iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VI **Month of publication:** June 2021

DOI: https://doi.org/10.22214/ijraset.2021.35755

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

Stock Market Prediction using Machine Learning

Prof. Gowrishankar B S[1], Pooja V[2], Girichandra R Karaveermath[3], Guruprasad M[4], Swaroop R Hosamani[5]
[1, 2, 3, 4, 5]*Dept. of Information Science, VVCE, Mysuru*

Abstract: ***Stock market is one of the most complicated and sophisticated ways to do business. Small ownerships, brokerage corporations, banking sectors, all depend on this very body to make revenue and divide risks; a very complicated model. However, this paper proposes to use machine learning algorithms to predict the future stock price for exchange by using pre-existing algorithms to help make this unpredictable format of business a little more predictable. The use of machine learning which makes predictions based on the values of current stock market indices by training on their previous values. Machine learning itself employs different models to make prediction easier and authentic. The data has to be cleansed before it can be used for predictions. This paper focuses on categorizing various methods used for predictive analytics in different domains to date, their shortcomings.***

Keywords: Data Analysis, Machine Learning, Predictive Analysis, Stock Prediction, KNN, SVM.

I. INTRODUCTION

Stock market predicting has been a prominent issue for discussion, because to its potential for strong gains. Predicting the stock market is a challenging feat, owing to the near similarity of a stock time series to a normal distribution. Economists, politicians, scholars, and market makers seem to have become more interested in predicting markets. The proposed paper aims to study and improve supervised learning algorithms for stock price prediction. Stock market quantitative traders purchase stocks, commodities, and shares at a reasonable price and then sell them at a hefty premium. Stock price forecasting is an important task for investment and financial decision making. Forecasting stock price or financial markets has been one of the biggest challenges to the AI community. Many of the algorithms already exist but the accuracy is not to the mark. Hence, understanding the need of the hour, we use KNN and SVM to forecast the stock price accurately. There are many algorithms that people have used say, LSTM, Artificial Neural Network, Regression, K-means, and so on.

II. LITERATURE SURVEY

Ashish Sharma. et al. [1] aim to test a survey of well-known efficient regression approaches based on market analysis to anticipate stock prices. The results of a multiple regression method might well be improved in the future by employing a bigger set of possibilities. Han Lock Siew. et al. [2] aim to nalyze the dataset's features, which are variables made up of statistical ratios. To generate the dataset for this study, all numerical data is transformed into ordinal or numbered values. The ordinal data is then evaluated using regression-based classifiers from WEKA as predictive analytics. The results are then compared and evaluated.

Ishita Parmar. et al. [3] aim to predict the future price of a corporation's economic stocks Machine learning, which produces predictions based on the values of present financial markets index by training on their previous values, is a recent trend in stock market prediction technology. Machine learning involves a variety of models to generate precise forecasts.

Kunal Pahwa. et al. [4] aim to test to aid make this unpredictable format of business a lot better foreseeable by combining open-source tools and pre-existing algorithms The methods used are data analysis, basic analysis, implementation, and linear regression.

Mehak Usman. et al. [5] aim to test Single Layer Perceptron (SLP), Multi-Layer Perceptron (MLP), Radial Basis Function (RBF), and Support Vector Machine (SVM) machine learning techniques are examined. Every one of these features are studied separately as well. When compared to other techniques, the MLP algorithm performed the best. The oil rate feature was determined to be the most significant in forecasting market performance.

Neelima Budhani. et al. [6] aim to test Artificial Neural Networks (ANN) are a kind of Artificial Intelligence (AI) technology (AI). It's a technique of recognizing samples (patterns) and creating a data model. The potential for particular problems with increasing learning and input output mapping are important characteristics of ANN. A neural network is a frequently used approach for categorizing unidentified, unobserved samples into input values that can be used to predict the stock market. To make forecasts, they utilized efforts to adopt neural networks with a backpropagation training algorithm.

Samuel Olusegun Ojo. et al. [7] aim to test adopts a stacked Long Short Term Memory network model for predicting stock market behaviour. Results obtained show that by making use of a stacked Long Short Term Memory network model, future stock market behavior can be predicted. The data used is composed of historic stock market data from the American Stock Exchange, NASDAQ Composite (IXIC).

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

III. METHODOLOGY

A. Proposed Solution

The main aim of this work is to develop a system using advanced algorithms that is used to predict stock market values in ease. As an alternative, machine learning methods such as Support vector machines (SVMs) and K Nearest Neighbor (KNN) have been established. Their practical success can be ascribed to VC theory's strong theoretical foundations. The project develops stock market analysis and distribution recommendation. It can be used to buy or sell stocks.

