iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

PREDICT STOCK MARKET - USING LINEAR REGRESSION AND PYTHON PROGRAMMING

S. Thoufeeka
School of Computing Sciences, VISTAS, Chennai, India
shaheenthoufeeqa@gmail.com
Dr. K. Dharmarajan
Professor, School of Computing Sciences, VISTAS, Chennai, India
dharmak07@gmail.com

ABSTRACT.
The most valuable indicator of a company's success is its stock price, which can rise in tandem with the price of a single share. For that reason, businesses advertise their stocks to their customers in an effort to get them to buy them. The volatility of stock prices makes it difficult for clients or stockholding companies to forecast the future value of a single stock. Therefore, stock market forecasting has emerged as the most well- liked topic in the corporate sector, and hence solving this problem has become so important for the betterment of the investors and buyers as many a times they have to face loss in their investment and this problem can be solve by various Machine learning algorithms .To solve this problem we are developing one stock price prediction website using Python and Linear Regression which is one of the best Machine Learning statistical method for predictive analysis. We are using historical Data for the prediction. Finding a method to use linear regression models to obtain more precise values is the major goal. To acquire more precise results from the linear regression models, it is possible to change the dataset that will be used to train the models. The purpose of this paper is to demonstrate that linear regression is the most appropriate and effective method for forecasting stock market analysis.

Keywords
Machine Learning, Linear Regression, Python, Django framework, Yahoo Finance.

INTRODUCTION

The stock market is a regulated marketplace where investors can buy or sell stocks publicly or privately. Companies often turn to the stock market to raise capital for business expansion, making it a popular investment option for investors. To make informed investment decisions, many investors rely on predictions based on past market trends. In recent years, the stock market has changed significantly, making it crucial to anticipate its future value or price given how dynamic the market is. Predictive analysis is frequently carried out using linear regression, a mathematical technique and supervised machine learning method. The linear regression model, which generates linear correlations between independent and dependent variables, is largely consistent with the continuous/real values of mathematical variables. The algorithm makes the predictions in accordance with the guidance provided in the training data set after first training with it. The other tools and technology used in the given website are html and CSS which are used for frontend. The library named scikit learn is used as it has all machine learning algorithms and functions inbuilt in it. The framework Django which comes under Python is used at the backend to access web services and resources which is inbuilt in python programming language. To fetch the data for the website which should be reliable and accurate, Yahoo finance site is used and connecting it with API. So the prediction of stocks of all the companies which are there in yahoo finance can be done through our website.

Here are summaries of 10 research papers on stock price prediction using linear regression algorithm as a machine learning approach:

1) This review paper by Shruti Shukla and Bhavya Shah presents a multiple linear regression model for stock price prediction based on historical price data and financial indicators.
2) This research paper by Nusrat Rouf, Saurabh Singh etc. analysed machine learning algorithms used for stock price prediction which is linear regression. The authors highlighted the importance of feature selection and data pre-processing in improving the accuracy of the prediction models.
3) This research paper by Vaishnavi Gururaj, Shriya V R and Dr. Ashwini K compared the performance of linear regression and support vector machine models for stock price prediction based on historical price

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

data and technical indicators.

4) This review by Nils Karlsson compared the performance of multiple linear regression and artificial neural network models for stock price prediction based on historical price data and financial indicators. The authors found that the artificial neural network model outperformed the linear regression model.
5) This research paper by Payal Soni, Yogya Tewari, Deepa Krishnan analysed the state-of-the-art machine learning techniques used for stock price prediction, including linear regression. The authors emphasized the importance of incorporating external factors such as news sentiment and macroeconomic indicators in the prediction models.
6) This research Paper by Meher Vijh, Deeksha Chandola etc. depicts that linear regression machine learning technique is best for predictive analysis of statistical data. The authors found that the linear regression model is perfect for predicting stock prices.
7) This review by Kastberg, Daniel This study compared the performance of linear regression and random forest regression models for stock price prediction based on historical price data and technical indicators. The authors found that the random forest regression model outperformed the linear regression model.
8) This paper by Ogulcan E. Orsel, Sasha S. Yamada compared the performance of various machine learning algorithms, including linear regression, for stock price prediction based on historical price data and financial indicators, random forest regression model and LSTM.
9) This study presented a Multivariate Regression and Logistic Regression by Jaydip Sen.
10) This research paper by Indu Kumar, Chetna Utreja etc. (2018): This study assessed the effectiveness of different machine learning techniques, including Forestry and Support vector machines.

