



# IJRASET

International Journal For Research in  
Applied Science and Engineering Technology



---

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

---

**Volume: 9      Issue: VII      Month of publication: July 2021**

**DOI: <https://doi.org/10.22214/ijraset.2021.37183>**

**[www.ijraset.com](http://www.ijraset.com)**

**Call:  08813907089**

**E-mail ID: [ijraset@gmail.com](mailto:ijraset@gmail.com)**

# Stock Prediction using Long Short-Term Memory, Support Vector Regression and Linear Regression Algorithms

Mr. V. Manoj Kumar<sup>1</sup>, A. Mohamed Shirajudeen<sup>2</sup>, B. Mukilan<sup>3</sup>, G. Sabarinathan<sup>4</sup>, K. Sameer Ahamed<sup>5</sup>

<sup>1</sup>Mentor: M.E (Assistant Professor), Department of Information Technology, Saranathan College of Engineering, Trichy-620012

<sup>2, 3, 4, 5</sup>Department of Information Technology, Saranathan College of Engineering, Trichy-620012

**Abstract:** Prediction is most important for stock market not only for traders but also for computer engineers who analyses stock data. We can perform this prediction by two ways one is using historical stock data and other by analyzing by information gathered from social media. It is based on model/pattern used to predict stock dataset, there are so many models are available for predicting stocks, simply model is algorithm that's from machine learning and deep learning. In the data set the two main parameters open and close value are used for stock prediction mostly but we can also predict by its volume too. So that data is preprocessed before it is used for prediction. In this paper we used various algorithm like linear regression, support vector regression and long short-term memory for better accuracy and to compare how it different from other algorithm and for predicting future stock.

**Keywords:** Stock Prediction, Data Analysis, Linear Regression, Support Vector Regression, Long Short-Term Memory, Machine Learning.

## I. INTRODUCTION

The economy growth is most important for any country all over the world. The economy growth is depending on stock market of the country, In our India is majorly depend on two national exchanges that is NSE and BSE. This indicator very much used for Indian economy growth. We are going to predict more than 15 companies in NSE and see the future stock values of these company by machine and deep learning algorithms. Before that we have to learn about data science and how we create model based on these algorithms and then how it predict. Data science is deals with data, that is used for extracting all kind of data using different methods and algorithms. Data science is working for automating transportation such as creating a self-driving car etc. it will help in several ways like survey for various thing, confirmation of tickets like flights, etc. there are two kind of data: one is structured and another is unstructured

Structure data is the normal student's database. Unstructured data is nothing but your "social media data like google and Facebook". Structure and an unstructured data need to manipulate for implementing your machine learning. Machine learning and Deep Learning is important part of data science. That provide model to run a machine that can act as a human brain. We use various algorithms to create model to predict stock values in data science .

The motive of proposed system is to predict the stocks minimal time with optimize solution. We going to include what does the right time invest on stocks. Forecast the stocks for next day. we use long short-term memory deep learning algorithm while building a model. Long short-term memory is a model used in the deep learning.it is artificial recurrent neural network and it has feedback connections. Its not like feed forward neural network It process a continuous or sequences of entire data not only a single data point. Linear regression is an algorithm which the output is continuous, and it is supervised machine learning algorithm.it has a constant slope. It defines relationship between two data, if x is increase and y also increase. Whereas support vector regression is a regression analysis, it is build based on the concept of support vector machine.

The objective of the project is implemented using historical data and building a model using machine learning and deep learning algorithms. It will be helpful to predict the future stock outcomes regarding a particular company stock. So that investor must invest and get profit based on prediction and avoiding any kind of risk. Different models such as linear regression, long short-term memory and support vector regression are used for giving better result.

## II. LITERATURE SURVEY

[8] had used different algorithms such as polynomial model, radial basis function(RBF) neural network and multilayer perceptron neural network, linear regression method.

They found feed forward neural network by comparing all these models and get accurate predict for open price value of the stock. They discuss the strategies of Buy low, Sell high. They evaluate and calculate errors like ME, RMSE, MAPE. Then obtain average opening price of the stock

[2] are combine to predict stock market by principle component analysis. In this paper, they predict market trends with help of a principal component analysis (PCA) with linear regression is investigated for high dimensionality problem of stock exchange .it help to improve the performance for prediction using machine learning methods it is reducing the data redundancy. This was carried out on a three stock exchanges such as: New York Stock Exchange, Karachi stock exchange and London Stock Exchange. The linear regression model accuracy is compared with PCA before and after applied. It shows that in general, PCA is improve the performance of machine learning if and only if relative correlation among input features is investigated and careful selection is done while choosing principal components. Root mean square error (RMSE) is used as an evaluation metric to evaluate the classification mode.

