IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, February 2022
Impact Factor: 6.252

Stock Market Prediction Using Random Forest Algorithm

Ashwini Patil, Rushikesh Rajegaonkar, Sakshi Sapkale, Aarya Sawale
Department of Computer Engineering
Guru Gobind Singh College of Engineering and Research Centre Nashik, Maharashtra, India
ashwinisanjay25@gmail.com, rushikeshrajegaonkar@gmail.com,
sakshijs.12601@gmail.com, aaryasawale@gmail.com

Abstract: *Investment is acknowledged as one of the powerful tools in the alleviation of poverty. Investors typically perform technical or fundamental analysis to determine favorable investment opportunities, and generally prefer to minimize risk while maximizing returns. They are supposed to be risk-adverse, safety oriented and guided by certainty of returns. The volatile nature of the stock market makes it difficult to apply simple time-series or regression techniques. The proposed work aims to involve more women in trading and investments in the stock market. The objective of this project is promoting women's sense of self-worth, their ability to determine their own choices, and their right to power up a social change for themselves and other fields. A defensive stock is a stock that provides consistent dividends and stable earnings regardless of the state of the overall stock market. With an aggregate level of knowledge and awareness, encouraging women to take part in the risk investment portfolios and change the behavior of looking towards a male dominated sector and become systematic in their investment in the stock market.*

Keywords: Women, Stock Market, Defensive Stock Market, Prediction Techniques, Portfolios

Problem Statement: Among Indian Stock Traders, 16.1% of them are women compared to 76.6% which are men. Womens just make savings to support their family and daily needs. Instead, we want them to achieve some personal goals by taking some risks and reap huge rewards. Women constitute half of the world's population but still lack in male dominated sector such as the stock market. In Economics, women can make enormous contributions whether in the stock market by investing some small amount in shares from their savings. We aim to promote women's sense of self- worth, their ability to determine their own choices, and their right to influence social change for themselves and others.

I. INTRODUCTION

The Stock Market is a very chaotic system. The paper looks forward to helping everyone by actually predicting the vicissitudes in the stock market through a suitable machine learning algorithm with better accuracy to bring about maximum profit through investing just small amounts of the savings. Investment In stock market is partially risky. The stock market is a highly wide range of expenditure and merchants who buy and sell shares show the price up or down to acquire the profit. It is in the vicinity impracticable to forecast stock prices to the T, owing to the instability of factors that perform a vital role in the movement of prices. Economics and bid prices are mainly contingent upon subjective impression about the stock market.

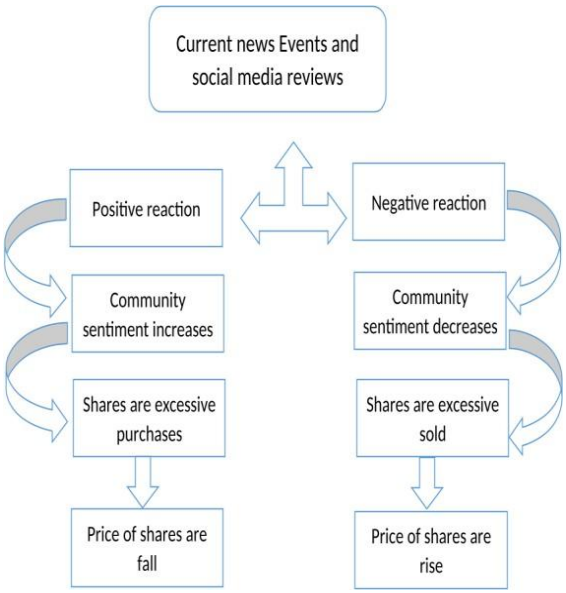
Nowadays people are stationing their commentary views on social media which can be shared with other people also. There are some mechanism for stock market prediction that comes under technical analysis:

1.1 Sentiment Analysis

Sentiment analysis is an approach that is used in relation to the latest trends. It observes the trends by analyzing news and social trends like tweet activity. Specifically, we analyze the stock's contents and extract financial sentiment using a

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, February 2022
Impact Factor: 6.252

set of text featurization and machine learning algorithms. The correlation between the aggregated daily sentiment and daily stock price movement is then studied. Using our algorithm for sentiment analysis, the correlation between the stock market values and sentiments in RSS news feeds are established. This trained model is used for prediction of stock market rates. Using our algorithm for sentiment analysis, the correlation between the stock market values and sentiments in RSS news feeds are established. This trained model is used for prediction of stock market rates.



