

STOCK MARKET PREDICTION USING MACHINE LEARNING

R.Dileep Kumar, Assistant professor, dileep.cse@cmrtc.ac.in

K.Sivadeep,BTech, Department of CSE, 187r1a05e4@cmrtc.ac.in

J.pavan,btech, department of CSE, 187r1a05e9@gmail.com

J.navaneeth,btech, department of cse, 187R1A05E1@cmrtc.ac.in

ABSTRACT: Stock market is place where people buy and sell shares of publicly listed companies. Every buyer and seller try to predict the stock market price movements to get maximum profits and minimum losses. Using cutting edge technology such as AI can improve prediction stock price. The stock market is characterized by extreme fluctuations, non-linearity, and shifts in internal and external environmental variables. Artificial intelligence (AI) techniques can detect such non-linearity, resulting in muchimproved forecast results. The application of Artificial Intelligence (AI) to financial investment is a research area that has attracted extensive research attention since the 1990s, when there was an accelerated technological development and popularization of the personal computer. Since then, countless approaches have been proposed to deal with the problem of price prediction in the stock market. This paper presents a systematic review of the literature on Artificial Intelligence applied to investments in the stock market based on a sample of 2326 papers from the Scopus website between 1995 and 2019. This culminating experience project used artificial intelligence (AI) technology to forecast and analyse the stock market and construct complex nonlinear relationships between the input data and the output data. These papers were divided into four categories:

portfolio optimization, stock market prediction using AI, financial sentiment analysis, and combinations involving two or more approaches. For each category, the initial introductory research to its state-of-the-art applications is described. In addition, an overview of the review leads to the conclusion that this research area is gaining continuous attention and the literature is becoming increasingly specific and thorough.

Keywords: *Artificial intelligence*

1. INTRODUCTION

Researchers have been studying different methods to effectively predict the stock market price. Useful prediction systems allow traders to get better insights about data such as: future trends. Also, investors have a major benefit since the analysis give future conditions of the market. One such method is to use machine learning algorithms for forecasting. This project's objective is to improve the quality of output of stock market predicted by using stock value. The future price of a stock is the main motivation behind the stock price prediction. In various cases like business and industry, environmental science, finance and economics motivation can be useful. The future value of the company's stock can be determining. In this proposed system, we focus on predicting the

stock values using machine learning algorithms like Random Forest and Support Vector Machines. We proposed the system “Stock market price prediction” we have predicted the stock market price using the random forest algorithm. we were able to train the machine from the various data points from the past to make a future prediction. We took data from the previous year stocks to train the model. We majorly used two machine-learning libraries to solve the problem. The first one was numpy, which was used to clean and manipulate the data, and getting it into a form ready for analysis. The other was scikit, which was used for real analysis and prediction. The dataframe features were date and the closing price for a particular day. We used all these features to train the machine on random forest model and predicted the object variable, which is the price for a given day.

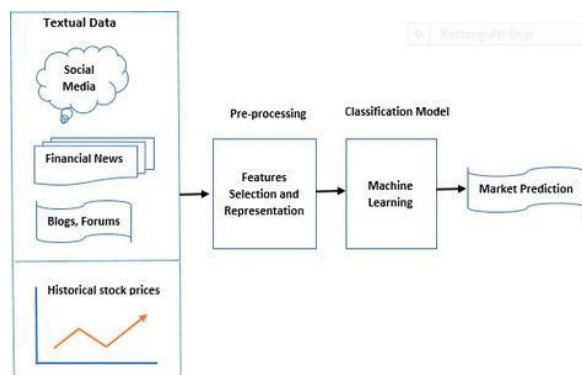


Fig.1: Stock market prediction using text mining

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. The movement in the stock market is usually determined by the sentiments of thousands of investors. Stock market prediction, calls for an ability to predict the effect of recent events on the investors. These events can be political events like a statement by a political leader, a piece of news on scam etc. It can also be an international event like sharp movements in currencies and commodity etc. All these events affect the corporate earnings, which in turn affects the sentiment of investors. It is beyond the scope of almost all investors to correctly and consistently predict these hyperparameters. All these factors make stock price prediction very difficult. Once the right data is collected, it then can be used to train a machine and to generate a predictive result.