[3], they predict the stock market by support vector machine (SVM). In this paper, they present a framework both empirical and theoretical to apply the Support Vector Machines algorithm to predict the stock market closing price. First, they select four specific company and six macroeconomic factors that may influence the stock trend for stock multivariate analysis. Second, Support Vector Machine is used for analyzing the relationship of those factors and predicting the performance of stock. The results shows that Support vector machine is a powerful predictive tool for predicting stock in the financial or stock market. They use federal reserve bank of St. louis' dataset for more companies.

[5] develop an survey on stockmarket prediction based on machinelearning algorithm. They used regression model for stock prediction, the regression models like polynomial regression,radial basis function regression, sigmoid regression and linear regression. Linear regression is most suited for the prediction purpose. They survey of well-known efficient regression approach to predict the stock market price from stock market data based. In future, they improve multiple regression approach using a greater number of variables.

[7], they done stock market prediction using Hybrid Approach. In this paper, the main objective is to construct a model using the opinion mining and clustering method to predict stock value moment and to predict National Stock Exchange (NSE). the Proposed methodology of this paper will give us two output set i.e., one is sentiment analysis, and another is clustering based prediction with help of some specialized parameters of stock exchange. they are examining both the results produces an efficient prediction. In this paper, they consider stocks with maximum capitalization within all the important sectors for empirical analysis.

[6] develop neural network for predicting stock market. In the proposed work, they presented an Artificial Neural Network approach to predict the stock market indices. It shows fairly accurate results unless there is variation in the actual data. They also applied forward feed and backward propagation for better results. They outlined the design of the Neural Network model with its salient features and customizable parameters.

They implement number of the activation functions along with the options for the cross-validation sets. they finally test our algorithm on the Nifty stock index dataset where they predict the values based on values from the past days. they achieve a best-case accuracy of 96% on the dataset.

[9], they predict stock market by big data approach rather than machine learning or deep learning. They introduce the concepts of financial derivatives such as "no arbitrage" principles and they use prediction models like emh and random walk theory. Due to type of data available they used emh.theyanalyse and predict based on news articles. In this prediction, there are several steps are involved such as data preparation, data analysis, aggregation, and visualisation. The data were collected on the basis and prediction were made on logistic regression. The output will be either positive,negative or neutral.

[4] they analyze comparative between linear regression and support vector regression. They compared these two model for data prediction based on past data or historical data. They use two function for regression technique such as LeastMedsq and SMOReg function.

The LeastMedSq function is used for predicting values more time than SMOReg . they also used two metrics for comparison such as MAE and RMSE.

### III. CATEGORIZATION OF MACHINE LEARNING AND DEEP LEARNING ALGORITHMS APPLIED FOR STOCK PREDICTION

TABLE 1: Analysis of machine and deep learning Algorithms

Algorithm used	Model Applied	Type
LR-Linear regression	Regression	Machine learning
SVR-Support vector regression	Regression	Machine learning
LSTM-Long shortTerm memory	Classification Or Regression	Deep learning

