**IOSR Journal of Engineering (IOSRJEN)** **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
PP 09-16

# Price Trend Prediction Using Data Mining Algorithm

Mrs. PranjaliKasture
*Assistant Professor, Department of Information Technology,*
*Thakur college of Engineering and Technology Mumbai, India*

***Abstract:*** *Data mining can be applied on past and present financial data to generate patterns and decision making algorithm. The system will analysis the historical stock data's of some companies. Based on different factors which affect the stock values of company like demand and supply of shares, popularity of company, profit earned. The proposed system will predict future price, buy and sell possibility of the share by representing it in graphical manner. Here user can select the company in which user is interested or who is interested to know about stock behavior of that company. Going through many surveys, the conclusion we come across is that it is not possible to consistently predict moments with better than average result as prices are affected by companies growth,revenue.*

***Keywords:*** *Data mining; Data preprocessing; Stock value; Stock exchange*

## I. Introduction

Price trend prediction of stock market application provides a way of predicting the prices of stock. The objective of the object analytics system is to generate efficient result of prediction for the user. The amount of data is tremendously increasing now a days due to advancement of storage techniques and digitization of work in every field.

The stock predictor is an application which provides the prediction of stocks shares' price of listed companies. In stock exchange, there are many losers and gainers in company in every day before the closing of Bombay stock exchange. This is a very helpful application for prediction of price of share of Bombay stock exchange by which user will get to know price of different shares directly in the application.

For any investor, there are two prices that are critical to know and those are thecurrent price and future price. The current priceis the investment he or she owns, or plans to own, andfuture price is its future selling price. Despite this, investors are constantly reviewing past pricing history and using it to influence their future investment decisions. Some investors won't buy a stock or index that has risen too sharply, because they assume it's due for a correction, while

other investors avoid a falling stock because they fear it will continue to deteriorate.Whenever reasonable predictions with less bias are produced by investors, great profit will be made. As the emergence of artificial intelligence (AI) algorithms has arisen in recent years, it has played an important role to help people forecast the future. In the stock market, many forecasting models were advanced by academy researchers to forecast stock price such as time series, technical analysis and fuzzy time-series models. However, there are some drawbacks in the past models: (1) strict statistical assumptions are required; (2) objective human judgments are involved in forecasting procedure; and (3) a proper threshold is not easy to be found.

The stock market process is full of uncertainty as well as is affected by many factors. Hence the Stock market prediction is one of the important exertions in finance and business. Price trend prediction of stock market provides the efficient solution for the user to predict future prices of stock. The successful prediction of stock's future price will yield significant profit to the investor.

## II. Literature survey

In [1], authors have proposed a data mining approach along with Artificial Intelligence for Prediction of Stock Market. The proposed system have used database of shares in the market. For prediction particularly ARMA (autoregressive-moving-average) algorithm is used. In this work, the authors have mentions the names of 10 methodology mostly used for stock value prediction for accuracy and performance along with its advantages and disadvantages.

In [2], authors have studied most important input parameters that have major impact on accuracy of predicting stock market prices like Technical parameters, Microeconomic parameters. The market price is highly dependent on the type of input parameter used for predicting it. Mostly the hybridized parameters like combination of technical and fundamental variable give better prediction accuracy.

In [3], authors have proposed a Natural Language Processing (NLP) algorithms to build a module for predicting a stock trend. The module analyzes the influence of news articles information which are rich in information and superior to numeric data based on grammatical analysis.

In [4], authors have proposed a text mining approach to analyze the current financial news items. They have used a Latent Dirichlet Allocation (LDA) based topic extraction mechanism to identify and characterize major events from
Financial news that impact the market.

In [5], authors have used an outlier data mining algorithm for predicting stock market values. A traditional algorithm of cluster analysis is used to detect the outliers which always predict an upward trend of the stock price. The result shows that, the proposed work will run effectively for stock market profit for long term uses.

In [6], authors have proposed service-oriented multi-kernel based learning framework (MKL) for Stock Market Volatility Prediction. The result shows that, as compared to existing single kernel methods a multi-kernel learning method has a higher degree of accuracy and a lower degree of false prediction. Both historical stock price information and news and trading volume data are integrated which significantly improves the effectiveness of stock market volatility prediction.

