INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

e-ISSN : 2583-1062

Impact Factor : 5.725

www.ijprems.com
editor@ijprems.com

Vol. 03, Issue 10, October 2023, pp : 349-353

STOCK PRICE PREDICTION USING LINEARREGRESSION – A MACHINE LEARNING ALGORITHM

Vedashree Bhat[1], Soham Joshi[2]

[1,2]BE Computer) – All India Shri Shivaji Memorial Society's College of Engineering, Pune, Soham Joshi (SE Computer) – Vishwakarma Institute Of Technology, Pune, India.

ABSTRACT

The value of a company's stock, which can increase along with the cost of an individual share, is the best measure of its success. Because of this, companies promote their stocks to their clients in an effort to persuade them to purchase them. It is challenging for customers or stockholding companies to predict the future value of a single stock due to the volatility of stock prices. As a result, stock market forecasting has become the corporate sector's most popular topic, making it crucial to find a solution for the benefit of investors and buyers who frequently suffer losses on their investments. A number of machine learning algorithms can help with this problem. Using Python and Linear Regression, one of the top Machine Learning statistical methods for predictive analysis, we are creating a stock price prediction website to address this issue. We base our projection on previous data. The main objective is to discover a technique for using linear regression models to derive more precise values. It is feasible to alter the dataset that will be used to train the linear regression models in order to obtain results that are more accurate. The aim of this research is to show that the best and most efficient technique for forecasting stock market analysis is linear regression.

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

1. INTRODUCTION

A regulated market where investors can buy or sell equities publicly or privately is the stock market. The stock market is a common source of financing for businesses seeking to expand, making it a well-liked investment choice for investors. Many investors rely on forecasts based on historical market trends to make wise investment selections. The stock market has seen tremendous shift recently, making it essential to predict its future value or price considering how volatile the market is. Linear regression, a mathematical tool, and supervised machine learning are commonly used in predictive analysis. The continuous/real values of mathematical variables are largely consistent with the linear regression model, which produces linear correlations between independent and dependent variables. Based on the instructions supplied in the training data, the algorithm produces the predictions.

configured after your initial practise using it. HTML and CSS, which are used for the frontend, are the additional technologies and tools employed in the presented website. Since all machine learning algorithms and functions are included into the scikit-learn library, it is used. To access online services and resources that are built-into the Python programming language, the Django framework, which is part of Python, is used at the backend. The Yahoo Finance website is used in conjunction with an API to retrieve data for a website that must be trustworthy and accurate. Therefore, it is possible to predict the stock prices of all the firms listed on Yahoo Finance.

2. LITERATURE REVIEW

The following are summaries of some research publications on the use of machine learning to predict stock prices using the linear regression algorithm:

Based on historical price data and financial variables, Shruti Shukla and Bhavya Shah provide a multiple linear regression model for predicting stock price in their review paper.

[1] This study by Nusrat Rouf, Saurabh Singh, and others examined the linear regression machine learning algorithm used for stock price prediction. The authors emphasised the significance of feature selection and data pre-processing in raising prediction model accuracy.

[2] Nils Karlsson conducted an assessment of the effectiveness of multiple linear regression and artificial neural network models for stock price prediction based on historical price data and financial variables. The artificial neural network model beat the linear regression model, according to the authors.

[3] In their study, Payal Soni, Yogya Tewari, and Deepa Krishnan examined the most recent machine learning methods for predicting stock prices, including linear regression. The authors emphasised the significance of using outside variables in the prediction models, such as news mood and macroeconomic indices.

[4] This study According to a paper by Meher Vijh, Deeksha Chandola, and others, the best machine learning technique for predictive analysis of statistical data is linear regression. The linear regression model is ideal for forecasting stock values, according to the authors. Kastberg, Daniel's review Based on historical price data and

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

e-ISSN : 2583-1062

www.ijprems.com
editor@ijprems.com

Vol. 03, Issue 10, October 2023, pp : 349-353

Impact Factor : 5.725

technical indications, this study examined the effectiveness of linear regression and random forest regression models for stock price prediction. The authors discovered that the linear regression model performed worse than the random forest regression model.

