



# PRICE PREDICTION OF STOCK MARKET USING MACHINE LEARNING

Dhananjay Patil<sup>1</sup>, Shraddha Magdum<sup>2</sup>, Harshwardhan Patil<sup>3</sup>, Pranav Nerkar<sup>4</sup>, Parmeshwar Kokare<sup>5</sup>

Department of E&TC, SKNCOE, SPPU, Pune

<sup>1</sup>dhananjaypatil.skncoe.entc@gmail.com, <sup>2</sup>shraddha\_magdum\_skncoe@sinhgad.edu,

<sup>3</sup>harshvardhanpatil.skncoe.entc@gmail.com, <sup>4</sup>pranavnerkar.skncoe.entc@gmail.com,

<sup>5</sup>pskokare@sinhgad.edu

**Abstract:** The aim of this project is to utilize machine learning techniques to make predictions about the future performance of a particular stock or a portfolio of stocks. The use of machine learning techniques has shown promising results in predicting stock prices and has become a popular approach in finance and investment research. The main goal of this project is to leverage stock price prediction using machine learning algorithms to facilitate more informed and accurate investment decisions. Our proposed system for predicting stock prices integrates mathematical functions, machine learning algorithms, and external factors to improve prediction accuracy and ultimately generate profitable trades. Although accurately predicting the actual stock price is challenging, it can develop a model that forecasts whether the price will increase or decrease using machine learning techniques.

**Keywords:** Stock Price Prediction, Linear Regression, Long and Short-Term Memory Network (LSTM).

## I. INTRODUCTION

The financial market provides a platform for individuals to trade a variety of financial products such as currencies, stocks, equities, and derivatives through virtual platforms facilitated by brokers. Through the stock market, investors have the opportunity to purchase shares of publicly traded companies by trading on either exchange or over-the-counter markets. Investing in the stock

market provides individuals with the opportunity to potentially gain wealth and achieve financial prosperity with a relatively low risk compared to starting a new business or pursuing a high-paying career, even with small initial investments. Various factors can impact the stock market, leading to high volatility and uncertainty. While humans can execute orders in the market, automated trading systems (ATS) powered by computer programs have the potential to perform better and with greater speed in submitting orders than any human. In order to assess and manage the performance of ATSs, it is necessary to implement risk strategies and safety measures that are informed by human judgments. When building an ATS, numerous factors are taken into account, including the trading strategy to be used, complex mathematical functions that reflect the condition of a particular stock, machine learning algorithms that facilitate predicting the future stock value, and relevant news associated with the stock being examined.

## II. LITERATURE REVIEW

[1] Mingxi Zhang, Xuemin Li, Shuibo Yue, and Liu Qian Yang. (2020). "An Empirical Study of TextRank for Keyword Extraction."

To conduct data analysis in this paper, it was necessary to obtain historical market data and extract the relevant features. This was followed by dividing the data into testing and training sets, training the algorithm for price prediction, and finally visualizing the data. Although the proposed system achieved an accuracy of 96 LSTM units, it did not account for market sentiments.

[2] Beliga, S. (2014). "Keyword extraction: a review of methods and approaches."

The proposed solution in this paper is structured into three main parts. The first involves selecting highly effective features, while the second entails examining and performing dimensionality reduction on the data. The proposed system, which has numerous features, has demonstrated decent accuracy. It is possible to increase the accuracy further by utilizing a combination of features, retaining some while discarding others. It's worth noting that the RFE algorithm is not sensitive to term lengths other than 2-day, weekly, and biweekly.

[3] Carretero-Campos, C., Bernaola-Galván, P., Coronado, A. V., & Carpena, P. (2013). "Improving statistical keyword detection in short texts: Entropic and clustering approaches. Physical."

This paper proposes a model in the field of Artificial Intelligence, Machine Learning, and Deep Learning, titled "FUTURE STOCK PRICE PREDICTION USING RECURRENT NEURAL NETWORK, LSTM, AND MACHINE LEARNING." Compared to existing models, this stock price prediction model demonstrates higher accuracy. Notably, the project includes a Graphical User Interface (GUI) where users can upload Train and Test Data and obtain the model results and future predicted graphs of stock prices for the next 30 days.