B. Data Collection and Data Preprocessing

Data Gathering/Collection is termed as accumulating data/information from various sources of interest in a systematic fashion or in a standard way to evaluate outcomes like testing hypotheses and formulating them to achieve a desired output. We use (Application Programming Interface) API 's to gather or collect information about stocks. Extracting the correlated data of multiple stocks and applying this information to the further process of prediction of the fluctuation patterns of a single stock. Which then leads us to the data preprocessing phase. The historical stock data is obtained from https://www.quandl.com/data/NSE and this data is utilized for the prediction of future stock prices.

Data preprocessing is a technique for turning impure data into a clean data set. Here, all the collected raw data related to the stocks which are obtained in scrambled and scattered form will be classified accordingly to a well-organized datasets. Which are then, are ready to feed into the machine learning algorithms for further processing of data. The pre-processing step includes

1) Data discretization.
2) Data transformation.
3) Data cleaning.
4) Data integration.

Attribute	Description
Date	Day of trading.
Open	Price of the first stock exchange.
High	Highest price of the stock on a particular day.
Low	Lowest price of the stock on a particular day.
Last	Closing price of the stock on the previous day.
Close	Closing price of the stock on that particular day.
Total trade	Volumes traded on that particular date.
Turnover	Turnover of a company.

Table 1: The preprocessed dataset attributes.

Table 1 gives the description of the attributes present in the dataset. The data from the dataset includes many instances and has 8 attributes as inputs and a turnover.

©IJRASET: All Rights are Reserved

C. Prediction Modeling

1) **Support Vector Machine (SVM)** is a supervised machine learning algorithm gaining knowledge of a set of rules which can be used for regression demanding situations. SVM essentially finds a hyper-plane that divides the various sorts of data. This hyperplane is nothing but a line in a two-dimensional space. Each data item in the dataset is charted in an N-dimensional space in SVM, where N is the number of features/attributes in the data. The very next task is to decide the best hyperplane for dividing the data. Simply put, support vectors are the coordinates of each unique observation. The SVM classifier is a classification frontier that performs well.
2) **The KNN algorithm** is a sort of supervised machine learning method that may be used to solve both classification and regression prediction problems. KNN is defined by the following two properties: KNN is a lazy learning algorithm since it does not have a separate training phase and instead trains with all of the data. Algorithm for non-parametric learning

The KNN method predicts the values of new data points based on how nearly they match the points in the training set, meaning that a number will be allocated to the new observation based on how nearly it matches the points in the training set. The steps below can help us in understanding how it actually does the work:

a) *Step 1:* We need a dataset to perform any technique. As an outcome, we must load both training and testing data at the first stage of KNN. The necessary stock value is represented by test data sets, meanwhile the trained stock market values are provided by train data sets.
b) *Step 2:* We should select the value of K, i.e., the data points that are nearest to that too. Any integer point could be used as K. K, which serves as a centroid point, is chosen from a test dataset.
c) *Step 3:* For each point in the test data- Using the Euclidean, Manhattan, or Hamming distance techniques, calculate the distance between test data and each picture of training examples. Based on the distance value, sort them in increasing order. Then after, the top K values from the input order will be chosen. This will allocate a class to the sample point based on the most frequent class of these pictures.

