International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 ISSN: 2582-3930

STOCK PRICE PREDICTION USING MACHINE LEARNING

Swarup G. Dhote[1], Shubham S. Deshmukh[2], Amruta H. Mohod[3],
Suyoga A. Nimbhorkar[4], Prof. A. R. Kale[5].
[1]Student, P.R. Pote Patil College of Engineering and Management, Amravati
[2]Student, P.R. Pote Patil College of Engineering and Management, Amravati
[3]Student, P.R. Pote Patil College of Engineering and Management, Amravati
[4]Student, P.R. Pote Patil College of Engineering and Management, Amravati
[5]Professor, P.R. Pote Patil College of Engineering and Management, Amravati

ABSTRACT— In today's economy, there is a profound impact of the stock market or equity market. Prediction of stock prices is extremely complex, chaotic, and the presence of a dynamic environment makes it a great challenge. Behavioral finance suggests that decision-making process of investors is to a very great extent influenced by the emotions and sentiments in response to a particular news. Thus, to support the decisions of the investors, we have presented an approach combining two distinct fields for analysis of stock exchange. The system combines price prediction based on historical and real-time data. SVM and Random Forest is used for predicting. It takes the latest trading information and analysis indicators as its input.

Stock market investment strategies are complex and rely on an evaluation of vast amounts of data. In recent years, machine learning techniques have increasingly been examined to assess whether they can improve market forecasting when compared with traditional approaches. The objective for this study is to identify directions for future machine learning stock market prediction research based upon a review of current literature. A systematic literature review methodology is used to identify relevant peer-reviewed journal articles from the past twenty years and categorize studies that have similar methods and contexts. Four categories emerge: artificial neural network studies, support vector machine studies, studies using genetic algorithms combined with other techniques, and studies using hybrid or other artificial intelligence approaches. Studies in each category are reviewed to identify common findings, unique findings, limitations, and areas that need further investigation. The final section provides overall conclusions and directions for future research.

I.INTRODUCTION

Stock price prediction is a classic and important problem. With a successful model for stock prediction, we can gain insight about market behavior over time, spotting trends that would otherwise not have been noticed. With the increasingly computational power of the computer, machine learning will be an efficient method to solve this problem. However, the public stock dataset is too limited for many machine learning algorithms to work with, while asking for more features may cost thousands of dollars every day. With the growth of the Internet, social networks, and online social interactions, getting daily user predictions is feasible job. Thus, our motivation is to design a public service incorporating historical data and users' predictions to make a stronger model that will benefit everyone.

II. LITERATURE REVIEW

The main methods for predicting stock prices are all built around the basic premises of fundamental and technical analysis, although recent research shows that stock prices have a strong correlation with company news reports. The financial analyst used hourly stock prices for 30 different stocks along with their corresponding news articles and tweets on the company from the Nasdaq website. They also collected tweets related to all 4,444 of these 30 actions over a period of six months. Song et al. Six data of years was collected from the Hong Kong stock market. They gathered all of the financial information for those particular companies and stocks and their corresponding contemporary news articles and tweets to essentially derive correlations between the news articles and stock market trends. Also, for a specific trading day, they record the open price, the close price, the high price and the low price of a specific stock for each company. Deep learning is currently having a major impact on the models used for sentiment analysis. The cutting-edge architecture is today the model of attention and its modifications.

Aspect-Based Sentiment Analysis (ABSA), was not often used at first due to the limitation that can only recognize one aspect and requires a large amount of -tagged data, but the development of has seen an increase. Wang et al. It is proposed to learn the integration vectors for each facet and add the input facet integrations to each word input vector to better benefit from the facet information.

Dogra et al. [8] conducted a detailed study of the efficiency of many classifiers such as KNN, Random

Forest, SVM (Support Vector Machines) and Naive Bayes [11][12][16] in predicting stock trends. Therefore, we can point out that while the accuracy of naive Bayesian models is 4044 significantly improved, SVM [10] and f-measures of some different 4044 algorithms such as random forests show that they generally have better performance. The deep learning model is very expensive. They also do not give a good idea of stock prices compared to other simple methods, so they are not appropriate choices for integration [15]. Also, simple traditional techniques such as open, close, high, low and moving averages are considered more useful in estimating the future of for a particular security. The current stock market trend shows the correlation with the past series of the security for a given time. The naive Bayesian approach [16] only aims to predict the closing price of a specific stock based on other market values such as the opening price on a specific date. In this case, calculating the price change over time can significantly improve the accuracy of the model. Bo et al. [9] proposed a feature of the measurement technique using a genetic algorithm.

