

# A Linear Regression Model Used to Analysis the Tesla Stock Price Prediction Using Machine Learning

R. Santhosh Kumar<sup>1\*</sup>, Somorjit Laishram<sup>2</sup>, Sriram B.<sup>3</sup>, Pushpa R.<sup>4</sup>,  
Jeevitha K.<sup>5</sup>, Priyadarshani P.<sup>6</sup>

## Abstract

*The stock market is a fascinating sector of the economic research. It comes in a number of varieties. Several specialists have been examining and investigating the several patterns that the stock market experiences fluctuations. Predicting the stock values of different companies using historical data has been one of the primary research projects. Stock price prediction can help people a great deal by helping them understand where and how to invest, which will lower their risk of losing money. Companies can also use this program to determine how much to aim for and how many shares to distribute during their initial public offering. There have been a lot of noteworthy advancements in this sector thus far. Deep learning and machine learning are being investigated by numerous researchers as potential methods to predict stock values. Two methods are used by the suggested system to function: regression and classification. The method is used in classification to predict whether the closing stock price will increase or decrease the following day and in regression analysis to predict the closing stock price of a company.*

**Keywords:** Stock prices, stock market, machine learning, linear regression, trading classification

## INTRODUCTION

The financial organizations known as stock exchanges permit the trading of various commodities amongst stockbroker constituents. Estimating the future value of a financial instrument traded on a market, like a stock, is the process of stock market prediction. The idea that buying and selling stocks or shares on the market is similar to gambling is another widespread misconception. It is possible to change this misconception and increase awareness of it. 90% of the world's data has been created in recent years due to the daily development of 2.5 quintillion bytes of data [1].

In this study, we propose effective ways to employ user Tesla data for linear regression. The remaining sections of the document are structured as follows: The article comprises an overview of the completed pertinent studies. focuses on the survey of literature. There is discussion of the datasets, data cleaning, data preparation, and other techniques used in this work. There is discussion of the experiments and their results. What and how the work might be further expanded completes the article.

### \*Author for Correspondence

R. Santhosh Kumar  
E-mail: [santhoshkumarramesh1978@gmail.com](mailto:santhoshkumarramesh1978@gmail.com)

<sup>1-6</sup>PG student, Department of Artificial Intelligence and Data Science, Sri Manakula Vinayagar Engineering College, Madagadipet, Puducherry, India

Received Date: February 16, 2024  
Accepted Date: May 17, 2024  
Published Date: June 29, 2024

**Citation:** R. Santhosh Kumar, Somorjit Laishram, Sriram B., Pushpa R., Jeevitha K., Priyadarshani P. A Linear Regression Model Used to Analysis the Tesla Stock Price Prediction Using Machine Learning. Journal of Advanced Database Management & Systems. 2024; 11(2): 8–13p.

## LITERATURE SURVEY

The data produced by the financial industry is massive. It is very difficult for a trader to recognize a pattern and then determine the best course of action for decisions. Predicting the performance of the stock market is among the most challenging tasks. The forecast takes into account a plethora of variables, including physiological and physical aspects, rational and irrational conduct, etc. The

---

amalgamation of these factors renders share prices very unpredictable and challenging to forecast with a high degree of precision. When anticipating stock market values, machine learning can be a game-changer. Machine learning approaches have the capacity to reveal trends and revelations. These can be used to create forecasts that are incredibly accurate, something we haven't seen previously. In the modern world, machine learning is expanding at an astounding rate [1].

Currently, the forecasting of stock market trends is being examined by numerous study teams with the aid of social media analytics. To understand the feeling behind every good in news/tweet has been determined to be polarity. Both the semi-supervised algorithm and the dictionary-based approach can be used to determine the polarity of any article or tweet. Dictionary-based algorithms determine each word's polarity in the news by comparing it to a list of words that has already been predetermined [2].

One of the most cutting-edge areas of study nowadays is stock market forecasting. Investors may see significant gains as a result of precise price prediction algorithms. As a result, extensive research has been carried out in this area. A compilation of the research conducted in this field is outlined in this section. While the majority of academics have used technical analysis techniques, others have demonstrated that news and events pertaining to a firm have an effect on changes in that company's stock price. Two popular methods in the field of artificial intelligence are fuzzy neural networks and artificial neural networks [3].

Horton looks at 349 stocks' candlestick patterns and concludes that there isn't any benefit to using them. According to Marshall et al., candlestick patterns have no forecasting ability for the Japanese market and the Dow Jones Industrial Average. Several bullish and bearish candlestick patterns are examined for profitability by Tharavanij et al. They draw the conclusion, based on experiments, that no candlestick pattern, not even when filtered by reputable stochastic oscillators, can accurately forecast market directions [4].