[4] M. Litvak, M. Last, H. Aizenman, I. Gobits, A. Kandel, “DegExt – “A Language- Independent Graph-Based keyphrase Extractor”.

The proposed model in this paper involves deploying a blending ensemble learning model that merges LSTM and GRU to tackle the challenging task at hand. The primary difference between LSTM and GRU is how memory content is exposed within each unit and how new information is processed. The LSTM unit controls the amount of memory content seen through the output gate, which determines the information that will be utilized in the next LSTM unit. This approach yields high accuracy and efficiency. However, existing approaches have failed to account for market sentiments or sudden news, which may impact the accuracy of the predictions.

[5] Ventura, J., & Silva, J. (2012). “Mining concepts from texts. Procedia Computer Science, An extended keyword extraction method. Physics Procedia.” 24, 1120-1127.

This paper suggests that data-mining techniques such as ANN and SVM are more effective at detecting stock-price manipulation compared to multivariate statistical techniques like discriminate analysis or LR. This is because data-mining techniques exhibit superior classification accuracy. The authors propose a new binary classification method for predicting corporate failure, which uses genetic algorithms, and they aim to validate its efficacy.

[6] Krishna, V. ScienceDirect . “NSE Stock Stock Market Market Prediction Prediction Using Using Deep-Learning.” (2021)

This paper explains that machine learning (ML) can be classified into two types: supervised learning and unsupervised learning. In supervised learning, a labeled dataset is used to train the model, and the resulting data is available for subsequent use. On the other hand, unsupervised learning uses an unlabeled dataset as input. The objective of supervised learning is to train the model to automatically associate and match new input data to a given output data when it is input into the model.

[7] Chandar, S.K. “Fusion model of wavelet transform and adaptive neuro fuzzy inference system for stock market prediction.” (2019)

The aim of this paper is to develop an effective soft computing technique for predicting stock prices. The authors proposed a time-series prediction model for the closing price, which combines the wavelet-adaptive network-based fuzzy inference system (WANFIS) with a fusion forecasting model capable of predicting stock market trends. The study used data obtained from internet sources and employed the discrete wavelet transform (DWT) to decompose financial time series data. The approximation obtained from the DWT is used as the input to the WANFIS model for prediction.

[8] Strader, T.J.; Rozycki, J.J.; Root, T.H.; Huang, Y.H.J. “Machine Learning Stock Market Prediction Studies: Review and Research Directions.” J. Int. Technol. Inf. Manag.(2020)

A group of studies in this paper primarily focuses on stock market prediction using artificial neural networks (ANNs), which are computational models based on biological neural networks.

ANNs consist of layers of nodes that transmit signals through connected nodes as they learn based on examples and attempt to minimize prediction error. These studies use ANNs to predict stock market performance.

[9] Zhang, C.; Cheng, J.; Tang, X.; Sheng, V.S.; Dong, Z.; Li, J. Novel. "DDoS Feature Representation Model Combining Deep Belief Network and Canonical Correlation Analysis." *Comput. Mater. Contin.* (2019)

This paper explains two methods for extracting data features using the deep belief network. The first method involves unsupervised training of the deep belief network, which is then used to initialize a feed forward neural network. In this method, all hidden layers of the feed forward neural network are used as feature extractors. The second method involves a layer-by-layer greedy approach used by the deep belief network.

[10] Ganser, A.; Hollaus, B.; Stabinger, S. "Classification of Tennis Shots with a Neural Network Approach." 2021

This article highlights the significance of data analysis in sports and explains the development of a novel approach in tennis for combining data generation with a wristband and classification using a deep convolutional neural network. The approach involves the creation of a shot detection trigger and a deep neural network capable of classifying tennis shots into three and five shot types using a dataset for training.