III.SYSTEM ARCHITECTURE

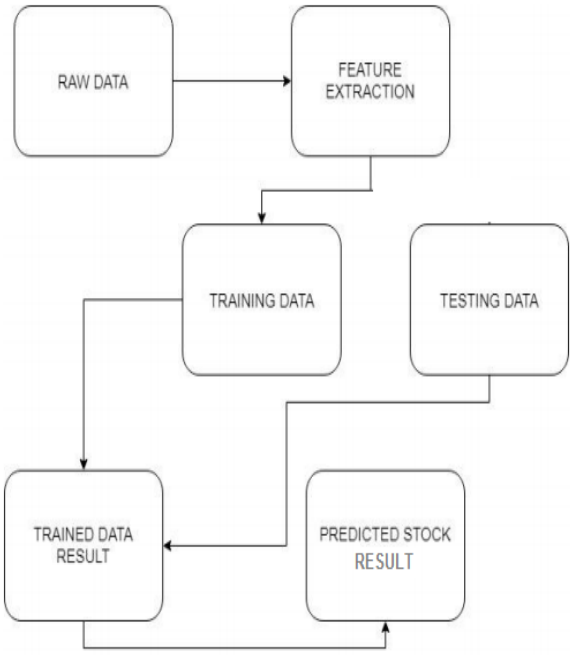

Fig 3.1: System Architecture of Stock Price Prediction

First the data is assembled. Then after that from the collected raw data feature will be extracted then data will be sent for training and testing .70% of a data will undergoes for training and 30% of a data will go for the testing purpose. When the train data result is out the testing data result is also collected at the block of train data result and from that block, we will predict the stock price which is important for the buyers to make the decision whether they have to buy the stock or not.

System design is the most important and vital part of any framework as it is used for the development of the system from its theory. This section includes the modules, architecture and various elements that are combined together to form the whole systems framework .We are loading the data set of about 5 years in both SVM and random forest model. We will read the stock data from Kaggle website the data will be stored in (open, high, low, close) format in a csv file.

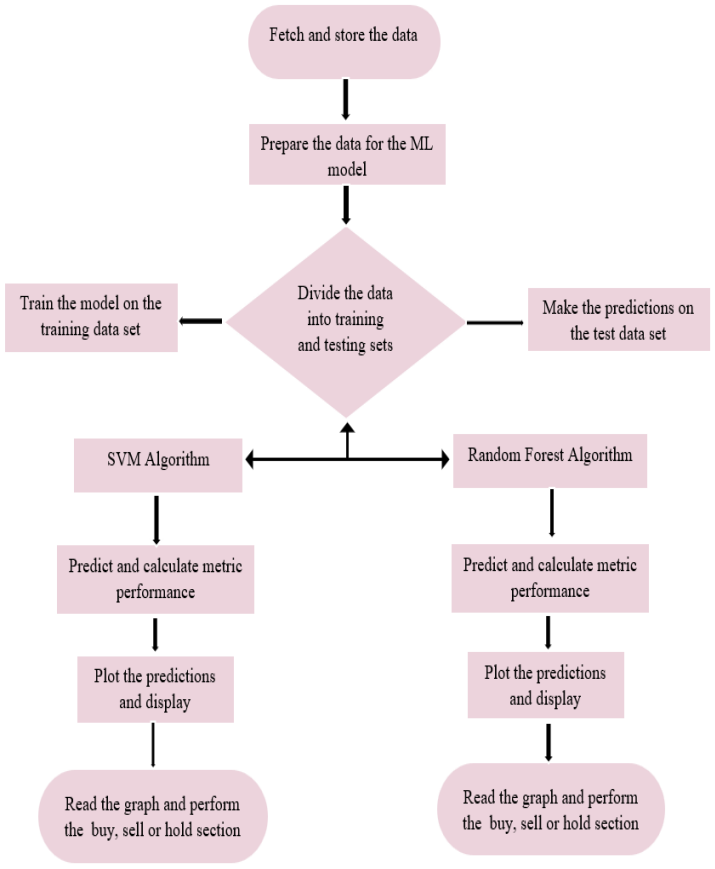

Fig 3.2: Flowchart of Proposed System

The data will be processed before use we will split the data into training and testing datasets. We are doing this so that we can evaluate the effectiveness of the model in the test