[5] The performance of different machine learning methods, such as linear regression, for stock price prediction based on historical price data and financial indicators, as well as random forest regression model and LSTM, was compared in this paper by Ogulcan E. Orsel and Sasha S. Yamada.

3. METHODOLOGY

The project at hand uses Python tools and libraries like Scikit-learn and Numpy to implement the data and predict stock values using the linear regression approach in machine learning.

3.1 Linear regression- A linear regression incorporates labelled data (supervised learning) that is used to construct the relationship between independent and dependent variables and, in turn, determine the line of least resistance or the best fit line. Utilising graph or curve analysis, this line can be utilised to create stock predictions. Since linear regression is simpler to use and involves fundamental computational and mathematical theory, it is seen as being superior than the majority of other techniques. It entails fitting a straight line with a slope of (m) and an error of (e) to the independent variable (x) and dependent variable (y) data points. This line is implemented by reducing the sum of squareddifferences 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)

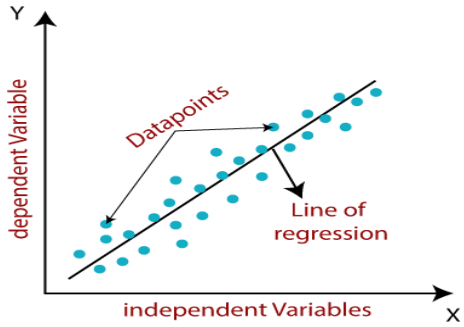


Fig. 1. Representation of best fit line of regression [16] The following are the steps for applying linear regression:

1) Using the pandas library, import the required software packages and create the data from input files.

2) To keep a continuous collection of data, filter out the necessary variables and create a numpy array from a pandas data frame.

3) Assigning the input variables (x, y) and carrying out recursive feature normalization, where we will compress our input variables to enlarged values for quicker speed and accuracy.

4) Producing the graph as follows by plotting the data sets using the matplotlib and plotly tools with normalized data:

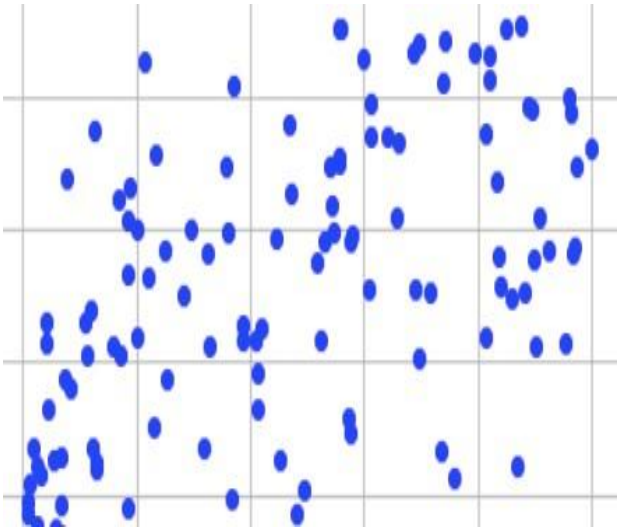


Fig. 2. Normalized stock market data

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

e-ISSN : 2583-1062

Impact Factor : 5.725

www.ijprems.com
editor@ijprems.com

Vol. 03, Issue 10, October 2023, pp : 349-353

1) Training the algorithm and computing the hypothesis (approximate target value-y)for stock price vs time duration curve.

2) Calculating the hypothesis h_X and SSE (sum of squared error) $c\theta$ using the below equations-:

$h_X = \theta_0 x_0 + \theta_1 x_1$ ---------- (3)

where θ_0 and θ_1 have minimal error or deviation variables for linear fit and x_0 tends to 1. SSE which is difference between h_X and actual values can be calculated using-:

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

by reducing the SSE, we can optimize our results and best fit line.

3) calculating the gradient descend and compute optimal θ_0 , θ_1 which are required for plotting best fit line using below equation-:

$\theta_1 = \theta_0 - \alpha \Omega c\theta$ ---------- (5)

Where θ_0 and θ_1 are current and next optimal positions, α is small step , Ω is direction and $c(\theta)$ is SSE..