2. LITERATURE REVIEW

2.1 Stock Market Prediction Using Machine Learning:

In the finance world stock trading is one of the most important activities. Stock market prediction is an act of trying to determine the future value of a stock other financial instrument traded on a financial exchange. This paper explains the prediction of a stock using Machine Learning. The technical and fundamental or the time series analysis is used by the most of the stockbrokers while making the stock predictions. The programming language is used to predict the stock market using machine learning is Python. In this paper we propose a Machine Learning (ML) approach that will be trained from the available

stocks data and gain intelligence and then uses the acquired knowledge for an accurate prediction. In this context this study uses a machine learning technique called Support Vector Machine (SVM) to predict stock prices for the large and small capitalizations and in the three different markets, employing prices with both daily and up-to-the-minute frequencies.

2.2 Forecasting the Stock Market Index Using Artificial Intelligence Techniques

The research work done by Lufuno Ronald Marwala A dissertation submitted to the Faculty of Engineering and the Built Environment, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science in Engineering. The weak form of Efficient Market hypothesis (EMH) states that it is impossible to forecast the future price of an asset based on the information contained in the historical prices of an asset. This means that the market behaves as a random walk and as a result makes forecasting impossible. Furthermore, financial forecasting is a difficult task due to the intrinsic complexity of the financial system. The objective of this work was to use artificial intelligence (AI) techniques to model and predict the future price of a stock market index. Three artificial intelligence techniques, namely, neural networks (NN), support vector machines and neuro-fuzzy systems are implemented in forecasting the future price of a stock market index based on its historical price information. Artificial intelligence techniques have the ability to take into consideration financial system complexities and they are used as financial time series forecasting tools.

2.3 Indian stock market prediction using artificial neural networks on tick data

The research work done by Dharmaraja Selvamuthu, Vineet Kumar and Abhishek Mishra Department of

Mathematics, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India. A stock market is a platform for trading of a company's stocks and derivatives at an agreed price. Supply and demand of shares drive the stock market. In any country stock market is one of the most emerging sectors. Nowadays, many people are indirectly or directly related to this sector. Therefore, it becomes essential to know about market trends. Thus, with the development of the stock 5 market, people are interested in forecasting stock price. But, due to dynamic nature and liable to quick changes in stock price, prediction of the stock price becomes a challenging task. Stock m Prior work has proposed effective methods to learn event representations that can capture syntactic and semantic information over text corpus, demonstrating their effectiveness for downstream tasks such as script event prediction. On the other hand, events extracted from raw texts lacks of common-sense knowledge, such as the intents and emotions of the event participants, which are useful for distinguishing event pairs when there are only subtle differences in their surface realizations. To address this issue, this paper proposes to leverage external common-sense knowledge about the intent and sentiment of the event.

2.4 The Stock Market and Investment

The research work done by Manh Ha Duong Boriss Siliverstovs. Investigating the relation between equity prices and aggregate investment in major European countries including France, Germany, Italy, the Netherlands and the United Kingdom. Increasing integration of European financial markets is likely to result in even stronger correlation between equity prices in different European countries. This process can also lead to convergence in economic development across European countries if

developments in stock markets influence real economic components, such as investment and consumption. Indeed, our vector autoregressive models suggest that the positive correlation between changes equity prices and investment is, in general, significant. Hence, 6 monetary authorities should monitor reactions of share prices to monetary policy and their effects on the business cycle.