In [7], authors have explained different methods for stock movement prediction using sentiment analysis and data mining approach. They have used a social media a ever-growing source of texts for collecting peoples thoughts and feeling. Then Sentiment classifiers have been built for factors such as product reviews, blog posts, and even twitter messages.

In [8], for Stock Market Prediction, authors have done fundamental analysis using social media data by applying sentiment analysis process and Technical analysis using historical data of stock prices by applying machine learning algorithms. Then the learning model is built which analyses correlation between the sentiments and stock values.

In [9], authors have proposed a time series mining method to find the outliers from time series financial data. The system is based on the fact that the abnormal fluctuations in financial data often contain important information. This paper presents a method that combines the activity and density of time series to discover the anomalies.

For efficient stock market prediction, various models have been developed which makes use of artificial neural network model and machine learning model [10-12].

## III. Proposed system

This proposed system, analyzes previous stock data for certain companies, with the help of certain parameters that affect stock value. The proposed system will implement these values in data mining algorithms. This will also help us to determine the values that particular stock will have in near future. Proposed system will determine the month's high stock prize and low stock prize with help of outlier data mining algorithms. The software maintains administrator level and user level. It generates report for prediction on previous stock. However, according to efficient market hypothesis, the market price will follow a random walkandpermanent prediction strategy.Atraditionalwaytopredictstocktrendisusing data mining technique on the basis of stock prices unfortunately, the data of stock price have many noises and for noisy data people always build stochastic volatility models to make prediction. Anomalies have long term predictability onthe stock trend. In our method system will utilize anomalies to remove noise so thatprediction can be done easily. Stock market prediction is the act of trying to determine the futurevalue of a company stock or other financial instrument added on an exchange. The successful prediction of a stock's future price could yield significant profit.

The main purpose of this project is to achieve faster and efficient solution for stock prediction of Bombay Stock exchange which will meet user's expectation and fulfil specified requirement. Stock market prediction is the act of trying to determine the future value of acompany stock or other financial instrument added onan exchange. The successful prediction of a stock'sfuture price could yield significant profit. The main purpose of this project is to achieve faster and efficient solution forstock prediction of Bombay Stock exchange which will meet user's expectation and fulfil specified requirement. Many banks, financial institutions, large-scale investors and stockbrokers have to buy and sell stock in shortest possible time

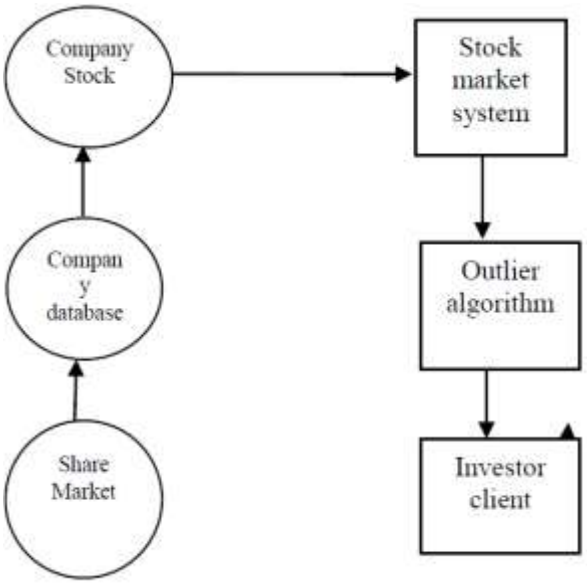


**Fig. 1.** DFD diagram for price trend prediction

## IV. Methodology

Past stock prices data of listed companies will be taken for prediction. Data will be selected for preprocessing to remove null values and holidays in data. Normalization is applied to reprocessed data for better comparisons of stockdata. Data mining algorithm is used to find the pattern recognition to find whether user should buy or sell stock. Figure 1 shows the following data flow diagram for price trend prediction and figure 2 shows flow diagram of price trendprediction