dataset, so basically, we are going to create an accuracy model and by using these two-machine learning algorithm that is SVM and random forest algorithm and from that we are going to calculate the accuracy of the stock. Calculating the metric performance, we are going to plot predictions and display to the user and then the user will decide whether to buy stock or not or whether he or she will be going to make profit or loss. After comparison we are going to choose the algorithm and the dataset associated with it which will give the highest amount of accuracy.

IV.OBJECTIVES

- Stock market prediction is basically defined as trying to determine the stock value and offer a robust idea for the people to know and predict the market and the stock prices.
- It is generally presented using the quarterly financial ratio using the dataset. Thus, relying on a single dataset may not be sufficient for the prediction and can give a result which is inaccurate. Hence, we are contemplating towards the study of machine learning with various datasets integration to predict the market and the stock trends.
- The problem with estimating the stock price will remain a problem if a better stock market prediction algorithm is not proposed. Predicting how the stock market will perform is quite difficult.

V. ADVANTAGES & DISADVANTAGES

ADVANTAGES

- The successful prediction will maximize the benefit of the customer.
- In this project we have discuss various algorithms to predict the same.
- We compared the accuracy of different machine learning algorithms.
- The major advantage of this method is that it is high in interpretability as the user can know which factor influences the price of stock more and by how much.

DISADVANTAGES

- However accuracy would decrease when setting more levels of stock market movement.
- These results indicate that the stock price is unpredictable when traditional classifier is used.
- The disadvantage includes that it is highly limited in its scope. Many predictors cannot be used, which is required to solve the stock price prediction problem.

VI. CONCLUSION AND FUTURE WORKS

CONCLUSION

There are specific problems in the world that push the capabilities of data science and the technologies available in this field to their edge. Among them is the stock market prediction. It is challenging for a person to create such a model, but there are ways through which this art can be learned. One can learn stock market prediction using machine learning projects on public forums such as Kaggle to understand how basic to intermediate level models can be created. This is an ever-evolving problem with new solutions being proposed by every generation of researchers and data scientists. There are still many novel techniques that can be used, and it is encouraged that the readers of this article give it a try, master these techniques and build models to predict the price of stock accurately.

FUTURE SCOPE

Future scope of this project will involve adding more parameters and factors like the financial ratios, multiple instances, etc. The more the parameters are taken into account more will be the accuracy. The algorithms can also be applied for analyzing the contents of public comments and thus determine patterns/relationships between the customer and the corporate employee. The use of traditional algorithms and data mining techniques can also help predict the corporation performance structure as a whole. In the future, we plan to integrate neural network with some other techniques such as genetic algorithm or fuzzy logic. Genetic algorithm can be used to identify optimal network architecture and training parameters. Fuzzy logic provides the ability to account for some uncertainty produced by the neural network predictions. Their uses in conjunction with neural network could provide an improvement for stock market prediction.

REFERENCES

1) Zhong X, Enke D. **Predicting** daily stock **returns** using dimensionality reduction. Expert Systems with Applications, 2017, **Vol.**

2) Hiransha M, Gopalakrishnan **EA,** Menon V K, Soman K P. NSE stock market **forecast** using **deep learning model.** Procedia **Informatique,** 2018, **Vol.**

3) Velay M, Fabrice D. Stock Chart Pattern Deep **Learning Recognition.** arXiv, 2018.

4) Parracho P, Neves R, Horta N. Trading in Financial Markets Using Pattern Recognition Genetic **Algorithm Optimization. Companion t o the** 12th Annual **Conference** Genetic and Evolutionary **Computing,** 2010, pp. 2105-2106.

5) Nesbitt **KV, Ballas** S. **Find** trading patterns in stock market data. IEEE Computer Graphics and Applications, 2004, 24(5), pp. 4555.