### IV. ARCHITECTURE DIAGRAM

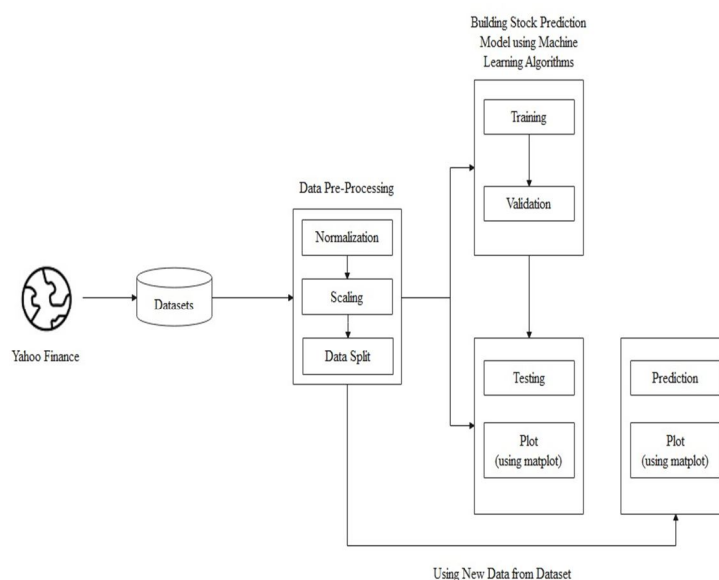


Figure IV-Architecture Diagram

### V. PSEUDOCODE

#### A. Support Vector Regression (SVR)

```

close_px = df['Adj Close']
mavg = close_px.rolling(window=100).mean()
dates_df = df.copy()
dates_df = dates_df.reset_index()
print(dates_df)
org_dates = dates_df['Date']
dates_df['Date'] = dates_df['Date'].map(mdates.date2num)
dates_df.tail()
dates = dates_df['Date'].to_numpy()
prices = df['Adj Close'].to_numpy()
dates = np.reshape(dates, (len(dates), 1))
prices = np.reshape(prices, (len(prices), 1))
    
```

```
svr_rbf = SVR(kernel='rbf', C= 1e3, gamma= 0.1)
svr_rbf.fit(dates, prices)
plt.figure(figsize = (12,6))
plt.plot(dates, prices, color= 'black', label= 'Data')
plt.plot(org_dates, svr_rbf.predict(dates), color= 'red', label= 'RBF model')
plt.xlabel('Date')
plt.ylabel('Price')
plt.legend()
Long-Short Term Memory (LSTM):
    Train_dates = pd.to_datetime(df['Date'])
    cols = list(df)[1:6]
    df_for_training = df[cols].astype(float)
    scaler = StandardScaler()
    scaler = scaler.fit(df_for_training)
    df_for_training_scaled = scaler.transform(df_for_training)
    trainX = []
    trainY = []
    n_future = 1
    n_past =14
    for i in range(n_past, len(df_for_training_scaled) - n_future+ 1):
        trainX.append(df_for_training_scaled[i-n_past:i, 0:df_for_training.shape[1]])
        trainY.append(df_for_training_scaled[i + n_future -1 :i +n_future, 0])
    trainX, trainY = np.array(trainX), np.array(trainY)
    model = Sequential()
model.add(LSTM(64,activation='relu',input_shape=(trainX.shape[1], trainX.shape[2]), return_sequences=True))
y_pred_future=scaler.inverse_transform(forecast_copies)[:,0]
forecast_dates =[]
    for time_i in forecast_period_dates:
        forecast_dates.append(time_i.date())
df_forecast=pd.DataFrame({'Date':np.array(forecast_dates),'Close':y_pred_future})
df_forecast['Date']=pd.to_datetime(df_forecast['Date'])
original = df[['Date','Close']]
original['Date']=pd.to_datetime(original['Date'])
original = original.loc[original['Date']>='2020-5-1']
sns.lineplot(original['Date'], original['Close'])
sns.lineplot(df_forecast['Date'],df_forecast['Close'])
```

### B. Linear Regression (LR)

```
close_col = df['Adj Close']
mvag = close_col.rolling(window=100).mean()
rd = close_col / close_col.shift(1) - 1
predict_days = 7
df['Prediction'] = df['Adj Close'].shift(-predict_days)
X = np.array(df.drop(['Prediction'], axis = 1))
X = X[:-predict_days]
y = np.array(df['Prediction'])
y = y[:-predict_days]
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2) #Splitting the data into 80% for training & 20% for
testing
```

```
linear_model = LinearRegression()
linear_model.fit(X_train, y_train)
linear_model_score= linear_model.score(X_test, y_test)
print('Linear Model score:', linear_model_score)
X_predict=np.array(df.drop(['Prediction'],1))[- predict_days:]
linear_model_predict_prediction=linear_model.predict(X_predict)
linear_model_real_prediction=linear_model.predict(np.array(df.drop(['Prediction'], 1)))
predicted_dates = []
recent_date = df.index.max()
display_at = 1000
alpha = 0.5
for i in range(predict_days):
    recent_date += str(timedelta(days=1))
    predicted_dates.append(recent_date)
plt.figure(figsize=(16, 9))
plt.plot(df.index[display_at:], linear_model_real_prediction[display_at:],label='Linear Prediction', color='blue',
alpha=alpha)
plt.plot(predicted_dates, linear_model_predict_prediction, label='Forecast', color='green', alpha=alpha)
plt.plot(df.index[display_at:], df['Close'][display_at:], label='Actual', color='red')
plt.legend()
```

## VI. LIMITATION OF EXISTING APPROACHES

In the Stock Price Prediction Using Data Analytics by [8], the limitations are they are using Feed Forward Neural network various efficient methods are available other than this, They used linear regression algorithm if we increase data it fails to fit complex datasets properly. Accuracy level is less in this project for predicting the stock price. Outliers of a data set are anomalies or extreme values that deviate from the other data points of the Distribution. Data outliers can damage the performance.

In the Prediction of Stock Market by Principle Component Analysis by [2], the limitations in this project are they using PCA (principle component analysis). principle component are not as readable and interpretable as original features. RME (root mean square error) is prone to outlier as it uses the same concept mean in computing each error value. In linear regression algorithm have underfitting problem when a situation that arises when a machine learning model fails to capture the data properly.

In the Stocks Market Prediction Using Support Vector Machine by [3], they conclude the large dataset is not suitable for support vector machine algorithm. It will overlap and does not perform very well when the data set has more. it will underperform when each data point exceeds the number of training data samples. it works by putting data points, above and below the classifying hyperplane, the classification has no probabilistic explanation.

