get the probability of the stock price lying in interval P2, when matrix B is multiplied iteratively. Then probability of the stock price lying in interval P3, when matrix C is multiplied iteratively. And probability of the stock price lying in interval P4, when matrix D is multiplied iteratively

Till now what we have explained is about how to calculate the stock cost. As mentioned earlier to implement this we will be using Machine leaning and HTML. Hence, we have made use of HTML/CSS to make different web pages to represent the predicted prices of different companies. All the different kinds of HTML/CSS pages are created on Notepad++ and then converted into proper document with .html and .php extension.

After the web pages are created with the Notepad++ software we used XAMPP server to connect the database to all the HTML files within a folder. To make it happen we placed the Project folder inside the htdocs folder which is present inside xampp folder. We then open the first login page from the localhost in Apache server. Hence the project folder opens in the web page format and displays the content that we have entered inside the project folder with .html and .php extension.

3. Results and Discussion

These are the different database that are added to the SQL which was mentioned earlier.

- Database Tables shown in phpMyAdmin

The First page of the which will be visible to the user as soon as he/she goes to our website. This page gives two options to the user i.e. either to Login if the user already has the account with the website or to Sign Up to register itself with the website and create an account for future logins.

- First page

The Sign-Up page which client will actually want to see as soon as he/she clicks on Sign Up option on the first page. Here user enters the basic details.

- Sign-Up page

The Login page which user will be able to see as soon as he/she clicks on Login option on the first page. Here user has to enter their login details to login into his account.

- Login page

The Home page of the website which will be visible after user logs in. Various are available on the page to choose from like predict the Future, About the Project, our team, Contact details etc. User can click on any of these options as per their needs. Predict the Future page where user gets various stocks to choose from. Prediction for any of the available stock can be made just by clicking on the company logo.

- Home page

The opening price prediction of Apple as soon as the user clicks on the Apple company's logo.

- opening price prediction of Apple

The prediction of closing price of Apple as when user select that option.

- closing price prediction of Apple

In this application we can find the accuracy with which user can make a decision on investing in a certain stock. This is an example of Apple Stocks.

- Accuracy of Apple

After a detailed analysis the application recommends if its investing is the right choice. This is an example of Apple Stocks

- Recommendation for investment in apple

This is an example of Microsoft Stocks Accuracy.

- Accuracy of Microsoft

This is an example of Microsoft Stocks Recommendation.

- Recommendation for investment in Microsoft
- Applications

The project would be able to analyze, predict and represent the values of stock market based on the past observed values with the help of machine learning and HTML.

This project can be applied on various investment markets, like share market, cryptocurrency such as bitcoin.

With proper data collection and analysis and application of machine learning strategy, market price prediction would be easier. Users can search various share companies and they can get a future prediction of the share bases on its past lows and highs.

4. Conclusion
To deal with the master direction accessible on numerous kinds of media, this initiative will include mathematical statistics that will make it easy for an individual to conclude and without the burden of collecting all their grey cells. At the simple click of a finger, this web application can deliver forecasts. The products he/she wishes to spend on are what the customer must be worried about. Although no forecast can be called flawless, this software aims to reach near-perfection although making predictions to decrease an investor's chances of losing. Our project implements the use of Markov chains on the stock prices and then represents them in a web application. The user has to just open the web application and then choose which companies stock price user wants to see. Upon selection different data predictions are shown to the user.

Future Enhancement

There is a lot of room for improvement in our project in the future. When used in combination with the Markov model, the Hidden Markov model can increase accuracy. If we want to go ahead with our project, we should base our estimates on deciles rather than quartiles (i.e.,10 intervals instead of 4) Although no definitive statement can be made about any forecast, it tends to be gathered that the field of prediction based on such mathematical models has a lot of potential. Machine Learning & Artificial Intelligence has created different techniques like Sentiment analysis, Neural networks and Fundamental

analysis. These approaches help make predictions associated with markets remarkably faster. Sentiment analysis, neural networks, and fundamental analysis are some of the techniques developed by Machine Learning and Artificial Intelligence. These methods aid in making business forecasts in a remarkably short amount of time.