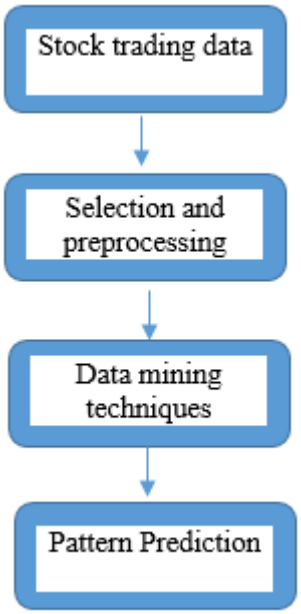


**Fig. 2.** Flow diagram for Price Trend Prediction

## V. Results

The result of proposed system shows the input interface for user as well as admin for mode of operation in system. Admin level has login which handles the database for stock values of listed companies. In user level, login as user consist of username and password as well as have to enter the number of company to see the stock values. User have to enter the start date and end date for corresponding companies in specified date format. Graph is generated which shows the flow of the stock values of companies for prediction ofstock price. Stock

values are normalized and preprocessed to remove the null values and holidays as well as plotted the stock price of specifiedcompanies.

**Before Normalization:**
Generally the data which are present in csv file are having some irregularity, noise,null values. Hence Normalization is necessary.It is shown in figure 3.

**After Normalization:**
In Normalization, stock prices are normalized as by removing null values,holidays,and irregular data. After normalization, we can easily compare the stock prices of different companies. Because the starting date will be same for all companies. It is shown in figure 4.

**Bollinger band(Buy or Sell signal):**
It is used to calculate the rolling mean of stock prices.It consists of lowerband and upperband which indicate the threshold value for stock prices. If stock prices will cross upperband then there has to be "sell" signal because after that the price has to come at mean and cross to lowerband then it has to be buy signal. It is shown in figure 5.

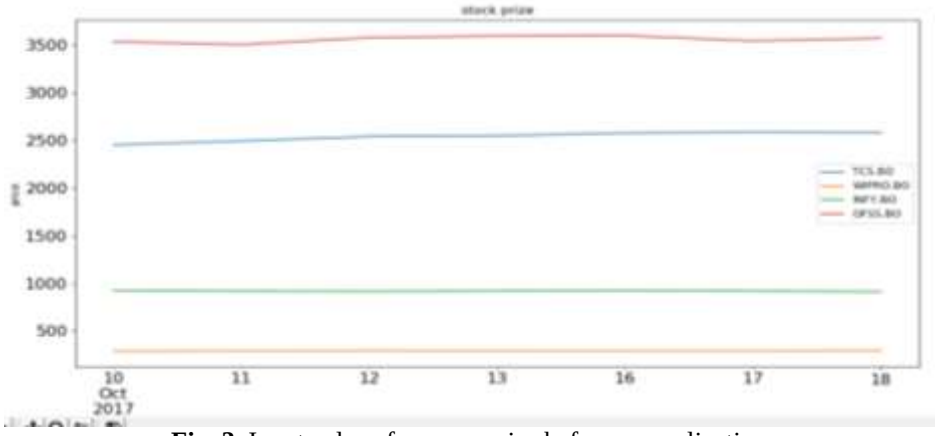


**Fig. 3.** Input values for companies before normalization

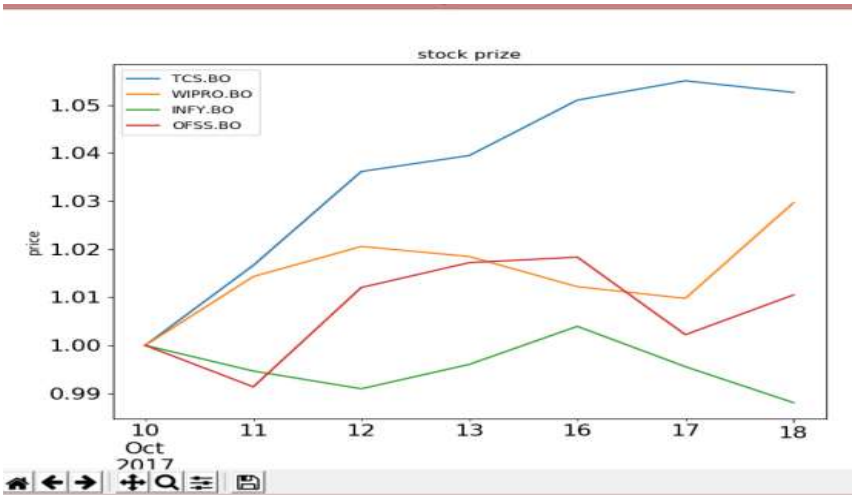


**Fig.4.** Stock values of companies after normalization


**Fig.5.** Bollinger band (Buy or Sell signal)

**Predicted price**

In this the whole one previous year stock prices of company is taken(Infosys). In this linear regression algorithm is used which gives the accuracy of 70%-80%. Past 1 year(Mar 2017-Mar 2018) data is taken.