This website can predict the stock prize for next 1 day to the next 360 days through graphical representation. Predicting stock market behavior has been a difficult task for researchers, but with the help of machine learning, it has become possible to accurately forecast stock prices.

METHODOLOGY

In the given project, stock prices are predicted using Linear regression algorithm in machine learning and data implementation using python tools and libraries like Scikit-learn, Numpy are done.

Linear regression

A linear regression contains the labelled data (supervised learning) which generates the relationship between independent and dependent variables using simple mathematical equations and thereby calculate the best fit line or line of minimum reluctance. This line can be used to calculate stock predictions using graph or curve analysis. Linear regression is considered to be better than most of the other techniques as it is easier to Implement and contains basic mathematical and computational theory. It involves fitting a straight line to the given data points of independent variable(x) and dependent variable(y) whose slope is (m) and error is (e). This line is implemented by reducing the sum of squared differences between actual values understood in the above equations and diagrams-

$y = mx + c + e$ - - - (1) - - - - - - - - - - - - - - - - - - -

here, 'c' is the intercept formed on dependent axis y. for multiple data sets with slopes m_1, $m_2 \ldots m_k$ we can use-:

$y = m_1x + m_2x \ldots . m_kx + c + e$ (2)

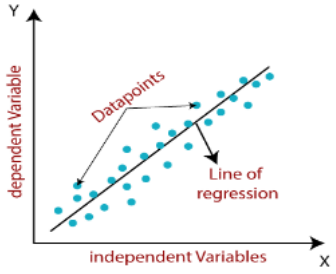


Fig. 1. Representation of best fit line of regression [16]

Steps for implementation of Linear regression includes-:

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

1) Importing necessary computing packages and creating the data from input files using pandas library.
2) Filter out the essential variables and generate a numpy array from pandas data frame to store continuous set of data.
3) Assigning the input variable (x,y) and performing recursive feature normalization , where we will compress our input variables to magnified values for faster speed and accuracy.
4) Plotting the data sets using matplotlib and plotly tools through normalized data and generate the graph as shown below-:

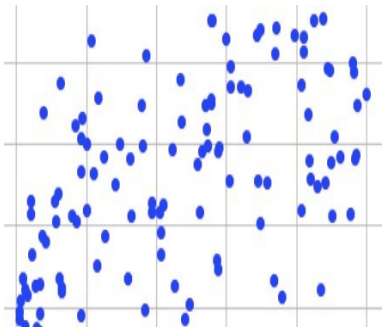


Fig. 2. Normalized stock market data

5) Training the algorithm and computing the hypothesis (approximate target value-y) for stock price vs time duration curve.
6) Calculating the hypothesis h_x and SSE (sum of squared error) $c\theta$ using the below equations-:

$$h_x = \theta 0\ x0 + \theta 1 x1 \qquad (3)$$

where $\theta 0$ and $\theta 1$ have minimal error or deviation variables for linear fit and $x0$ tends to 1. SSE which is difference between h_x and actual values can be calculated using-:

$$c_\theta = \tfrac{1}{2}\, m \sum_{\theta=0}^{n} (h\theta x - y)^2 \qquad (4)$$

by reducing the SSE, we can optimize our results and best fit line.
7) calculating the gradient descend and compute optimal $\theta 0$, $\theta 1$ which are required for plotting best fit line using below equation-:

$$\theta 1 = \theta 0 - \alpha\, \Omega\, c\theta \qquad (5)$$

Where $\theta 0$ and $\theta 1$ are current and next optimal positions, α is small step , Ω is direction and $c(\theta)$ is SSE..
8) model testing, plotting best fit line and graph, deployment-:

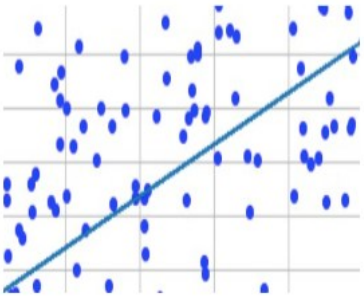


Fig. 3. Implementation of best fit line and predicted graph

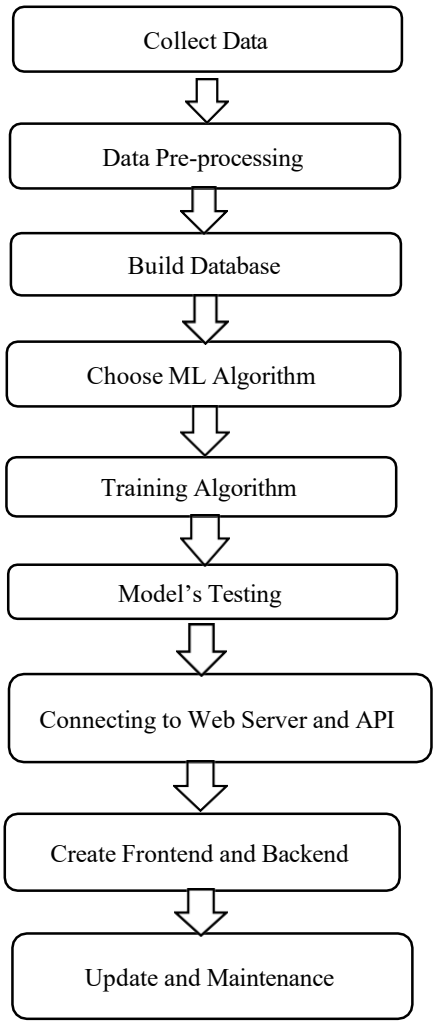
iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

Flow of stock market prediction website -

In order to develop GUI (graphical user interface) to make this predictor tool accessible to users, predictor is converted to a website which will show real time stock market prediction.

The first step in this is to collect and fetch the data from reliable source site Yahoo- finance and creating a DBMS (data base management system) using SQLite. Data preprocessing is yet another important step in which inaccuracy or inconsistencies in data had been removed and data is transformed into user understandable format using pandas library in python. In the backend part numPy, plotly, matplotlib are used for multiple operation execution. The prediction part is performed using linear regression technique of machine learning. To access the machine learning tools and resources for the website, scikit learn library and Django framework are used respectively. For UI/UX part Html5 and CSS3 versions are used. The website is maintained at regular intervals and fetch new and current status of data everytime.

Flowchart



iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

RESULTS AND DISCUSSIONS

The primary objective of this project was to predict stock market prices at a higher accuracy rates and it had been achieved at an approximate accuracy ranging in between 75% to 85 %. The stock market prediction website named "J3 predictor" can predict stock values from 1 day to 365 days (1 year). The website has stock data of multiple stocks connected to the data server through yahoo finance by API path. Website has ticker information of all the stocks available at the third tab, from which ticker value can be inserted into predictor page of second tab and get the results as shown below.

Symbol	ADANIGREEN.NS
Name	AdaniWilmar
Last Sale	408.3
Net Change	1.2
Percent Change	102320.83
Market Capital	India
Country	1750568
IPO Year	ConsumerFood
Volume	nan
Sector	nan
Industry	nan

Fig. 4. Display of stock information on result page

Fig. 5. Recent and prediction stock display for ADANIGREEN upto 90 days.

The results obtained at the first tab of the website are for top 6 companies and the remaining stock prediction occurs at prediction section. The comparison between actual stock prices and predicted prices for various NIFTY trading companies for 3 days duration from 10 may to 12 may are shown in table given below-:

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

Table 1. actual and predicated value

COMPANY NAME	ACTUAL VALUES	PREDICTED VALUES
Adani Green	896.45 INR	894.16 INR
Tata Motors	513.80 INR	515.70 INR
Axis Bank	911.00 INR	908.80 INR
Reliance Industries	2481.90 INR	2479.80 INR
State Bank of India	578.10 INR	575.00 INR

If the user enters wrong ticker symbol or wrong time duration then the predictor will give error that the data is inadequate. In this way, the stock prediction using machine learning has proven very useful for stock market prediction and can help many people who are newcomers to stock market or want to make money using stocks.