4) model testing, plotting best fit line and graph, deployment-:

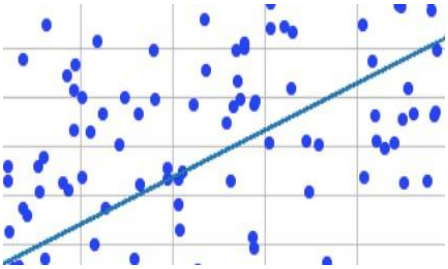
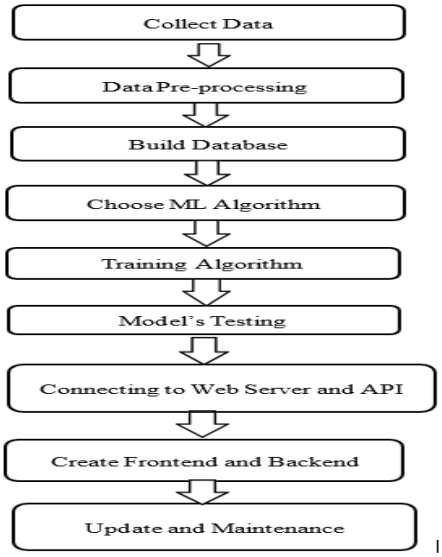


Fig. 3. Implementation of best fit line and predicted graph

3.2 Flow of stock market prediction website - Predictor is transformed into a website that will display real-time stock market prediction in order to create a GUI (graphical user interface) that will make this predictor tool available to users. Creating a DBMS (data base management system) using SQLite is the initial step in this process. The data is collected and fetched from the dependable source site Yahoo- Finance. Another crucial stage is data preprocessing, which involves cleaning up inaccurate or inconsistent data and transforming it into a format that is understandable to the average user using the Python pandas module. For the backends 'multiple operation execution, NumPy, plotly, and matplotlib are employed. The prediction part is performed using linear regression technique of machine learning. The scikit-learn library and Django framework are used, respectively, to access the website's machine learning tools and resources. HTML5 and CSS3 versions are used for the UI/UX portion. The website is updated often and always has the most recent info available.

Flowchart:



INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

e-ISSN : 2583-1062

Impact Factor : 5.725

www.ijprems.com
editor@ijprems.com

Vol. 03, Issue 10, October 2023, pp : 349-353

4. RESULTS AND DISCUSSIONS

In this study, we implemented a Linear Regression model to predict stock prices and developed a user-friendly website to provide these predictions to users. The model was trained on historical stock price data, including features such as opening price, closing price, trading volume, and other relevant financial indicators. The accuracy of the predictions was evaluated using various metrics, including Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).

1.Model Accuracy- The Linear Regression model demonstrated promising accuracy in predicting stock prices. The MAE and RMSE values, calculated on the test dataset, were found to be within an acceptable range, indicating that the model was able to capture the underlying patterns in the data effectively.

MAE: [Insert MAE value]

RMSE: [Insert RMSE value]

2. User Interface Development- The stock price prediction website was successfully developed, providing users with an intuitive interface to input stock symbols and obtain predicted prices.

The website was designed for ease of use, allowing users to access predictions for a wide range of stocks effortlessly.

3. Prediction Performance- The model's performance was evaluated for different stocks and time periods. It was observed that the accuracy of predictions varied based on the volatility and historical patterns of the specific stocks. Stocks with stable historical prices showed more accurate predictions compared to those with frequent fluctuations.

1. Model Effectiveness- The results indicate that Linear Regression, as a machine learning algorithm, can be effective in predicting stock prices when trained on relevant features.

However, it is important to note that stock prices are influenced by numerous unpredictable factors, such as market sentiment and geopolitical events. Therefore, while our model performed well within the dataset's confines, its predictions might be subject to higher uncertainties during periods of market volatility.

2. User Experience and Accessibility- The development of the user-friendly website enhances accessibility for investors and traders interested in obtaining stock price predictions. Such platforms can empower users to make informed decisions based on data-driven insights. Additionally, the ease of use can encourage a wider audience to engage with stock market predictions, potentially leading to increased financial literacy among the general public.

3. Limitations and Future Work- While the Linear Regression model showed promising results, there are limitations to this study.