### III. IMPLEMENTATION

This model explains that regression analysis is a statistical technique used to model the relationship between a dependent variable and one or more independent variables. The Project also highlights that there are two categories of prediction methods: statistical and artificial intelligence methods. Examples of statistical methods are the logistic regression model and ARCH model, while artificial intelligence methods include multi-layer perceptron, convolutional neural network, naive Bayes network, and others. The study in question uses a Long Short-Term Memory (LSTM) network for prediction. The model offers two options for prediction: regression and LSTM (Long-Short Term Memory). LSTM (Long-Short Term Memory) is the preferred method for prediction because it has unique advantages over regression techniques.

### IV. METHODOLOGY

#### Linear Regression :

Linear regression is one of the easiest and most popular Machine Learning algorithms. It is a statistical method that is used for predictive analysis. Linear regression makes predictions for continuous/real or numeric variables such as sales, salary, age, product price, etc. Linear regression

algorithm shows a linear relationship between a dependent (y) and one or more independent (x) variables, hence called as linear regression. Since linear regression shows the linear relationship, which means it finds how the value of the dependent variable is changing according to the value of the independent variable. The linear regression model provides a sloped straight line representing the relationship between the variable.

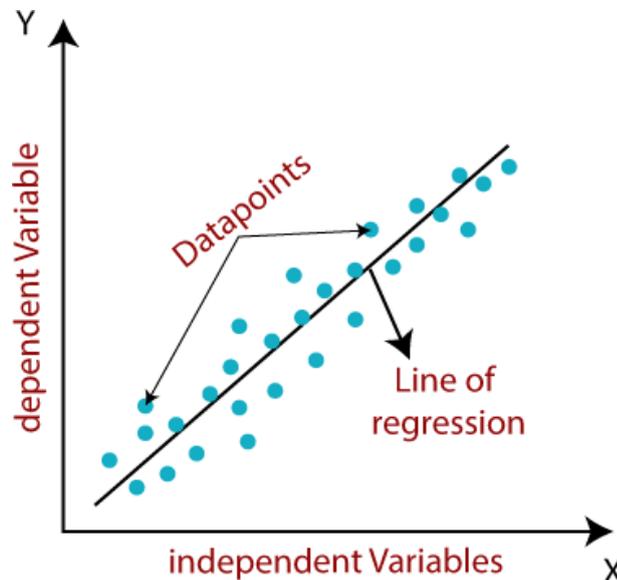


Figure 1 : Linear Regression

### LSTM (Long Short Term Memory) :

Long Short Term Memory is kind of recurrent neural network. In RNN output from the last step is fed as input in the current step. LSTM was designed by Hochreiter & Schmidhuber. It tackled the problem of long-term dependencies of RNN in which the RNN cannot predict the word stored in the long-term memory but can give more accurate predictions from the recent information. As the gap length increases RNN does not give an efficient performance. LSTM can by default retain the information for a long period of time. It is used for processing, predicting, and classifying on the basis of time-series data.