**Fig.6.** Predicted price for Infosys

**Testing:**

As machine learning is used for prediction of stock price .Hence every run (no of jobs) of program will help to analyze previous stock data and learning to analyze. In this no of jobs is taken 10 times which gives accuracy 79% for 1 year data. Hence to increase the accuracy of system, datashould be more at least 6 month or greater.

**Comparison(Predicted price vs. Real Price):**

In this we have compared the real time stock prices of company with predicted prices. In this we have taken asstart_date='2017-10-26' and end date='2017-11-24'.

In this graph shown in figure 8,the dark blue color represents the predicted price for Infosys and light blue color represent the actual price for Infosys from 2017-10-26 to 2017-11-24.

As graph shows the predicted and actual prices are nearly same for Infosys.

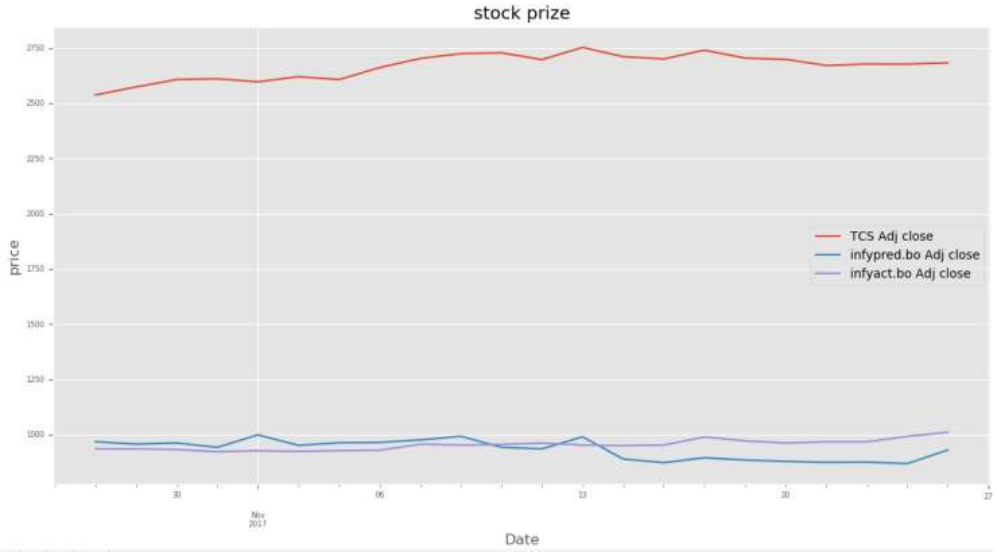


**Fig.7.** Comparison of Predicted price vs. Real Price of Infosys

Similarly, we can show the results of other listed companies also. The list of different companies is shown below in figure 8. We can add more number of company names as per the requirements.