2.5 Data Compression Techniques for Stock Market Prediction.

Advanced data compression approaches are presented here for predicting stock market behaviour using generally established financial market models. Technical analysis, portfolio theory, and nonlinear market models can all benefit from the techniques. Lossy and lossless compression approaches, according to the authors, are ideally suited for anticipating stock values as well as market patterns like strong trends and substantial modifications. They also discuss unique uses of multispectral compression techniques in portfolio theory, stock correlation, interest rate effects, transaction costs, and taxes.

2.6 Application of K-Nearest Neighbor (KNN) Approach for Predicting Economic Events:

The k-Nearest Neighbor categorization algorithm was investigated for economic forecasting in this work. Finance distress prediction models have been one of the most appealing fields of financial study due to the implications of organisations' financial difficulty on stakeholders. The number of insolvent enterprises has increased in recent years, following the global financial crisis. Because financial difficulty is the first step of bankruptcy, researchers and economic and financial organizations have focused too much effort on utilizing financial ratios to anticipate

financial distress. Although studies on forecasting financial distress in Iranian enterprises have expanded in recent years, the majority of efforts have relied on traditional statistical approaches, with only a few studies using nonparametric methods. Recent research shows that this strategy is more capable than others.

2.7 Validation of nearest neighbor classifiers

This study shows how to generate error bounds for k-nearest neighbour classifiers that are probably approximately right. The method uses some training data as a validation set to limit the holdout classifier's error rate, which is based on the remaining training data. The approach then uses the validation set to constrain the difference in error rates between the holdout classifier and the all-data classifier. The result is a bound on the classifier's out-of-sample error rate based on all training data.

3. IMPLEMENTATION

The existing system fails when there are rare outcomes or predictors, as the algorithm is based on bootstrap sampling.

- The previous results indicate that the stock price is unpredictable when the traditional classifier is used.
- The existence system reported highly predictive values, by selecting an appropriate time period for their experiment to obtain highly predictive scores.
- The existing system does not perform well when there is a change in the operating environment.
- It doesn't focus on external events in the environment, like news events or social media.

- It exploits only one data source, thus highly biased.

In this proposed system, we focus on predicting the stock values using machine learning algorithms like Random Forest and Support Vector Machines. We proposed the system “Stock market price prediction” we have predicted the stock market price using the random forest algorithm. In this proposed system, we were able to train the machine from the various data points from the past to make a future prediction. We took data from the previous year stocks to train the model. We majorly used two machine-learning libraries to solve the problem. The first one was numpy, which was used to clean and manipulate the data, and getting it into a form ready for analysis. The other was scikit, which was used for real analysis and prediction. The data set we used was from the previous years stock markets collected from the public database available online, 80 % of data was used to train the machine and the rest 20 % to test the data. The basic approach of the supervised learning model is to learn the patterns and relationships in the data from the training set and then reproduce them for the test data. We used the python pandas library for data processing which combined different datasets into a data frame. The tuned up dataframe allowed us to prepare the data for feature extraction. The dataframe features were date and the closing price for a particular day. We used all these features to train the machine on random forest model and predicted the object variable, which is the price for a given day.

ADVANTAGES:

- The investors are more confident to invest on stocks according to the predictions.
- The using of different algorithms according to the events to give the accurate prediction.

- Multiple algorithms will be applied at the moment of event.