1.2 Random Forest Algorithm

The Random forest algorithm is used in many sectors. One of the most popular being stock market prediction itself. The random forest algorithm is a supervised learning algorithm. Random Forest builts multiple decision trees and merges them together to get a more accurate and stable prediction. Random forest arrives at a decision or prediction based on the maximum number of votes received from the decision trees

1.3 Mean Reversion

Mean reversion, or reversion to the mean, is a theory used in finance that suggests that asset price volatility and historical returns eventually will revert to the long-run mean or average level of the entire dataset. Mean Reversion model based on specific data that successfully demonstrates the results that indicate that ignoring structural breaks that arise from the liberalization of emerging markets can lead to incorrect inference.

II. Related Work

It is widely believed incorrectly that share market is a part of gambling and is just exclusively available for expert men. According to recent studies done by Choose broker in 2021, the global ratio of people investing in the share market is given as 76 percent male to 24 percent female. It has also found that in Emerging Markets, the Philippines has a female investor ratio of 44 percent which is the highest number of female investors; while in India the ratio decreases straight to 21 percent.

While, if observed meticulously and believing studies from over the last 10 years, it proves that when women investors are compared to male investors, women investors outperformed their male counterparts by 40 percent. Adding up to this, on an average, women investors. A significant amount of development gain in relation to economic growth, poverty eradication and the well-being of families and communities can be achieved by ensuring the participation of women in areas that were earlier thought to be dominated by men. Information about extending the involvement and the obstacles faced by women while investing in the stock market will support the state government to take restorative

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, February 2022
Impact Factor: 6.252

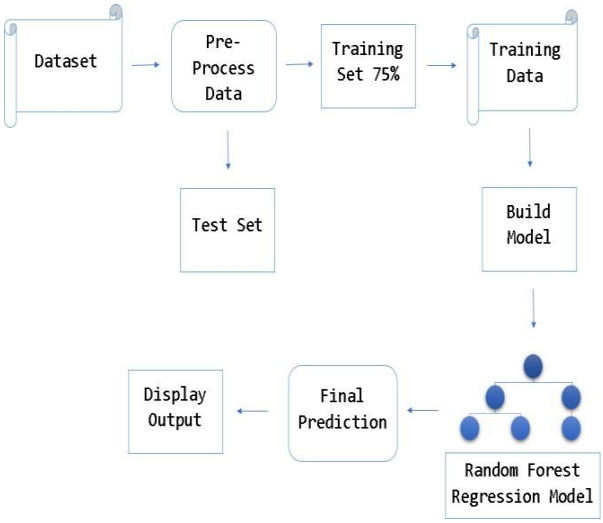
measures to ensure equal involvement of women in all economic spheres. The final output value does not depend on a single classifier, it is dependent on multiple classifiers that is why the output using random forest is more accurate.

Daily text content on Twitter is analyzed by mood tracking tools and investigated for predicting the changes of DJIA closing values, which was based on a self-organizing fuzzy neural network. One of the data mining techniques called a decision tree classifier is used to take the decision in the stock market for buying or selling stocks based on the knowledge extracted from the historical prices of stock.

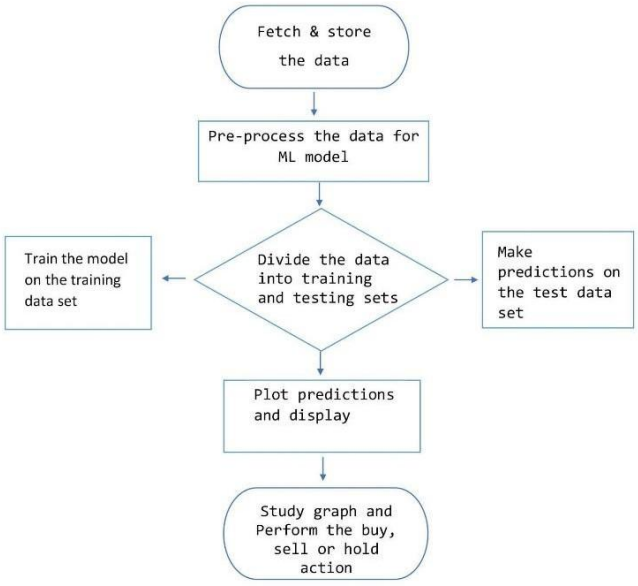
The increasing importance of stock markets is pushing researchers into finding newer methods for prediction which not only helps researchers but trending alike. Predicting stock market returns is a challenging task due to consistently changing stock values which are dependent on multiple parameters which form complex patterns.