Acknowledgment

References:
1. A. Sharma, D. Bhuriya and U. Singh, "Survey of stock market prediction using machine learning approach," 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, 2017.
2. P. V. Chandrika, K. Visalakshmi and K. Sakthi Srinivasan, "Application of Hidden Markov Models in Stock Trading," 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2020.
3. L. Xiong and Y. Lu, "Hybrid ARIMA-BPNN model for time series prediction of the Chinese stock market," 2017 3rd International Conference on Information Management (ICIM), Chengdu, 2017, pp. 93-97, doi: 10.1109/INFOMAN.2017.7950353.
4. S. M. Idrees, M. A. Alam and P. Agarwal, "A Prediction Approach for Stock Market Volatility Based on Time Series Data," in IEEE Access, vol. 7, pp. 17287-17298, 2019.
5. M. Kesavan, J. Karthiraman, R. T. Ebenezer and S. Adhithyan, "Stock Market Prediction with Historical Time Series Data and Sentimental Analysis of Social Media Data," 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2020, pp. 477-482, doi: 10.1109/ICICCS48265.2020.9121121.
6. V. Rajput and S. Bobde, "Stock market prediction using hybrid approach," 2016 International Conference on Computing, Communication and Automation (ICCCA), Noida, 2016, pp. 82-86, doi: 10.1109/CCAA.2016.7813694.
7. Abhishek Panda, Yerriswamy T," Finger Gesture Mouse Using OpenCV", TEST Engineering and Management, ISSN: 0193-4120 Page No. 5018-5022, May - June 2020
8. S. A. Dwivedi, A. Attry, D. Parekh and K. Singla, "Analysis and forecasting of Time-Series data using S-ARIMA, CNN and LSTM," 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), Greater Noida, India, 2021, pp. 131-136, doi: 10.1109/ICCCIS51004.2021.9397134.
9. J. Qin, Z. Tao, S. Huang and G. Gupta, "Stock Price Forecast Based on ARIMA Model and BP Neural Network Model," 2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE), Nanchang, China, 2021, pp. 426-430, doi: 10.1109/ICBAIE52039.2021.9389917.
10. Dr. Lokesh, Yerriswamy T," A Proportional analysis of a collection of techniques in sequential rules mining and sequential patterns mining", International Journal of Mechanical and Production Engineering Research and Development, E-ISSN:2249-8001, Nov-2018.
11. G. W. R. I. Wijesinghe and R. M. K. T. Rathnayaka, "ARIMA and ANN Approach for forecasting daily stock price fluctuations of industries in Colombo Stock Exchange, Sri Lanka," 2020 5th International Conference on Information Technology Research (ICITR), Moratuwa, Sri Lanka, 2020, pp. 1-7, doi: 10.1109/ICITR51448.2020.9310826.
12. A. Site and Z. Işık, "The Analysis and Forecasting of Stock Price with Deep Learning," 2020 28th Signal Processing and Communications Applications Conference (SIU), Gaziantep, Turkey, 2020, pp. 1-4, doi: 10.1109/SIU49456.2020.9302045.
13. Dr. Lokesh, Yerriswamy T "Design, Improvement, Development, and Performance analysis of a collection of models developed from Naive Bayes and maximum entropy opinion mining classifiers for movie reviews", International Journal of Engineering and Technology, June-2018
14. N. Sirimevan, I. G. U. H. Mamalgaha, C. Jayasekara, Y. S. Mayuran and C. Jayawardena, "Stock Market Prediction Using Machine Learning Techniques," 2019 International Conference on Advancements in Computing (ICAC), Malabe, Sri Lanka, 2019, pp. 192-197, doi: 10.1109/ICAC49085.2019.9103381.
15. Y. Huang, W. Hsieh, Shih-Wei and Lin, "A Trend based Forecasting Model Using Fuzzy Time Series and PSO Algorithm," 2019 IEEE International Conference on Computation, Communication and Engineering (ICCCE), Fujian, China, 2019, pp. 21-24, doi: 10.1109/ICCCE48422.2019.9010899.
16. A. Fazeli and S. Houghten, "Deep Learning for the Prediction of Stock Market Trends," 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 2019, pp. 5513-5521, doi: 10.1109/BigData47090.2019.9005523.
17. Y. Xu and V. Keselj, "Stock Prediction using Deep Learning and Sentiment Analysis," 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 2019, pp. 5573-5580, doi: 10.1109/BigData47090.2019.9006342.