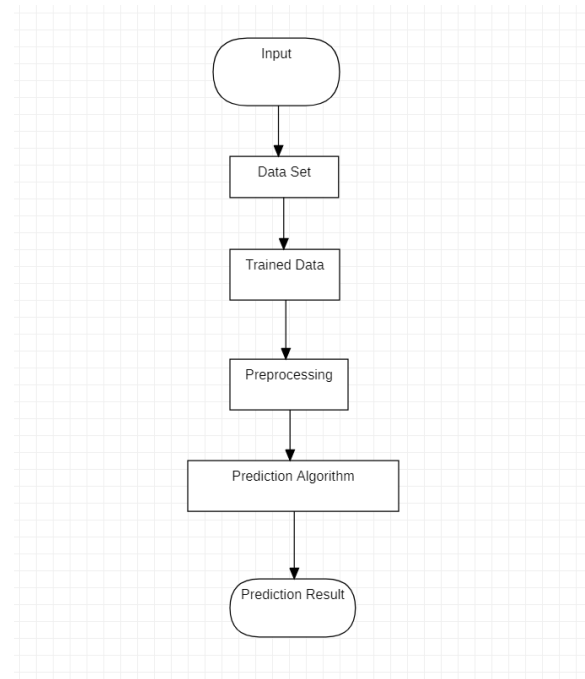


Fig.2: System architecture

4. ALGORITHMS

Machine Learning:

Machine learning and deep learning have been widely embraced, and even more widely misunderstood. In this article, I'd like to step back and explain both machine learning and deep learning in basic terms, discuss some of the most common machine learning algorithms, and explain how those algorithms relate to the other pieces of the puzzle of creating predictive models from historical data. Machine learning is a class of methods for automatically creating models from data. Machine learning algorithms are the engines of machine learning, meaning it is the algorithms that turn a data set into a model. Which kind of algorithm works best (supervised, unsupervised, classification, regression, etc.) depends

on the kind of problem you're solving, the computing resources available, and the nature of the data.

A machine learning algorithm is the method by which the AI system conducts its task, generally predicting output values from given input data. The two main processes of machine learning algorithms are classification and regression.

There are four types of machine learning algorithms: supervised, semi-supervised, unsupervised and reinforcement. Simply put, machine learning allows the user to feed a computer algorithm an immense amount of data and have the computer analyze and make data-driven recommendations and decisions based on only the input data. Supervised machine learning models are trained with labeled data sets, which allow the models to learn and grow more accurate over time. For example, an algorithm would be trained with pictures of dogs and other things, all labeled by humans, and the machine would learn ways to identify pictures of dogs on its own.

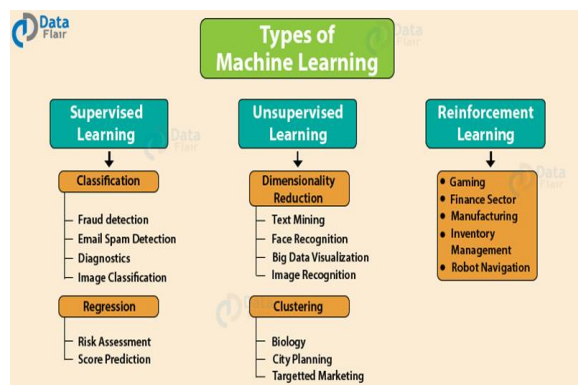


Fig.3: Machine learning models

The different types of machine learning paradigms. We went through supervised, unsupervised and reinforcement learning. We also discussed the several algorithms that are part of these three categories.

Then, we went through the various real-life applications of these algorithms.

