Stock Market Prediction Using Machine Learning Algorithms

Pallavi N[1], Deepak Nair[2]
[1] Assistant Professor, Dept of Computer Science & Engineering
[2] Dept of Computer Science & Engineering
[1, 2] Atria Institute of Technology, Bangalore

Abstract- *These days prediction of stock market has become difficult. Individuals and organizations are getting highly involved in buying company shares and are dependent on prediction technologies to receive maximum gain. The most important aspect here is the data set of stock market prices from previous year. In this paper, techniques like random decision forests and support vector machine will be discussed. The dataset will be processed against support vector machine and random decision forest and their outcomes would be recorded. With accurate predictions of stock market, the stock investors would be highly benefited.*

Keywords- Dataset, Machine Learning, Stock, Stock investors, Stock Market

I. INTRODUCTION

The stock market is where investors connect to buy and sell investments, most commonly known as stocks, which are shares of ownership in a public company. It is a market where the trading of company stock, both listed and unlisted securities takes place. It is different from stock exchange as it includes all of the stock exchanges in the country.

Determining the future value of stock market based on previous outcomes is called as stock market prediction. The system needs to be tuned with respect to real world situations. With the advancement in technologies, techniques involving artificial neural networks, recurrent neural networks, which are implementations of machine learning.

The stock market price develops a linear curve over a long period of time. Investors usually buy those stocks whose prices are predicted to increase in the future. The uncertainty in stock market often intimidates people and a lot of people don't invest in them. Hence, there is a need of predicting the stock market as accurately as possible. Stock market prediction methods include machine learning, conventional time series and predicting variable stock market.

Compared to other supervised learning techniques, support vector machines have better generalization performance anddo not try to over-fit the problem. The support vector machine can be used for both classification and regression. In problem like ours, support vector machines are highly used.

The existing system has a handful of disadvantages, which has troubled the investors a lot. The current model does focus on social media or news aspects. When using traditional classifiers, it is very unpredictable as seen in previous results of stock price. When there is a change in operating environment, the performance of the existing model reduces. Also, since it only exploits one data source, it is highly biased.

II. LITERATURE SURVEY

A. Survey of Stock Market Prediction Using Machine Learning

At the present time, stock market prediction has become a major issue. The present analytical techniques are insufficient to acquire the expected results. After considering all factors in the stock market, predictions are made and usually investments are done based on these predictions. Regression was the technique which was used to attain this. There are different ways of fitting linear regression like, using the least square approach, diminishing the "lack of fit" in some other norm, or by diminishing a handicapped version of the least squares loss function.

B. Financial Ratios and Technical Analysis Using Random Forests

In today's worlds the usage of machine learning and artificial intelligence is growing. This has added up to the variety of ways in which stock market could be predicted. The outcome of each of these techniques varies even if the provided dataset is same. However, the price of stock can be influenced if factors like investor's emotion, company opinion in the eye of the public, news and some events can also impact the entire stock market. Hence, the accuracy of the stock price prediction model can be increased by using the financial ratio along with model.

Stock Price Prediction Using Support Vector Machines

Investors have always tried to use various kinds of techniques and models to try and achieve higher profit. The main feature about stock forecasting is that just a few rate focuses can benefit with a large amount of dollars for the investors.

Stock Market Prediction Based on Independent Component Analysis (ICA) and Support Vector Machines (SVM)

Various financial institutions have worked on time series prediction problem. The combination of both the SVM and independent component analysis is introduced for stock market prediction, called SVM-ICA. SVM solves the regression problem in non-linear classification and time series analysis. Various important features are extracted from the dataset using the ICA technique.

Machine learning Approach In Prediction Of Stock Market

Several techniques were introduced to predict stock market, but these techniques were unreliable. Hence, in order to achieve better accurate results, machine learning and artificial intelligence are implemented. The accuracy of the outcomes increased when CS-SVM hybrid was proposed in comparison to ANN. Similarly, the performance of CS-SVM is better in forecasting of stock value prediction.

Prediction Using Historical Data Analysis

Multiple factors can affect the stock market prediction significantly. For business and finance, stock market plays an essential role. With the help of sentimental analysis process, technical and fundamental analysis is done. Social media also has a very high impact on stock market. By using machine learning algorithms on historical data of stock prices, technical analysis is done. Social media data is collected in order to extract sentiments expressed by individuals. All the data points are connected with a relationship and this relationship is used to predict future stock values.
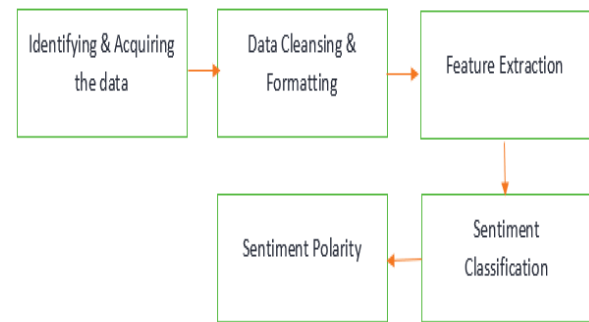
III. METHODOLOGIES

Based on the common attributes, a set is analyzed and categorized. This is an instance of supervised learning. From the observed values, conclusions are drawn. If more than one input is given then classification will try to predict one or more outcomes for the same.