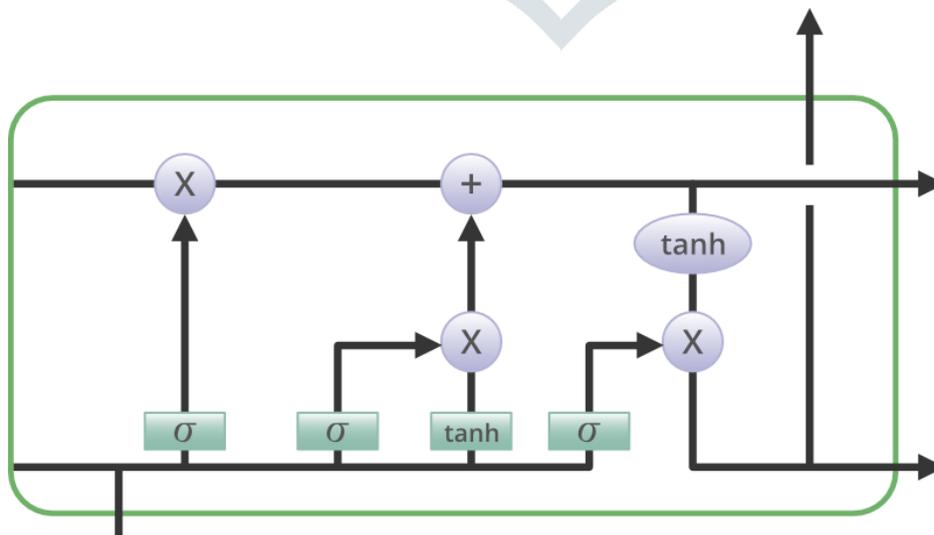


Figure 2 : LSTM

**Block Diagram :**

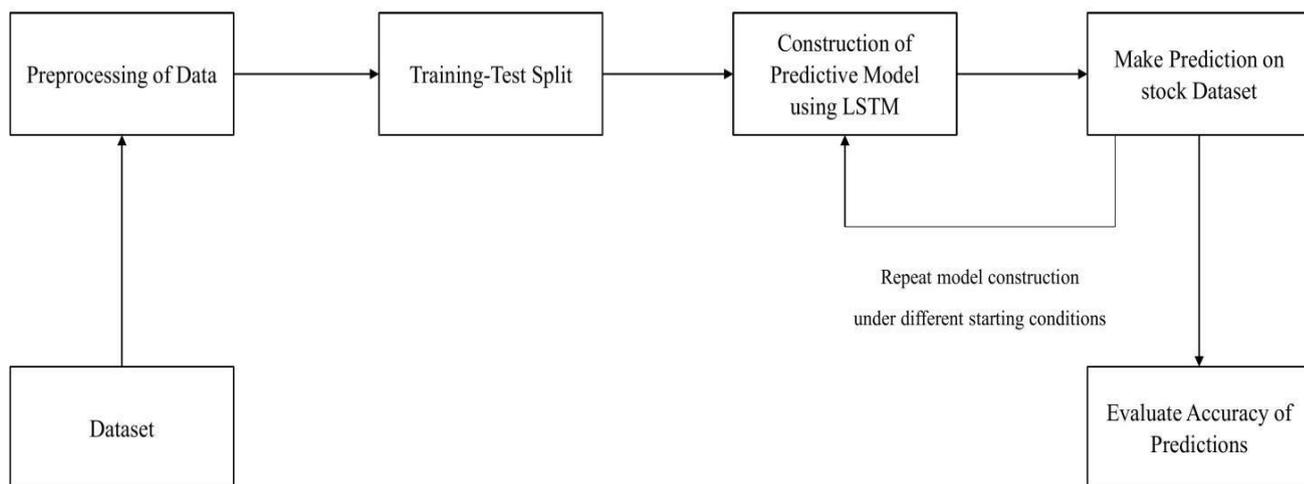


Figure 3 : Block Diagram

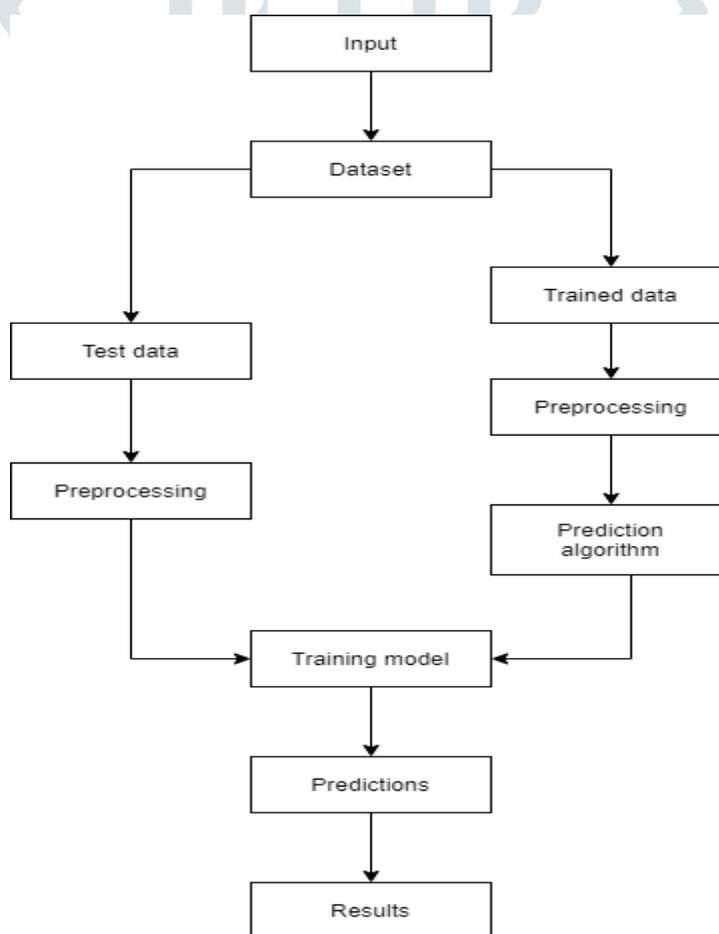


Figure 4 : Flow Chart

The flowchart provides an overview of the model's functionality. The model incorporates a dataset of CSV files containing data on multiple stocks and uses either regression