By using various machine learning algorithm models, the price of the stock is predicted on the basis of closing value and stock price. Accuracy is calculated by comparing machine learning algorithms.

III. ARCHITECTURE



IV. FLOWCHART



IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, February 2022
Impact Factor: 6.252

V. METHODOLOGY

5.1 Dataset

The dataset used here is from kaggle. The data illustrates the specifics of stocks concomitant to 160 Indian companies whose stocks actually influence the market. It is the data that was collected from 2018 to 2021 while it is the most recent and most useful data. It is a raw dataset that should be pre- processed and analyzed before using it for the given work.

Feature	Description
Timestamp	Time and Date with reference to Greenwich Mean Time
Open	Starting price at which a stock is traded in a day
High	Highest price of equity symbol in a day
Low	Lowest price of share in a day
Close	Final price adjusted for splits
Volume	The number of shares that changed hands during a given day

5.2 Data pre-processing

Data pre-processing is a method in which we rebuild our data before giving it to the algorithm. In this, we clean the data set as per the needs of our algorithm. When we gather any such data we need to perform some filtration like filling up or expelling the null values, discarding the noise to turn it into usable data.

5.3 Classifier

Random forest is a supervised machine learning algorithm used to perform classification and regression on datasets. This is the most suitable algorithm that can be used for predicting the future prices of the stock market. It builds different decision trees by making random subsets of the given dataset and increases the accuracy of the presented model.

What is Decision Tree?

In the machine learning world, decision trees are a kind of non parametric model that can be used for both classification and regression. This means that decision trees are flexible models that don't increase their number of parameters as we add features (if we build them correctly), and they can either output a categorial prediction (like if a plant is of a certain kind or not) or a numerical prediction (like the price of house).

5.4 Long Short-Term Memory (LSTM) Network:

LSTM network is a type of deep RNN model composed of LST-M units. RNN is a deep learning network with internal feedback between neurons. These internal feedbacks enable the memorization of significant past events and incorporate past experience. RNN shares parameters across all the parts of a model, so it can be generalized to sequence lengths that have not been seen during training. presents an example of RNN architecture that produces an output at every time step, and has recurrent connections among Sustainability hidden neurons the LSTM block contains a memory cell and three multiplicative gating units; an input, an output, and a forget gate. There are recurrent connections between the cells, and each gate provides continuous operations for the cells. The cell is responsible for conveying "state" values over arbitrary time intervals, and each gate conducts write, read, and reset operations for the cells.

The data collected was normalized for values between 0 and 1 using LSTM:

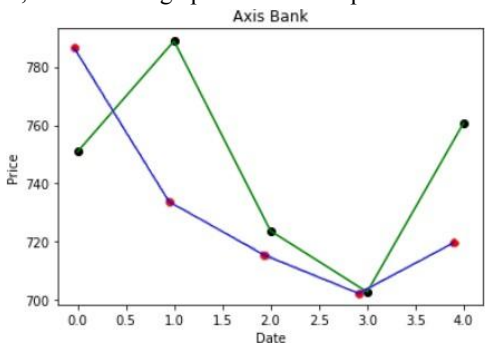
$$Xnew = \frac{X - Xmin}{Xmax - Xmin}$$

By using various machine learning algorithm models, the price of the stock is predicted on the basis of closing value and stock price. Accuracy is calculated by comparing machine learning algorithms.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, February 2022
Impact Factor: 6.252

VI. RESULTS

As we see, these types of problems whose output is to be predicted can be solved using Random Forest Algorithm which is a part of unsupervised learning. This algorithm is best suited when we know the input but we need our output as the answer. The following graph shows predicted results with respect to the real time closing price. The green line-graph indicates the real time prices. Whereas, the blue line graph indicates the predicted closing prices of Axis Bank.