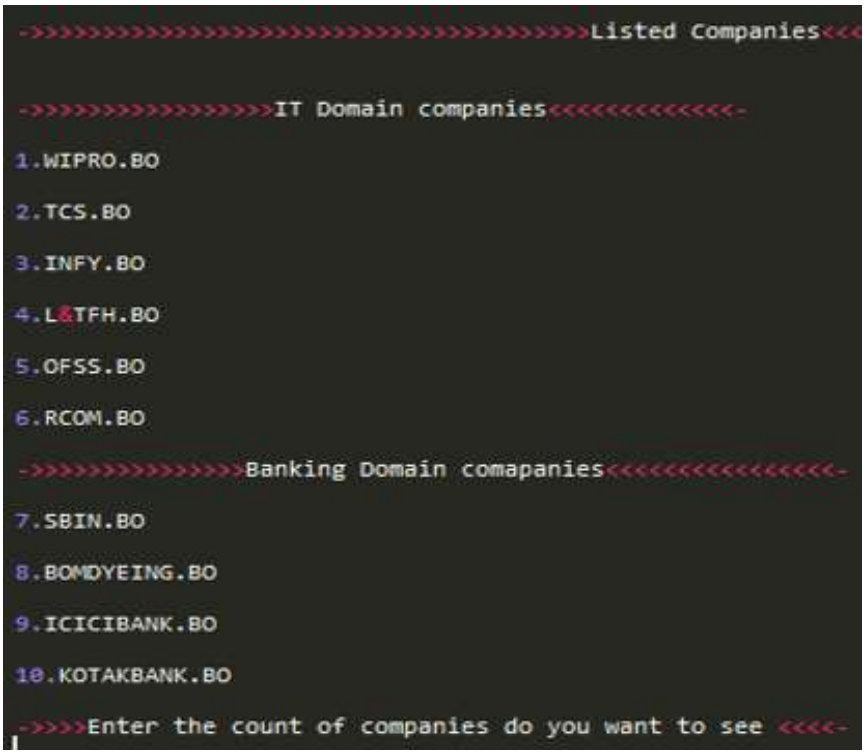


**Fig.8.** List of different companies

The following table shows the prediction values of three companies stock:Infosys, Wiproand OFSS for the period 2017-10-26 to 2017-11-22 (2 months).

**Table 1:** Stock prediction values for companies for 2 months

| **Date** | **INFOSYS** | **OFSS** | **WIPRO** |
|---|---|---|---|
| 2017-10-26 | 967.161071 | 3515.042478 | 267.084147 |
| 2017-10-27 | 956.292413 | 3507.727610 | 283.777036 |
| 2017-10-30 | 961.459968 | 3540.689153 | 270.183260 |
| 2017-10-31 | 941.612934 | 3523.870622 | 251.636673 |
| 2017-11-01 | 998.187053 | 3535.270540 | 267.143011 |
| 2017-11-02 | 950.657507 | 3502.476389 | 291.558488 |
| 2017-11-03 | 962.389291 | 3730.966880 | 282.516667 |
| 2017-11-06 | 963.900027 | 3479.409987 | 279.553962 |
| 2017-11-07 | 975.171910 | 3498.287850 | 297.874334 |
| 2017-11-08 | 991.431042 | 3479.250363 | 287.507045 |
| 2017-11-09 | 941.835377 | 3481.267710 | 290.778711 |
| 2017-11-10 | 934.456265 | 3603.277516 | 290.958496 |
| 2017-11-13 | 989.484357 | 3499.511685 | 285.026233 |
| 2017-11-14 | 888.639395 | 3486.237999 | 279.524194 |
| 2017-11-15 | 872.410279 | 3518.736509 | 298.280893v |
| 2017-11-16 | 894.479869 | 3489.721608 | 292.570050 |
| 2017-11-17 | 883.969456 | 3518.807137 | 282.259997 |
| 2017-11-20 | 878.233664 | 3477.872980 | 286.947381 |
| 2017-11-21 | 873.653883 | 3547.550196 | 260.959972 |
| 2017-11-22 | 874.672747 | 3514.059216 | 279.994085 |

The table 2 shows actual price and predicted price for Infosys on a given date. The accuracy of algorithm is based on amount of data pass to the it. If data passed is of one year, then it gives more accuracy than data passed for months. Over all, the average accuracy of the algorithm is 80-90%.

**Table 2:** Sample of predicted prices vs. Actual prices for Infosys

| Date | INFOSYS | | %Error |
|---|---|---|---|
| | (Predicted Price) | (Actual price) | |
| 2017-10-26 | 967.161071 | 934.460815 | 3.49 |
| 2017-10-27 | 956.292413 | 934.411499 | 2.34 |
| 2017-10-30 | 961.459968 | 931.650024 | 3.19 |
| 2017-10-31 | 941.612934 | 921.650024 | 2.16 |
| 2017-11-01 | 998.187053 | 926.250000 | 7.76 |
| 2017-11-02 | 950.657507 | 922.900024 | 3.0 |
| 2017-11-03 | 962.389291 | 926.650024 | 3.85 |
| 2017-11-06 | 963.900027 | 928.599976 | 3.8 |
| 2017-11-07 | 975.171910 | 955.700012 | 2.03 |

## VI. Future scope

In proposed system, machine learning algorithm is applied which is linear regression. To provide better support and prediction of stock market, sentiment analysis can be applied by which additional factor also can be taken for prediction. Additional factor includes top headlines,news related to each companies also affect the prediction of stock price of companies.