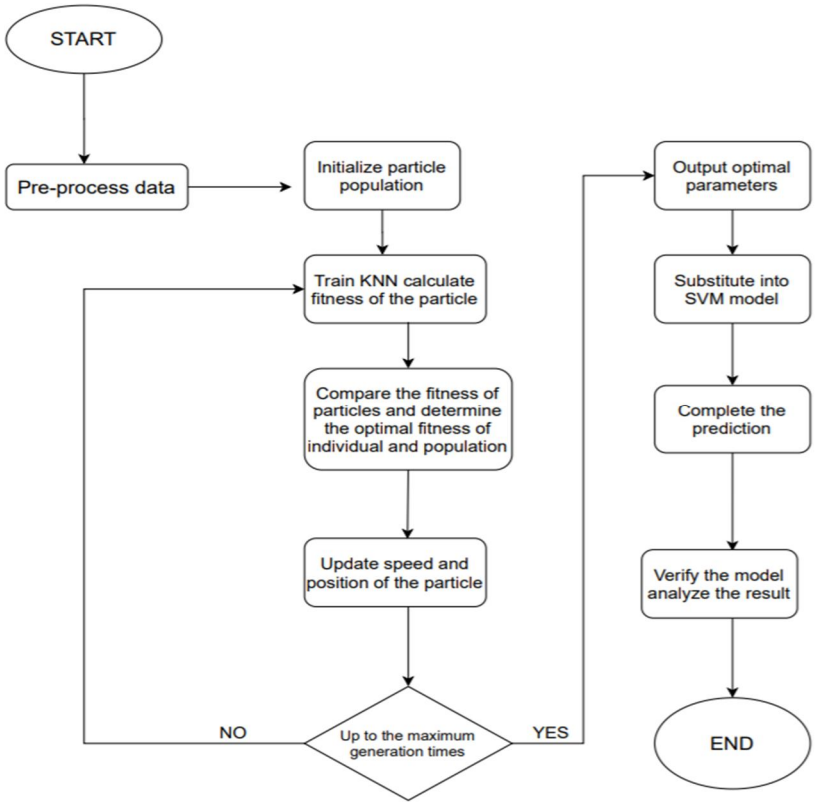


Figure 1: System Architecture

Figure 1 represents the system architecture. Data is first gathered and preprocessed followed by predictive analysis and visualization and then the trained data is fed into the algorithms which then gives the required results.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

IV. COMPARISON

Several analyses have been carried out for stock market prediction using algorithms such as LSTM (Long Short Term Memory), Regression models, Random Forest, ANN (Artificial Neural Network) and so on.

A. Long Short Term Memory (LSTM)

LSTM is the advanced version of Recurrent-Neural Networks (RNN) where the information belonging to the previous state persists. This indicates that the interval of information is relatively smaller than that to LSTM [3]. The LSTM is good at forecasting the following day's values; in fact, the next day's prediction is quite close to the previous day's actual value. These findings show that LSTM cannot forecast the stock market's value for the following day.

B. Random Forest

Random Forest can be used for regression and classification problems as well. Although the main limitation of random forest is that it is enormous and has a huge number of trees which will make the algorithm too slow and futile in the real-world scenario.

C. Artificial Neural Network (ANN)

The Artificial Neural Network (ANN) is the most often utilized method. Due to the high number of parameters to correct and limited prior user information about the significance of inputs in the studied problem, Artificial Neural Networks are prone to overfitting.

D. Linear Regression

Linear regression infers linear relationships between both input and output variables. This deteriorates to fit complex datasets properly. In real world scenarios, these relationship between variables of training and testing datasets is not linear and this leads to not fitting a straight line properly during the course of turnover.

Algorithms	Accuracy
LSTM	53.6%
Random Forest	54.12 %
Neural Network	88 %
Linear Regression	56.11%

Table 2: Accuracies of different machine learning algorithm

Table 2 shows the prediction accuracies of the algorithms after training and testing. In the case of LSTM, it gives an accuracy of 53.6%. While the highest accuracy was given by the Neural network of 88%. Although SVM gives an accuracy of 96.15% which is the highest and gives almost accurate results