To evaluate the worth of shares in the Taiwan Stock Exchange capitalization, a study was conducted employing an adaptive network-based fuzzy inference system (ANFIS) along with an empirical mode decomposition (EMD) empirical model. Balanced the EMD-ANFIS was shown to be more important than earlier study findings utilizing different approaches by the Hang Seng Stock Index and the Taiex Stock Index. Finally, research conducted utilizing architectural simulation of the Gradient Boosted Decision Tree (GBDT) model and the Logistic Regression model across two distinct stock indices—namely, the S&P 500 Composite Stock Price Index, the NASDAQ Composite Index, and the Shanghai Stock Exchange Composite Index—aimed to predict the market value of these indices. When it comes to stock index forecasts, LR2GBDT outperforms GBDT, SVM, NN, and TPOT models in terms of accuracy [5].

With the use of social media mining, a study by Yaojun and Yaoqing Wang clarified the impact of stock rating information from social media on the stock price. Taking this element into account, in addition to other significant factors, produced more accurate results. This study successfully used SVM with an emotion index [6].

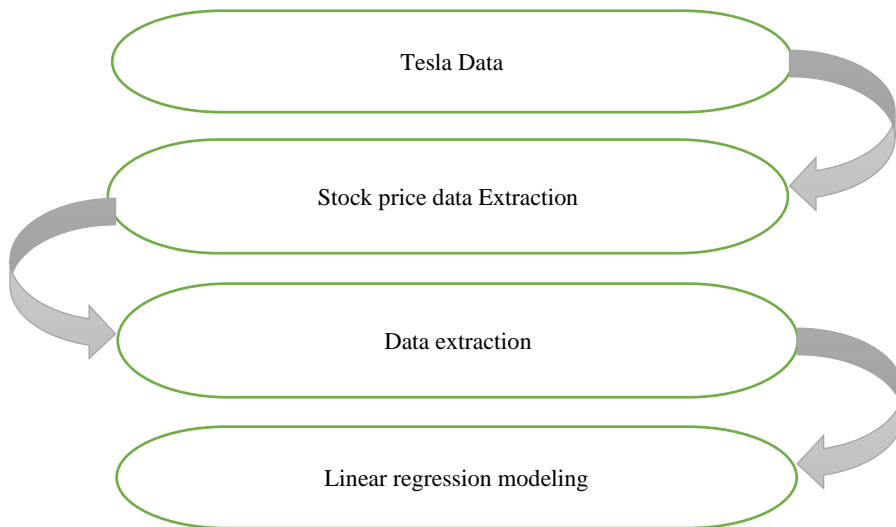
Ashish Sharma et al. developed a study that provided a theoretical approach to stock price prediction using basic regression models [7]. The difficulty is that the anticipated price of the stock has to be somewhat close to the actual price. That algorithm is therefore the most trustworthy algorithm. In their study, Yuiqing Dai and Yuning Zhang found that utilizing all 16 features in the SVM model predicted the greatest accuracy of 79% over the following 44 days [8].

However, they noted that they needed to validate the predictor's robustness using a variety of stock categories. Mustafa Abdul Salam, Osman Hegzay, and Osman S. Soliman [9]. Consequently, by using RNN-LSTM to anticipate the shares of the two firms, the research done by [10] that uses RNN-LSTM to predict future share prices provided satisfactory findings [11]. In related research, the trading price