VII. ADVANTAGES

1. If we do it accurately, fast and secretly, for as long as the market assumptions stay stationary, you will get rich very quickly with relatively little labor input.
2. Stock market prediction aims to determine the future movement of the stock value of a financial exchange. The accurate prediction of share price movement will lead to more profit investors can make.
3. For more advanced decision making, the more data that the software makes available the better-informed decisions you can make. It can identify patterns and trends in large amounts of data, providing organizations insight that previously may not have been available.
4. Stock Market Prediction through Mean Reversion helps to forecast and manage the type of investment, to make organizations more efficient, and help to optimize the amount of investment and increased revenue. It helps proactively improve their investment and take appropriate actions when needed.
5. Determining customer responses and purchases is very important in marketing strategies, and Predicting stocks helps to analyse data to identify new opportunities to attract or retain customers.

VIII. APPLICATIONS

1. Forecasting and predicting the trends of the market are the most important applications of stock market prediction models.
2. It also uncovers the future market behaviour which always helps the investors to understand when and what stocks can be purchased for the growth of their investment.
3. Predictions are used to determine customer responses or purchases, as well as promote cross-sell opportunities.
4. Prediction of the stock market helps businesses attract, retain and grow their most profitable customers.
5. Many companies use predictive models to forecast inventory and manage resources.
6. Better use in analyzing market ups and downs.

IX. CONCLUSION

We have effectively executed the following process with the help of a random forest algorithm. Among all the methods this method is more accurate and best suitable than other methods. Random Forest gives 80.7% accuracy than all other classifiers. We also found the advantages as well as the applications of the following algorithm. Thus we conclude that Random Forest successfully demonstrates better results.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, February 2022
Impact Factor: 6.252

REFERENCES

[1]. Rajni Jindal, Nikhil Bansal, Nitin Chawla, Sanskriti Singhal: "Improving Traditional Stock Market Prediction Algorithms using Covid-19 Analysis."

[2]. Ashwini Pathak, Sakshi Pathak: "Study of Machine learning Algorithms for Stock Market Prediction, 2020. Vol. 9 pp 295-300."

[3]. Jaspreet Kaur Theth,Jaspreet Kaur Theth,Hitakshi Pate,Prof.Vaishali Shirsath :"Stock Market Prediction and Portfolio Management using ML techniques"

[4]. Mandeep Kaur, Tina Vohra: "Women and Stock Market Participation A Review of Empirical Evidence."

[5]. Mandeep Kaur, Tina Vohra: "Women and Stock Market Participation A Review of Empirical Evidence."

[6]. Shri Bharathi, Angelina Geetha: "Sentiment Analysis for Effective Stock Market Prediction"

[7]. Jaspreet Kaur Thethi, Aditi Pandit, Hitakshi Patel, Prof. Vaishali Shirsath: "Stock Market Prediction and Portfolio Management using ML techniques."

[8]. M.Ferni Ukrit1, A.Saranya, Rallabandi Anurag: "Stock Market Prediction using Long Short Term Memory"

[9]. Syed Saad Azhar Ali, Kamran Raza, Syed Hasan Adil, Mehak Usmani : "Stock Market Prediction using Machine Learning Techniques"

[10]. Ray Chen, Marius Lazer: "Sentiment Analysis of Twitter Feeds for the Prediction of Stock Market Movement "

[11]. Sahil vazirani, Abhishek sharma and Pavika sharma: "Analysis of various machine learning algorithm and hybrid model for stock market prediction using python"

[12]. Manolis Maragoudakis, Dimitrios Serpanos: "Towards Stock Market Data Mining Using Enriched Random Forests from Textual Resources and Technical Indicators"

[13]. Subba Rao Polamuri, K. Srinivas, A. Krishna Mohan: "Stock Market Prices Prediction using Random Forest and Extra Tree Regression", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8 Issue-3, (September 2019)

[14]. Jingyi Shen and M. Omair Shafq: "Short-term stock market price trend prediction using a comprehensive deep learning system"(2020).