V. CONCLUSION

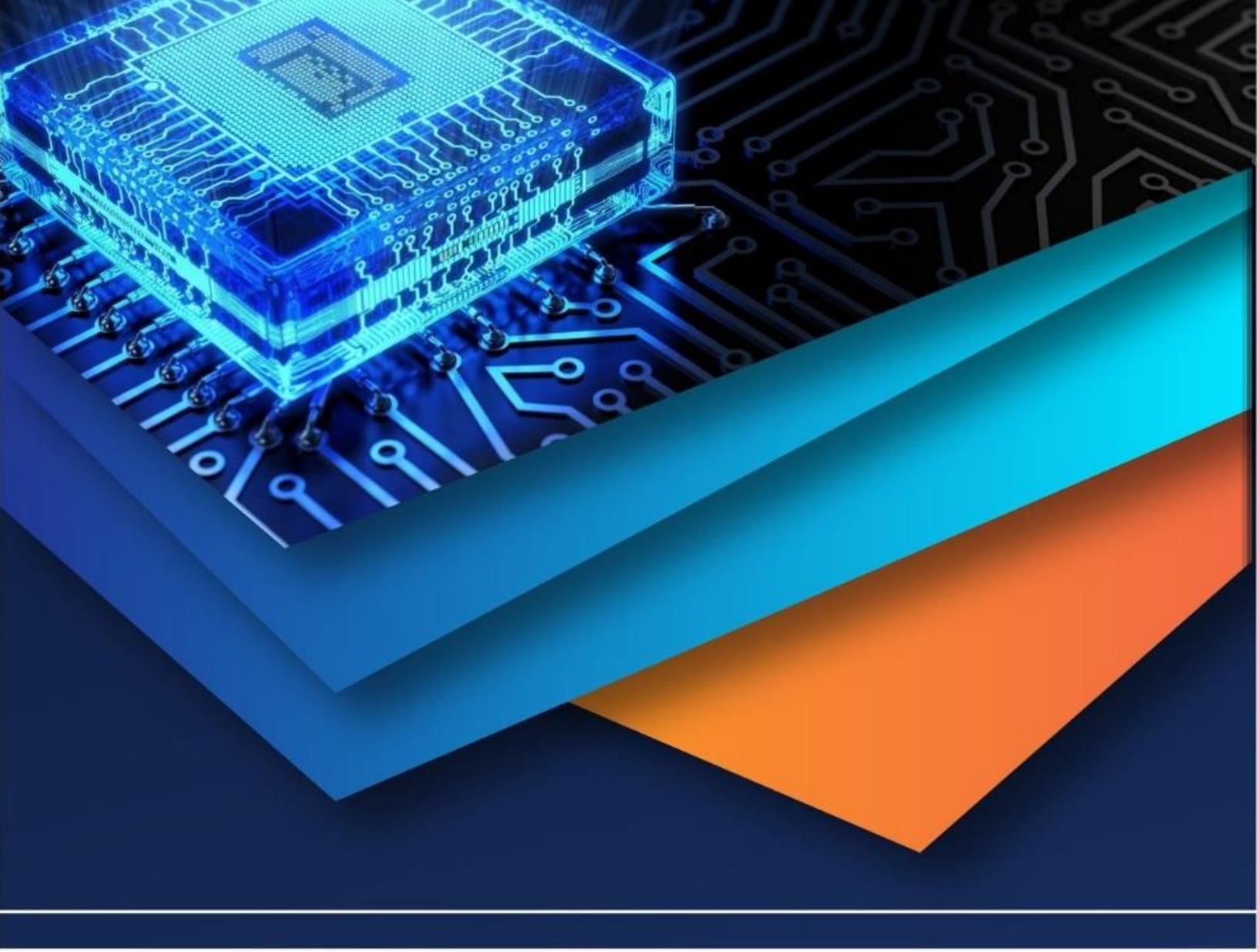
Predicting stock market returns is a challenging task due to consistently changing stock values which are dependent on multiple parameters which form complex patterns. The historical dataset available on the company's website consists of only a few features like high, low, open, close, adjacent close value of stock prices, volumes of the shares traded etc., which are not sufficient enough. To obtain higher accuracy in predicted price value will be working on all possible factors using Machine Learning.

©IJRASET: All Rights are Reserved

REFERENCES

[1] Ashish Sharma, Dinesh Bhuriya, Upendra Singh, "Survey of Stock Market Prediction Using Machine Learning Approach", IEEE, 2017.
[2] Han Lock Siew, Md Jan Nordin, "Regression Techniques for the Prediction of Stock Price Trend", IEEE, 2016.
[3] Ishita Parmar, Navanshu Agarwal, Sheirsh Saxena, Ridam Arora, Shikhin Gupta, Himanshu Dhiman, Lokesh Chouhan, "Stock Market Prediction Using Machine Learning", IEEE, 2018.
[4] Kunal Pahwa, Neha Agarwal, "Stock Market Analysis using Supervised Machine Learning", IEEE, 2019
[5] Meghna Misra, Ajay Prakash Yadav, and Harkiran Kaur, "Stock Market Prediction using Machine Learning Algorithms: A Classification Study", IEEE, 2018.
[6] Mehak Usmani , Syed Hasan Adil , Kamran Raza, and Syed Saad Azhar Ali, "Stock Market Prediction Using Machine Learning Techniques", IEEE, 2016.
[7] Neelima Budhani, Dr. C. K. Jha, Sandeep K Budhani, "Prediction of Stock Market Using Artificial Neural Network", IEEE, 2014.
[8] Samuel Olusegun Ojo, Pius Adewale Owolawi, Maredi Mphahlele, and Juliana Adeola Adisa, "Stock Market Behaviour Prediction using Stacked LSTM Networks", IEEE, 2017.
[9] https://www.tutorialspoint.com/machine_learning/index.htm
[10] Machine Learning: What it is and why it matters | SAS

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)