5. EXPERIMENTAL RESULTS

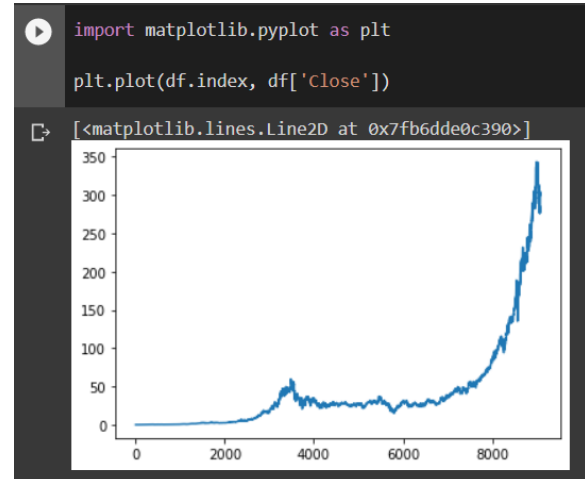


Fig.4: Graph of the given data

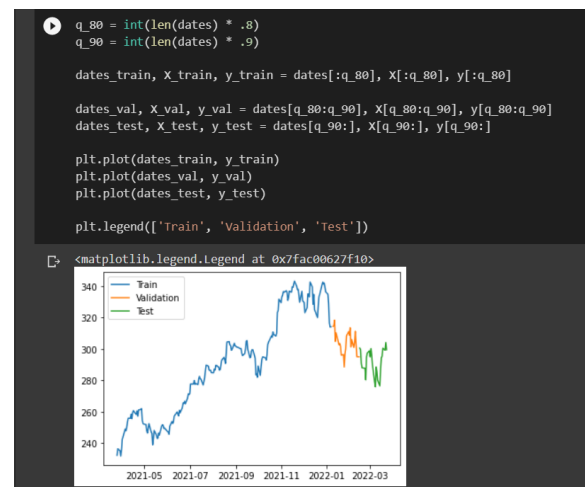


Fig.5: Training the data

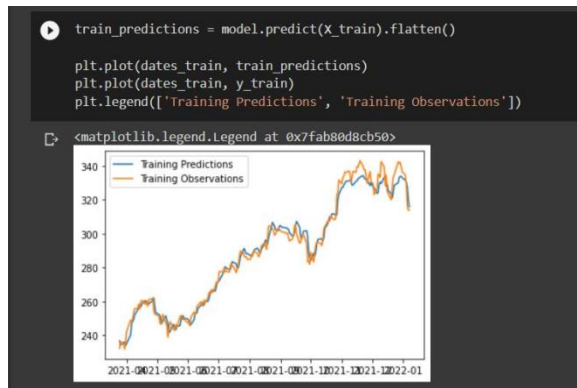


Fig.6: Training prediction

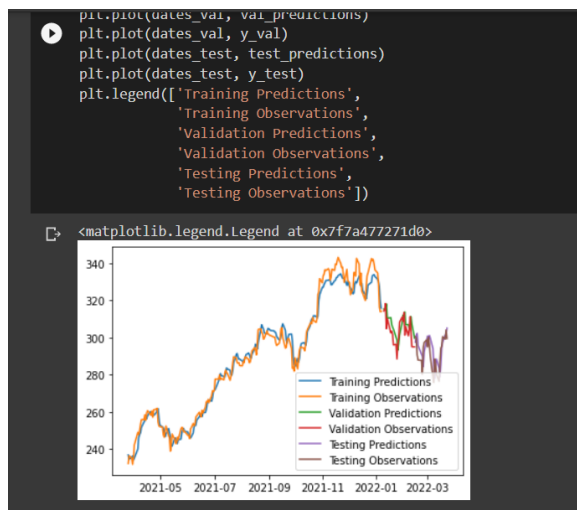


Fig.7: visualization

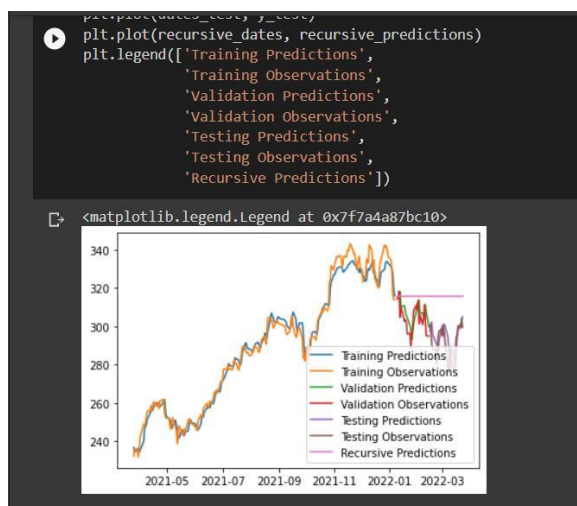


Fig.8: Predicted graph

6. CONCLUSION

By estimating the parameters of accuracy, it was found that the most appropriate method for stock market prediction is to use multiple algorithms such as RNN, LSM, LSTM and combine their results for predictions, and also considering all the factors affecting the stock market price such as news sentiment, currency, commodity prices and other international stock exchange market data. High accuracy was achieved by doing all this. Considering other factors affecting the stock price improves accuracy of prediction. Trying to predict the stock market only when its more predictable and avoiding when market is more uncertain gives greater profits.