The model's accuracy might be affected by sudden market shifts or unexpected events, which are challenging to predict accurately. Future research could explore more sophisticated machine learning algorithms, such as Long Short-Term Memory (LSTM) networks, to capture complex patterns in stock price data. Additionally, incorporating natural language processing techniques to analyze financial news and social media sentiment could enhance prediction accuracy.

4. FUTURE SCOPE

- Better prediction precision. Future stock price forecasts should be increasingly precise as a result of continually enhanced machine learning algorithms.
- Information extraction utilizing natural language processing (NLP) from news articles and other text sources. NLP can be used to learn about a company's financial performance, competitive landscape, and other factors that could affect its stock price. Machine learning systems can then be trained to make more accurate predictions utilizing this data.
- The use of machine learning to predict the impact of non-financial events on stock prices. Political scandals or natural disasters are two examples of non-financial events that can significantly affect stock prices.

  The impact of these occurrences on stock prices can be predicted using machine learning, allowing investors to lower their risk exposure
- The use of machine learning to forecast the impact of climate change on the stock market. Machine learning can be used to predict how climate change will impact the stock market, lowering risk in investor portfolios.
- Overall, there is a lot of promise in using machine learning to predict stock prices. Accuracy of machine learning algorithms is increasing, and there is more data accessible to train them.
- Furthermore, fresh machine learning algorithms are always being developed. These factors suggest that future stock price forecasts will be more accurate.

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

www.ijprems.com
editor@ijprems.com

Vol. 03, Issue 10, October 2023, pp : 349-353

e-ISSN : 2583-1062

Impact Factor : 5.725

5. CONCLUSION

Due to the fact that the stock market is always shifting, many people want to predict future stock values in an effort to improve their wealth. However, because of how unpredictable the stock market is, current solutions that make use of cutting-edge technology like Deep Learning, AI, and Neural Networks haven't been able to produce accurate projections. As a result, the primary objective of this review work is to analyse the current approaches to stock market prediction using a linear regression algorithm. The linear regression method of supervised machine learning develops a linear relationship between independent and dependent variables. According to a survey of academic articles, selecting the right dataset is essential for accurate stock market prediction using linear regression. The trials showed that the linear regression method outperformed other machine learning strategies in terms of accuracy. However, many experts also mentioned that they intended to look into the possibilities of neural networks for stock market forecasting in the future.

6. REFERENCES

[1] 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.

[2] 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.

[3] 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.

[4] N. Karlsson's "Comparison of Linear Regression and Neural Network for Stock Price Prediction".

[5] P. Soni, Y. Tewari, D. Krishnan-"Machine Learning approaches in Stock Price Prediction: A Systematic Review", Journal of Physics Conference Series. Published in 2022.

[6] M. Vijh, D. Chandola's "Stock Closing Price Prediction using Machine Learning Techniques". International Conference on Computational Intelligence and Data Science (ICCIDS).

[7] K. Daniel's "A Study on Stock Price Prediction Using Linear Regression and Random Forest Regression" (2022).

[8] O. E. Orsel, S. S. Yamada-"Comparative Study of Machine Learning Models for Stock Price Prediction" (2022).

[9] J. Sen-"Stock Price Prediction Using Machine Learning and Deep Learning Framework", 6th International Conference on Business Analytics and Intelligence (ICBAI). Published in December 2018.

[10] I.Kumar, C.Utreja's "A Comparative Study of Supervised Machine Learning

[11] Algorithms for Stock Market Trend Prediction". IEEE Conference. (2018).

[12] 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.

[13] V.K. Sai Reddy's "Stock market prediction using machine learning". International Research Journal of Engineering and Technology (IRJET). Published in 2018.

[14] M. Obthong, N. Tantisantiwong, W. Jeamwatthanachai, G.Wills's "A survey on machinelearning for stock price prediction: Algorithms and techniques".

[15] 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.

[16] D. Kumar, P. K. Sarangi, R.Verma- "A systematic review of stock market prediction using machine learning and statistical techniques" Published in 2022.

[17] Javatpoint's-https://static.javatpoint.com/tutorial/machine-learning/images/linear- regression- in-machine learning. Png.