or LSTM algorithms. The resulting output is presented in the form of charts and graphs.

## V. RESULT

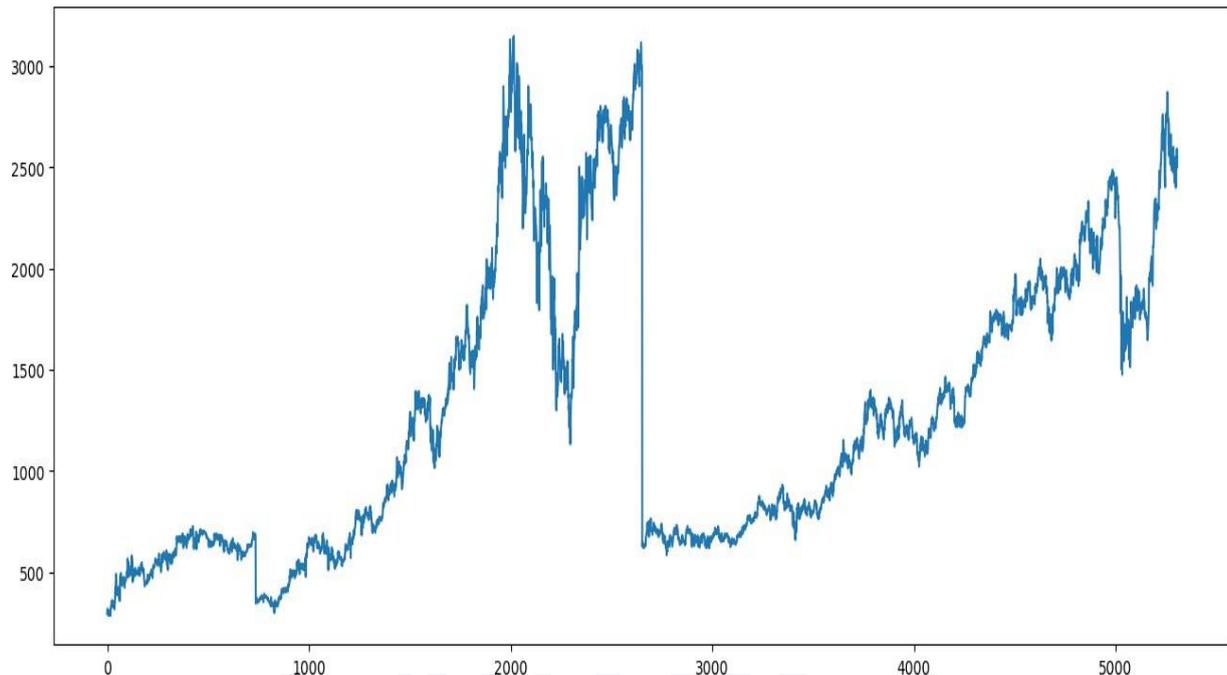


Figure 4 : Result-1

In this result we have shown the closing price of the stock in future. In the x-axis the price is given and at the y-axis the dates are shown. Closing prices of the stock are important to prediction the future of the stock.