7. FUTURE SCOPE

We want to extend this application for predicting cryptocurrency trading. We want to add sentiment analysis for better analysis.

REFERENCES

- [1] Stock Price Prediction Using LSTM on Indian Share Market by Achyut Ghosh, Soumik Bose¹, Giridhar Maji, Narayan C. Debnath, Soumya Sen.
- [2] S. Selvin, R. Vinayakumar, E. A. Gopalkrishnan, V. K. Menon and K. P. Soman, "S "Stock price prediction using LSTM, RNN and CNN-sliding window model," in I International Conference on Advances in Computing, Communications and Informatics, 2017.
- [3] Murtaza Roondiwala, Harshal Patel, Shraddha Varma, "Predicting Stock Prices Using LSTM" in Undergraduate Engineering Students, Department of InformationTechnology, Mumbai University, 2015.

[4] Xiongwen Pang, Yanqiang Zhou, Pan Wang, Weiwei Lin, "An innovative neural network approach for stock market prediction", 2018

[5] Ishita Parmar, Navanshu Agarwal, Sheirsh Saxena, Ridam Arora, Shikhin Gupta, H Himanshu Dhiman, Lokesh Chouhan Department of Computer Science and Engineering National Institute of Technology, Hamirpur – 177005, INDIA - Stock Market Prediction Using Machine Learning.

[6] Pranav Bhat Electronics and Telecommunication Department, Maharashtra Institute of Technology, Pune. Savitribai Phule Pune University - A Machine Learning Model for Stock Market Prediction.

[7] Anurag Sinha Department of computer science, Student, Amity University Jharkhand Ranchi, Jharkhand (India), 834001 - Stock Market Prediction Using Machine Learning.

[8] V Kranthi Sai Reddy Student, ECM, Sreenidhi Institute of Science and Technology, H Hyderabad, India - Stock Market Prediction Using Machine Learning.

[9] Stock Price Prediction Using LSTM on Indian Share Market by Achyut Ghosh, Soumik Bose1, Giridhar Maji, Narayan C. Debnath, Soumya Sen

[10] S. Selvin, R. Vinayakumar, E. A. Gopalkrishnan, V. K. Menon and K. P. Soman, "Stock price prediction using LSTM, RNN and CNN-sliding window model," in International Conference on Advances in Computing, Communications and Informatics, 2017.

[11] Murtaza Roondiwala, Harshal Patel, Shraddha Varma, "Predicting Stock Prices Using LSTM" in

Undergraduate Engineering Students, Department of Information Technology, Mumbai University, 2015.

[12] Xiongwen Pang, Yanqiang Zhou, Pan Wang, Weiwei Lin, "An innovative neural network approach for stock market prediction", 2018

[13] Ishita Parmar, Navanshu Agarwal, Sheirsh Saxena, Ridam Arora, Shikhin Gupta, Himanshu Dhiman, Lokesh Chouhan Department of Computer Science and Engineering National Institute of Technology, Hamirpur – 177005, INDIA - Stock Market Prediction Using Machine Learning.

[14] Pranav Bhat Electronics and Telecommunication Department, Maharashtra Institute of Technology, Pune. Savitribai Phule Pune University - A Machine Learning Model for Stock Market Prediction.

[15] Anurag Sinha Department of computer science, Student, Amity University Jharkhand Ranchi, Jharkhand (India), 834001 - Stock Market Prediction Using Machine Learning.