Although there has been some study regarding trend analysis and trading of stock market indices; a detailed study is, however, required for developing an intelligence system for making decisions regarding buying and selling of stocks and portfolios. In this regard, it is suggested to develop effective learning procedures for taking accurate decisions in stock market trading.

## VII. Conclusion

The main idea of prediction of stock is based on data mining algorithm. In this system, linear regression algorithm is implemented which generally gives accuracy of 80-90%. For analyzing previous stock data, a machine learning algorithm is used. Linear regression is used to find multiple factors that can affect the stock prices of company as daily return value, open price, volume etc. Data is normalized and preprocessed for better comparison among different companies. Another data mining algorithm can be applied as KNN and SVM which gives the accuracy nearly as linear regression.For better prediction of stock market, more previous data (previous stock prices of companies) is given which gives better analysis.

## References


[1]. G. S. Navale,NishantDudhwala,KunalJadhav, "Prediction of Stock Market using Data Mining and Artificial Intelligence", International Journal of Computer Applications (0975 – 8887) Volume 134 – No.12, January 2016

[2]. Prashant S. Chavan, Shrishail. T. Patil, " Parameters for Stock Market Prediction", ISSN:2229-6093 , IJCTA | Mar-Apr 2013

[3]. Ruchi Desai, Snehal Gandhi, " Stock Market Prediction Using Data Mining", 2014 IJEDR | Volume 2, Issue 2 | ISSN: 2321-9939

[4]. Anuj Mahajan, LipikaDey, Sk. MirajulHaque," Mining Financial News for Major Events and Their Impacts on the Market", 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology

[5]. Zhao, Lei, Wang, Lin," Price Trend Prediction of Stock Market Using Outlier Data Mining Algorithm", 2015 IEEE Fifth International Conference on Big Data and Cloud Computing

[6]. Feng Wang, Ling Liu, Chenxiao Dou, "Stock Market Volatility Prediction: A Service-Oriented Multi-Kernel Learning Approach", 2012 IEEE Ninth International Conference on Services Computing

[7]. Vivek Rajput1, SarikaBobde," Stock Market Forecasting Techniques: Literature Survey", ISSN 2320–088XIJCSMC, Vol. 5, Issue. 6, June 2016

[8]. VivekKanade, BhausahebDevikar,SayaliPhadatare, "Stock Market Prediction: Using Historical Data Analysis", ISSN: 2277 128X, Volume 7, Issue 1, January 2017,International Journal of Advanced Research in Computer Science and Software Engineering.

[9]. Chen-Ming Guo, Ling-Yu Xu, Hui-Fang Liu, Lei Wang, " The Financial Data of Anomaly Detection Research based on Time Series", 2015 International Conference on Computer Science and Applications

[10]. Ashraf S. Hussein, Ibrahim M. Hamed, Mohamed F. Tolba, "An Efficient System for Stock Market Prediction",

[11]. Springer International Publishing Switzerland 2015, DOI: 10.1007/978-3-319-11310-4_76

[12]. Osman Hegazy , Omar S. Soliman, Mustafa Abdul Salam," A Machine Learning Model for Stock Market Prediction", International Journal of Computer Science and Telecommunications [Volume 4, Issue 12, December 2013]

[13]. MingyueQiu, Yu Song, "Predicting the Direction of Stock Market Index Movement Using an Optimized Artificial Neural Network Model", Published online 2016 May 19. doi: 10.1371/journal.pone.0155133

[14]. http://www.investopedia.com/articles/07/mean_reversion_martingale.asp

[15]. http : // www.nseindia.com/content/research

[16]. http: //en.wikipedia.org/wiki/stock_market_prediction