### Visualization of Actual Price and Predicted Price

```
In [36]: graph= dfr.head(30)
```

```
In [37]: graph.plot(kind='bar')
```

```
Out[37]: <AxesSubplot:>
```

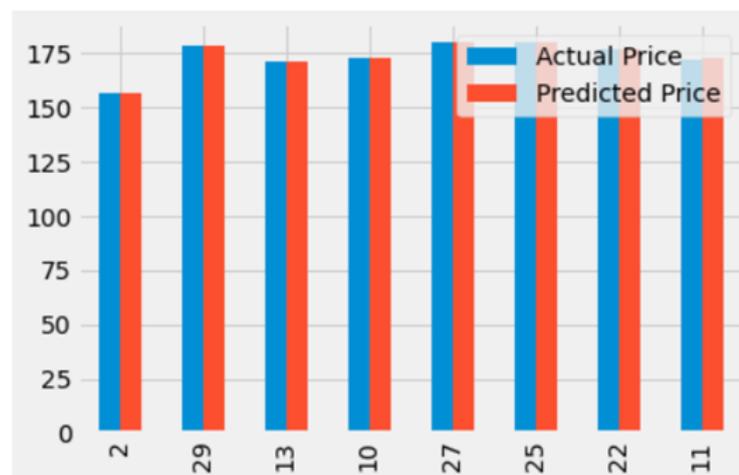


Figure 4 : Result-2

This fig shows the candlestick pattern of the actual price and the predicted price Blue candles show the actual price The orange candle show the predicted price. So we can get idea of the stock price movements on next day.

## VI. CONCLUSION

By enabling smarter solutions, we can create a more secure and brighter future. Our Stock Market Price Prediction project employs Machine Learning and LSTM models to analyse past stock prices, process the data, and forecast future prices, reducing losses and maximizing profits. This approach provides a path for stock market trading and helps mitigate the risk of significant losses. Although there is no guarantee of profit, utilizing this approach can help minimize the losses incurred in the stock market.

## VII. REFERENCES

- [1] Mingxi Zhang, Xuemin Li, Shuibo Yue, and Liu Qian Yang. (2020). An Empirical Study of TextRank for Keyword Extraction.
- [2] Beliga, S. (2014). Keyword extraction: a review of methods and approaches. University of Rijeka, Department of Informatics, Rijeka.
- [3] Carretero-Campos, C., Bernaola-Galván, P., Coronado, A. V., & Carpena, P. (2013). Improving statistical keyword detection in short texts: Entropic and clustering approaches. *Physica A: Statistical Mechanics and its Applications*, 392(6), 1481-1492.
- [4] M. Litvak, M. Last, H. Aizenman, I. Gobits, A. Kandel, "DegExt – A Language- Independent Graph-Based keyphrase Extractor" in Proc. of the 7th AWIC 2011, pp. 121-130, Switzerland, 2011. M. Litvak, M. Last, H. Aizenman, I. Gobits, A. Kandel, "DegExt – A Language-Independent Graph-Based keyphrase Extractor" in Proc. of the 7th AWIC 2011, pp. 121-130, Switzerland, 2011.
- [5] Ventura, J., & Silva, J. (2012). Mining concepts from texts. *Procedia Computer Science*, 9, 27-36.
- [6] Hong, B., & Zhen, D. (2012). An extended keyword extraction method. *Physics Procedia*, 24, 1120-1127.
- [6] Krishna, V. ScienceDirect ScienceDirect NSE Stock Stock Market Market Prediction Prediction Using Using Deep-Learning (2021)
- [7] Chandar, S.K. Fusion model of wavelet transform and adaptive neuro fuzzy inference system for stock market prediction. (2019)
- [8] Strader, T.J.; Rozycki, J.J.; Root, T.H.; Huang, Y.H.J. Machine Learning Stock Market Prediction Studies: Review and Research Directions. *J. Int. Technol. Inf. Manag.*(2020)
- [9] Zhang, C.; Cheng, J.; Tang, X.; Sheng, V.S.; Dong, Z.; Li, J. Novel DDoS Feature Representation Model Combining Deep Belief Network and Canonical Correlation Analysis. *Comput. Mater. Contin.* (2019)
- [10] Ganser, A.; Hollaus, B.; Stabinger, S. Classification of Tennis Shots with a Neural Network Approach. 2021