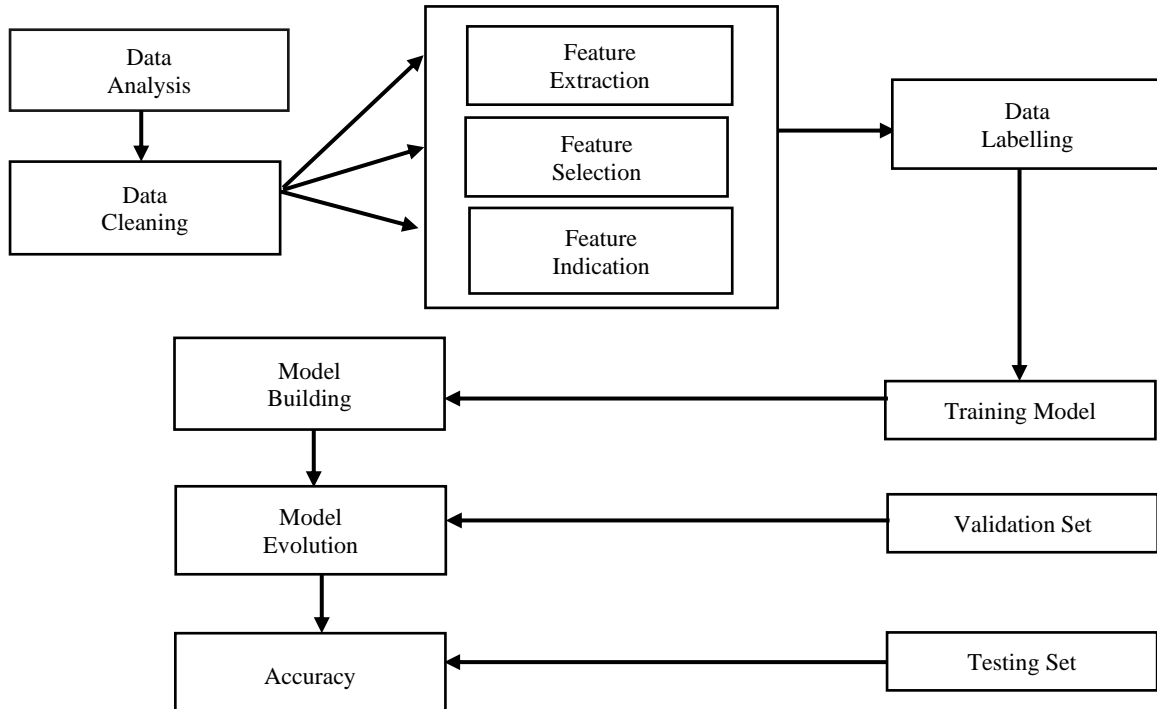
of shares for five businesses was predicted using Ann and Random Forest; Ann results showed a higher predictive value than random forest [12].

### METHODOLOGY

Visualizing the relationship between two variables and the regression line that best fits the data is necessary to create a diagram that illustrates linear regression. The following is a process for creating a linear regression diagram (Figures 1 and 2).



**Figure 1.** Overview of Telsa data diagram.



**Figure 2.** Data preprocessing block diagram.

### Data Collection

In the data collection, obtain historical information on the target security's stock price. Incorporate trading volumes, other market-related factors, and pertinent financial indicators. Make sure there is enough time range in the dataset to capture a variety of market circumstances (Figure 3).

### Data Preprocessing

This data preparation technique uses imputation or removal to manage missing data. To standardize numerical features, normalize them or bring them to a common scale.

Consider the incorporation of technical indicators or derived features based on domain knowledge (Figure 4).

### Train-Test Split

Divide the dataset into training and testing sets. Ensure a reasonable split ratio, for example, 80% for training and 20% for testing. Maintain chronological order to mimic real-world forecasting scenarios.

### Linear Regression Model

A variable's value can be predicted using linear regression analysis based on the value of another variable. Let's know what a linear regression equation is.

	Date	Open	High	Low	Close	Adj Close	Volume
count	2193	2193.000000	2193.000000	2193.000000	2193.000000	2193.000000	2.193000e+03
mean	2014-11-04 14:37:15.841313024	175.652882	178.710262	172.412075	175.648555	175.648555	5.077449e+06
min	2010-06-29 00:00:00	16.139999	16.629999	14.980000	15.800000	15.800000	1.185000e+05
25%	2012-08-29 00:00:00	33.110001	33.910000	32.459999	33.160000	33.160000	1.577800e+06
50%	2014-11-04 00:00:00	204.990005	208.160004	201.669998	204.990005	204.990005	4.171700e+06
75%	2017-01-09 00:00:00	262.000000	265.329987	256.209991	261.739990	261.739990	6.885600e+06
max	2019-03-15 00:00:00	386.690002	389.609985	379.350006	385.000000	385.000000	3.716390e+07
std	NaN	115.580903	117.370092	113.654794	115.580771	115.580771	4.545398e+06

Figure 3. Overview on data collection.



Figure 4. Overview on data preprocessing.

The formula for linear regression equation is given

$$y = \beta_0 + \beta_1 X$$

Y is the dependent variable

X is the independent variable  
 $\beta_0$  is the intercept  
 $\beta_1$  is the slope

In this linear regression, implement a linear regression model using a graph based on the on my project basis (Figure 5).

### Model Training

In this model, train the linear regression model using historical data on the training set. Assess the influence of various features on the model's performance. employing techniques like cross-validation for more robust training.