Helpful Hints

4.1. Figures and Tables

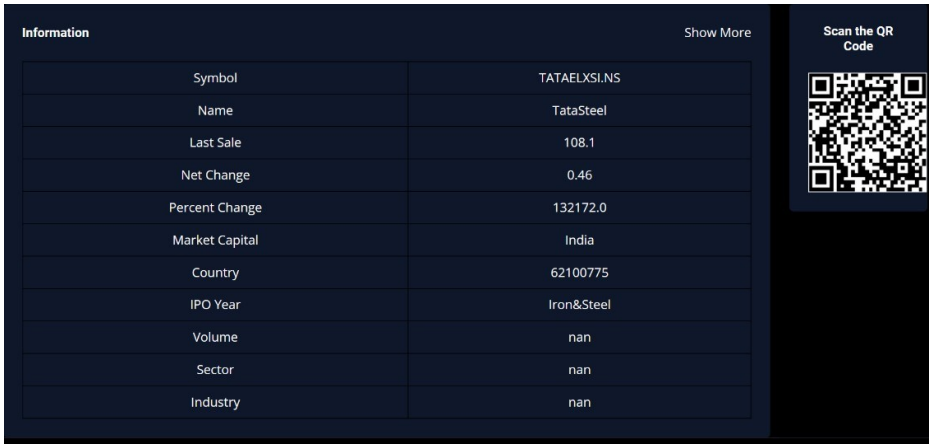


Fig. 6. Predicted page of Tata Elxsi

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

Fig. 7. Prediction Graph of Tata Elxsi

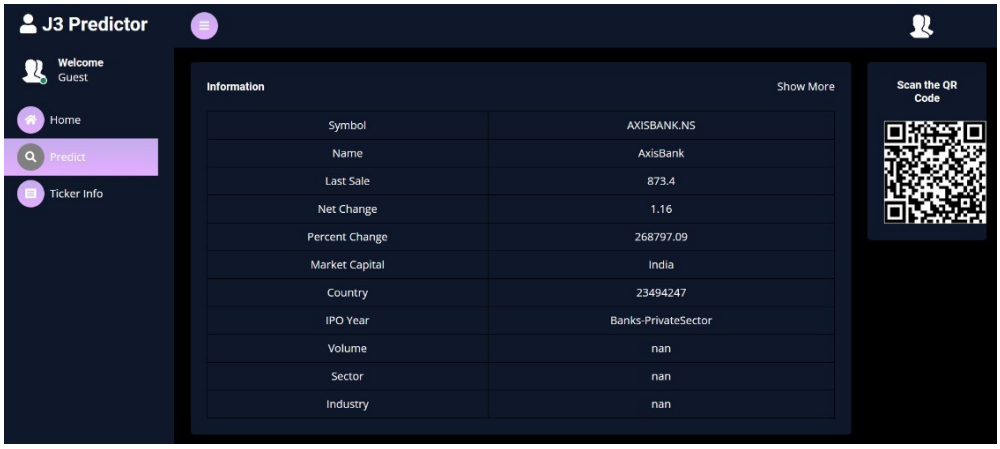


Fig. 8. Prediction page of AXIS BANK.

Fig. 9. Prediction graph of AXIS BANK

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

FUTURE SCOPE

Increased prediction accuracy. Future stock price projections are likely to be more accurate thanks to continuously improving machine learning algorithms

Information extraction from news articles and other text sources using natural language processing (NLP). A company's financial performance, competitive environment, and other elements that might affect its stock price can all be learned about using NLP. Then, using this data, machine learning systems can be trained to produce more precise predictions.

The application of machine learning to forecast how non-financial events would affect stock values. Natural disasters or political scandals are examples of non-financial events that can have a big impact on stock values. Machine learning can be used to forecast how these occurrences would affect stock prices, enabling investors to reduce risk in their portfolios.