The Survey is created for stock market prediction Using Machine Learning Approach by A. Sharma, [5], the limitations are It is easily affected by outlier's regression solution will be likely dense (because no regularization is applied), The learning curve is steep.

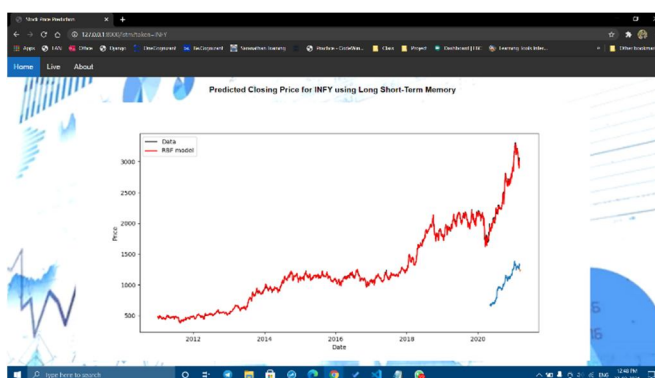
In the Neural Networks through Stock Market Data Prediction by [6], the limitations are Needs enormous amount of data mainly for architecture, Long training times for deep networks, Architecture has to tuned out to achieve best performance. There are design decisions that must be made, from the no of layers and no of nodes in each layer to activation functions, and an architecture that works well to some one problem very often does not generalized.

In the Stock Market Prediction Using Hybrid Approach by [7], Sentimental analysis is not efficient for analyzing large amounts of data without error. The no of clusters often unknown in different datasets, No particular data point relevant, Sentimental analysis limitations dependent on the restraints you place on degree the input will be modified.

In the comparative analysis on linear regression and support vector regression by [4], they compare these two algorithm based on RMSE, but it doesn't focus on time taken to build model. Linear regression takes more time to build model than support vector regression and it lack in high efficiency. It gets better result with less error rate.

In the stock market prediction using big data approach by [9], it lacks on large amount of testing data. It perform with small dataset that produce low accuracy than training dataset.

## VII. SCREENSHOTS



## VIII. CONCLUSION

The main aim of our project is to predict stock market values by three different algorithms such as linear regression, support vector regression and long short term memory to show better prediction and its accuracy to investor, by that investor has taken right decision to invest in that particular company or go for another company. The input dataset is taken from the yahoo finance and the final output will be displayed in website by Django.

## IX. ACKNOWLEDGMENT

We are the students of Saranathan college of engineering, would like to present our gratitude to our mentor Mr.V.Manojkumar.M.E, without whom this research paper could not have been possible. Her guidance has been of utmost importance in the completion of this research paper.



### REFERENCES

- [1] D. Bhuriya, G. Kaushal, A. Sharma and U. Singh, "Stock Market Predication Using A Linear Regression," in International Conference on Electronics, Communication and Aerospace Technology, 2017.
- [2] P. Guo, M. Waqar, H. Dawood, M. B. Shahnawaz and M. A. Ghazanfar, "Prediction of in International Conference on Electronics, Communication and Aerospace Technology, 2017. Stock Market by Principle Component Analysis," in 13th International Conference on Computational Intelligence and Security, 2017. International.
- [3] Z. Hu, J. Zhu and K. Tse, "Stocks Market Prediction Using Support Vector Machine," in 6th International Conference on Information Management, Innovation Management and Industrial Engineering, 2013.
- [4] S. Kavitha, S. Varuna and R. Ramya, "A Comparative Analysis on Linear Regression and Support Vector Regression," in Online International Conference on Green Engineering and Technologies (IC-GET), 2016.
- [5] A. Sharma, D. Bhuriya and U. Singh, "Survey of Stock Market Prediction Using Machine Learning Approach," in International Conference on Electronics, Communication and Aerospace Technology, 2017.
- [6] R. Verma, P. Choure and U. Singh, "Neural Networks through Stock Market Data Prediction,
- [7] V. Rajput and S. Bobde, "Stock Market Prediction Using Hybrid Approach," in International Conference on Computing, Communication and Automation (ICCCA2016), 2016
- [8] S. Tiwari, A. Bharadwaj and S. Gupta, "Stock Price Prediction Using Data Analytics," in International Conference on Advances in Computing, Communication and Control (ICAC3), Mumbai, India, 2017
- [9] G. V. Attigeri, M. P. M, R. M. Pai and A. Nayak, "Stock Market Prediction: A Big Data Approach," in TENCON 2015 - 2015 IEEE Region 10 Conference, Macao, China, 2015.





10.22214/IJRASET



45.98



IMPACT FACTOR:  
7.129



IMPACT FACTOR:  
7.429



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24\*7 Support on Whatsapp)