### RESULT

Trading professionals are very interested in predicting short-term stock price patterns since doing so could result in lucrative rewards. The selection of the machine learning algorithm relies on the attributes of the data. Coordinate backslide may be a common choice for straightforwardness and interpretability, though linear regression algorithm. Data preprocessing steps, tallying, taking care of misplaced values, scaling highlights, and encoding categorical components are essential for the execution of the appear. Portioning the dataset into planning and testing sets is critical to evaluating the model's execution on subtle data. Clear models like coordinate backslide are basic to decode but may not capture complex associations.

The Min-Max Scaler is a frequently employed feature scaling method in machine learning, utilized to scale and normalize dataset values within a predefined range, typically spanning from 0 to 1. The accuracy metric is commonly used in classification problems but may not be directly applicable to regression problems, such as stock price prediction. The following metrics are calculated using train and test data (Table 1).



**Figure 5.** Overview on linear regression.

**Table 1.** Overview of the result and discussion.

Metric	Train	Test
R2_score	0.8658871776828707	0.8610649253244576
MSE	1821.3833862936174	1780.9875394188446

### CONCLUSION

In summary, there are advantages and disadvantages associated with employing linear regression for predicting stock prices. A basic understanding of the link between independent factors and stock prices

can be obtained through the use of the straight-forward and easily interpretable linear regression model. It is very helpful for looking at linear patterns over time.

It is crucial to recognize the difficulties and presumptions that come with using linear regression models. The financial markets are intricate and subject to a multitude of factors that may not exhibit linear correlations. The stock market is a dynamic and ever-changing environment; therefore, the model's assumption of a consistent relationship between variables may not always hold true.

### Acknowledgments

We are very thankful and grateful to our beloved guide, Mr. Pragash (Assistant Professor), whose great support in valuable advice, suggestions, and tremendous help enabled us to complete our project. He has been a great source of inspiration to us.

### REFERENCES

1. Ravikumar S, Saraf P. Prediction of stock prices using machine learning (regression, classification) algorithms. In: 2020 International Conference for Emerging Technology (INCET); 2020 Jun 5-7; Belgaum, India: IEEE; 2020.
2. Sarode S, Tolani HG, Kak P, Lifna CS. Stock price prediction using machine learning techniques. In: 2019 International Conference on Intelligent Sustainable Systems (ICISS); 2019 Dec 20-21; Palladam, Tamil Nadu, India: IEEE; 2019.
3. Kaushal A, Chaudhary P. News and events aware stock price forecasting technique. In: 2017 International Conference on Big Data, IoT and Data Science (BID); 2017 Dec 20-22; Pune, India: IEEE; 2017.
4. Udagawa Y. Predicting stock price trend using candlestick chart blending technique. In: 2018 IEEE International Conference on Big Data (Big Data) 2018 Dec 10 (pp. 4162-4168). IEEE.
5. Moedjahedy JH, Rotikan R, Roshandi WF, Mambu JY. Stock price forecasting on telecommunication sector companies in Indonesia Stock Exchange using machine learning algorithms. In: 2nd International Conference on Cybernetics and Intelligent System (ICORIS); 2020 Oct 28-29; Makassar, Indonesia: IEEE; 2020.
6. Wang Y, Wang Y. Using social media mining technology to assist in price prediction of stock market. In: 2016 IEEE International conference on big data analysis (ICBDA) 2016 Mar 12 (pp. 1-4). IEEE.
7. Sharma A, Bhuriya D, Singh U. Survey of stock market prediction using machine learning approach. In: 2017 International Conference on Electronics, Communication and Aerospace Technology (ICECA); 2017 Apr 20-22; Coimbatore, India: IEEE; 2017.
8. Dai Y, Zhang Y. Machine learning in stock price trend forecasting. Stanford University, <http://cs229.stanford.edu/proj2013/DaiZhang-MachineLearningInStockPriceTrendForecasting.pdf>. Erişim Tarihi. 2013;2:2023.
9. Nishitha SN, Bano S, Reddy GG, Arja P, Niharika GL. Stock price prognosticator using machine learning techniques. In: 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA) 2020 Nov 5 (pp. 1-7). IEEE.
10. Moghar A, Hamiche M. Stock market prediction using LSTM recurrent neural network. *Procedia computer science*. 2020 Jan 1;170:1168-73.
11. Vijn M, Chandola D, Tikkiwal VA, Kumar A. Stock closing price prediction using machine learning techniques. *Procedia Comput Sci*. 2020; 167: 599–606. doi: 10.1016/j.procs.2020.03.326.
12. Ivaşcu CF. Option pricing using machine learning. *Expert Systems with Applications*. 2021 Jan 1;163:113799.