The application of machine learning to forecast how climate change would affect the stock market. To reduce risk in investor portfolios, machine learning can be used to forecast how climate change would affect the stock market. Overall, the potential of employing machine learning to anticipate stock prices is very positive. Machine learning algorithms are improving in accuracy, and more data is being made available to train them. Additionally, new machine learning algorithms are always being created. Future stock price predictions are likely to be more accurate because of these considerations.

CONCLUSION

Many people desire to forecast future stock prices in order to increase their fortune because the stock market is continuously changing. However, due to the volatile nature of the stock market, current solutions that employ cutting-edge technology like Deep Learning, AI, and Neural Networks have not proven successful in making precise forecasts. As a result, the analysis of the current methods for stock market prediction using a linear regression algorithm is the main goal of this review work. A supervised machine learning approach called linear regression establishes a linear relationship between independent and dependent variables. The choice of an appropriate dataset is crucial for successful stock market prediction using linear regression, according to a review of research papers. According to the experiments, the linear regression method performed better in terms of accuracy than other machine learning techniques. However, a lot of experts also stated that they planned to investigate neural networks' potential for stock market prediction in the future.

ACKNOWLEDGEMENT

The research project received support from a renowned institution in Pune that is actively involved in the field of education and engineering sciences. The head of the Department of DESH, Prof. Chandrashekhar Mahajan, and the coordinator Dr. Sachin Sawant from this institution provided valuable assistance at every level of the project, and their contributions are highly appreciated. According to the experiments, the linear regression method performed better in terms of accuracy than other machine learning techniques. However, a lot of experts also stated that they planned to investigate neural networks' potential for stock market prediction in the future.

REFERENCES

- S.Shukla and B.Shah's "Stock Price Prediction Using Multiple Linear Regression Models"International Journal of Engineering Science Invention (IJESI). Published in October 2018.
- N.Rouf and S.Singh's "Stock Price Prediction Using Machine Learning Techniques: A Decade Survey on Methodologies, Recent Developments, and Future Directions" MDPI. Published in November 2021.
- V. Gururaj, S. V R and A.karthick's "Stock Price Prediction Using Linear Regression and Support Vector Machines".International Journal of Applied Engineering Research. Published in 2019.
- N. Karlsson's "Comparison of Linear Regression and Neural Network for Stock Price Prediction".
- P.Soni, Y.Tewari, D. Krishnan-"Machine Learning approaches in Stock Price Prediction: A Systematic Review", Journal of Physics Conference Series. Published in 2022.
- M.Vijh, D.Chandola's "Stock Closing Price Prediction using Machine Learning Techniques". International Conference on Computational Intelligence and Data Science (ICCIDS).
- K.Daniel's "A Study on Stock Price Prediction Using Linear Regression and Random Forest Regression" (2022).
- O. E. Orsel, S. S. Yamada-"Comparative Study of Machine Learning Models for Stock Price Prediction" (2022).
- J.Sen-"Stock Price Prediction Using Machine Learning and Deep Learning Framework", 6th International

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

Conference on Business Analytics and Intelligence (ICBAI). Published in December 2018.

- I.Kumar, C.Utreja's "A Comparative Study of Supervised Machine LearningAlgorithms for Stock Market Trend Prediction". IEEE Conference. (2018).
- M.Nikou, G.Mansourfar, J.Bagherzadeh's "Stock price prediction using DEEP learning algorithm and its comparison with machine learning algorithms". Intelligent Systems in Accounting, Finance and Management. Published in 2019.
- V.K. Sai Reddy's "Stock market prediction using machine learning". International Research Journal of Engineering and Technology (IRJET). Published in 2018.
- M. Obthong, N. Tantisantiwong, W.Jeamwatthanachai, G.Wills's "A survey on machine learning for stock price prediction: Algorithms and techniques".
- M. Umer, M. Awais, M. Muzammul- "Stock market prediction using machine learning (ML) algorithms" (ADCAIJ) Advances in Distributed Computing and Artificial Intelligence Journal. Published in year 2019.
- D. Kumar, P. K. Sarangi, R.Verma- "A systematic review of stock market prediction using machine learning and statistical techniques" Published in 2022.
- Javatpoint's-https://static.javatpoint.com/tutorial/machine-learning/images/linear- regression-in-machine-learning.png.