One of the type of classifier is random forest classifier, it is a supervised algorithm. Results are obtained from the decision trees created by it. Another type of classifier is the SVM classifier. This is a type of discriminative classifier. It uses a labeled training data. Associate learning algorithms are used in SVM during classification and regression. Random forest classifier has the same hyper parameters as that of a decision tree. Decision tool has a model similar to that of a decision tree. From the several decision tree outcomes, aggregate is taken, this happens after the algorithm randomly selects observation and features to build several decision trees.
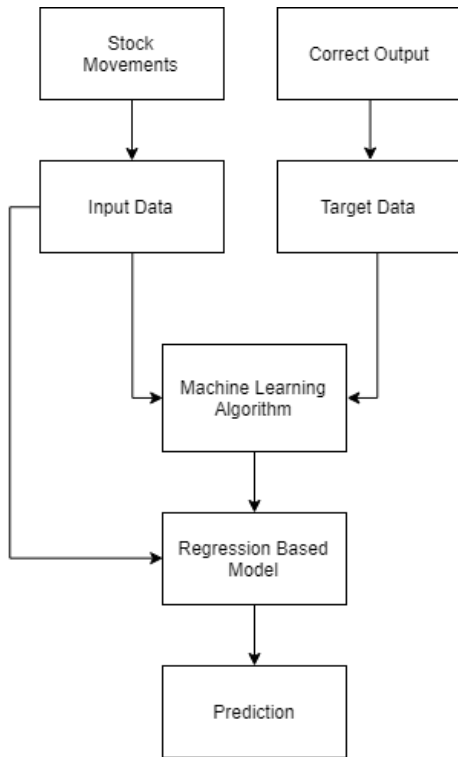
Fig 1: Regression based model

IV. PROPOSED SYSTEM

The proposed system is focused on the values of Support Vector Machines and Random Forest algorithms based on the concept of machine learning. In this proposed system, the machine was trained with datasets from previous stock market outcomes in order to predict future stock market prices. A couple of libraries from machine learning were used to solve the problem. They are scikit and numpy. Scikit was used for real analysis and prediction. Numpy was used to clean and manipulate the data and getting it ready for analysis.

Among the datasets which were obtained from previous outcomes of stock market, majority of those were used to train the system and the rest was used to test the system. The patterns and relationships in the data are the fundamentals in supervised learning approach.

Data processed was done by using python pandas library, which combined different datasets into a data frame. The date and closing price were the features of data frame. The machine was trained by using all the features of data frame and the price for a given day was predicted.

V. SYSTEM ARCHITECTURE

The dataset being used in the proposed project is obtained from Kaggle. Kaggle is an online community where the datasets are provided by data miners. This community allows the users to use the datasets in order to build models and achieve solutions for real life data science problems. However, these datasets are in raw format and need to be processed first. Feature extraction is used for this purpose, only the attributes which are useful for prediction are taken into consideration.

So, from the whole list of attributes, key attributes are extracted. This is followed by a classification process where the data obtained from extraction is divided into segments. Here, the training data has a higher proportion than testing data. To analyze the data, random forest algorithm uses a collection of random decision trees. These are then classified into different segments, known as data splitting.
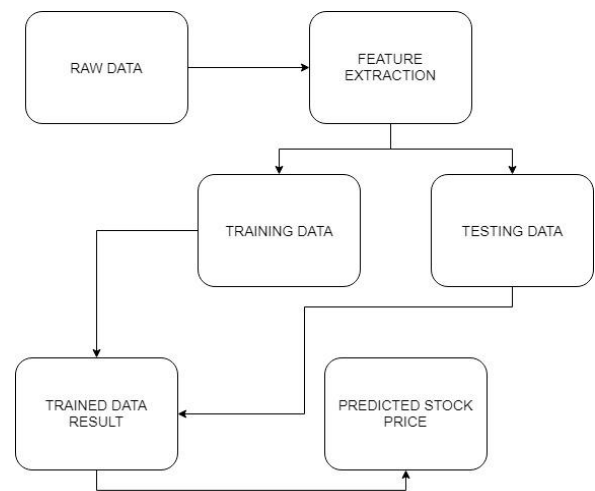

Fig 2: System Architecture

VI. FUTURE ENHANCEMENT

More parameters and factors would be involved in the future scope of this project. The accuracy will increase even more when the parameters increase. This project can be tweaked and used in other sectors as well. The sectors where outcome could be decided on the basis of datasets obtained from previous encounters. Using the traditional techniques and algorithms, prediction of a corporation performance can be determined, this helps the corporation as a whole.

VII. CONCLUSION

After trying out different algorithms and measuring their accuracies, we came up with the most appropriate algorithm to predict the market price on the basis of the datasets obtained from previous changes in stock market, which is the random forest algorithm. This algorithm will be useful for a lot of brokers and investors who invest money in stock market. This model of machine learning predicts stock

values more accurately than the previous models of machine learning. Hence, more accurate results.

VIII. ACKNOWLEDGEMENT

REFERENCES

[1] Ashish Sharma, Dinesh Bhuriya, Upendra Singh. "Survey of Stock Market Prediction Using Machine Learning Approach", ICECA 2017.

[2] SachinSampatPatil, Prof. Kailash Patidar, Asst. Prof. Megha Jain, "A Survey on Stock Market Prediction Using SVM", IJCTET 2016.

[3] VivekKanade, BhausahebDevikar, SayaliPhadatare, PranaliMunde, ShubhangiSonone. "Stock Market Prediction: Using Historical Data Analysis", IJARCSSE 2017.

[4] Loke.K.S. "Impact Of Financial Ratios And Technical Analysis On Stock Price Prediction Using Random Forests", IEEE, 2017.

[5] RautSushrut Deepak, ShindeIshaUday, Dr. D. Malathi, "Machine Learning Approach In Stock Market.

[6] Prediction", IJPAM 2017.

[7] Pei-Yuan Zhou , Keith C.C. Chan, Member, IEEE, and Carol XiaojuanOu, "Corporate Communication Network and Stock Price Movements: Insights From Data Mining", IEEE 2018