International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 05 | May 2021 www.irjet.net p-ISSN: 2395-0072

Stock Market Analysis

Nikita Bhilare[1], Meet Mota[2], Rinkal Shah[3] , Dhanashree Toradmalle[4]

[1]Student, Department of IT Engineering, Shah and Anchor Kutchhi Engineering College, Maharashtra, India.
[2]Student, Department of IT Engineering, Shah and Anchor Kutchhi Engineering College, Maharashtra, India.
[3]Student, Department of IT Engineering, Shah and Anchor Kutchhi Engineering College, Maharashtra, India.
[4]Professor, Department of IT Engineering, Shah and Anchor Kutchhi Engineering College, Maharashtra, India.

Abstract - *In today's fast moving, competitive world, a complimentary source of income for any individual, has become a necessity. In such adverse conditions, stock market trading proves to be one of the most appropriate, smart and reliable sources of income. One cannot trade without having a proper analysis and the background study of the respective stock. On the contrary, study of all stocks simultaneously becomes quite tedious and the individual might lose on to some great opportunities to trade. Choosing stocks for trading is really a tough task. Here in this paper, we aim to build a system that examines the certainty of financial movement along with some technical indicators and sentiment analysis. Investment timing plays a crucial role and that's only possible only when one has an overview of the stock and the respective sector.*

***Key Words*: complimentary, reliable, investment, predictability, indicators, sentiment, sector.**

1.INTRODUCTION

Financial analysts initially find it very difficult to figure out which stocks would be apt for investments and which would help them yield huge profits. Constant analysis of this information is tough manually. Hence, the demand for having a automated system increases. Combination of machine learning algorithms with the technical study of stocks, makes the system more efficient and reliable.

Stock returns or exchange prediction is a crucial monetary aspect. Different monetary have aspects such as rates of interests and exchange, The samples of info include economic variables like interest rates and exchange rates, information based on business factors such as production value and company oriented data such as dividend yields. All of these account to focus on the stock market returns and profit yields.

2.EXISTING SYSTEMS

Recent researches use input file from varied sources and multiple forms. Some systems use historical stock information, some use money news articles, some use professional reviews whereas some use a hybrid system that takes multiple inputs to predict the market. Also, a large variety of machine learning algorithms area unit accessible which will be the accustomed style the system. Some models perform mathematical synthesis on historic data for prediction whereas some perform sentiment analysis on monetary news articles and highly recommended reviews for prediction. However, thanks to the volatility of the securities market, no system includes an excellent or correct prediction.

3.PROPOSED SYSTEMS

We are trying to come up with a model that focuses on depicting the technical and analytical study of certain stocks along with an extent of algorithmic trading in it. The model will have a certain set of technical tools that will help to the user to identify which points were suitable for buying/selling in the stock and what trend it followed. It will have a section of algorithmic trading and stock market sentiment analysis that will project the status of that particular stock. Technical analysis of historical data of stocks will help the user to predict or analyse the direction or scope of the future of that particular stock.

3.1 Why our system is better?

Its a combined model that will help to determine the status of the stock.

Its user friendly and highly efficient.

It not just predicts the futuristic price, but also, shows the current status of stock using sentiment analysis.

Unlike any other system, it focuses on depicting the technical study using technical indicators.

4.LITERATURE REVIEW

The paper 'Predicting the effects of news sentiments on stock market' [1], proposes the Development of a sentiment analysis wordbook for the monetary sector. The sentiment scores projected through the analysis of these articles indicate a strong signal for the prediction of short run trades. The model is in a position to realize an accuracy of 270.59%

The paper [2], "Forecasting directional movements of stock prices for intraday trading using LSTM and random forests" Introduce a multi-feature setting having solely of the returns with reference to the closing costs, however conjointly with reference to the gap costs and intraday returns. use each random forest on one hand and LSTM networks on the opposite hand as coaching methodology and apply an equivalent mercantilism strategy as in

Krauss et al. (2017) and Fischer & Krauss (2018). The Daily returns for LSTM are 0.64% and for Random Forest are 0.54%. Share of Positive returns: LSTM: 69.67%, Random forests: 65.85%, Therefore they conclude that LSTM outperforms random forests.

The paper [3], "Machine Learning Stock Market Prediction Studies: Review and Research Directions" Aims to see the value or the trend of a definite stock for the approaching end-of-day. FFNN is employed for classification, regression, and pattern recognition issues. RNN is employed for sequencing information. SVM model, at the side of completely different kernel algorithms square measure, used. SVR in predicting the direction of today's shut worth with reference to yesterday's shut worth is employed. Directional Accuracies: SVM, had the foremost for all four thought of stocks as compared with DNN, RNN, SVR. Hence this paper concludes that, SVM outperforms SVR, DNN, and RNN.

The paper [4], "Stock Market Prediction Using Machine Learning" uses a model based on two things: Regression and LSTM. Factors thought of are open, close, low, high, and volume. Regression-based mostly Model & gradient descent algorithmic rule is employed for predicting continuous values through freelance values. A successive model is formed by stacking 2 LSTM layers on the prime of every alternative with an output worth 256. Regression model: The R-square confidence check resulted in an exceeding confidence score of 0.86625. LSTM Model: Train Score = 0.00106 MSE (0.03 RMSE) Check Score of 0.00875 MSE (0.09 RMSE). Thus, this paper concludes that LSTM Model offered a lot of accuracy than the Regression-based mostly Model.

The paper [5], "Stock Market Prediction Using Machine Learning" Propose a Machine Learning (ML) approach which will be trained from the on the market stocks knowledge associated gain intelligence and so uses the non inheritable information for a correct prediction. Therefore, the YALE data processing Environments used for completing the experiments. Prediction of the securities market is completed by SVM and Radial Basis operate (RBF). SVM doesn't provide a drawback of over fitting. The model yields a higher profit as compared to the predictable mark.

ALGORITHMS

Different supervised and unsupervised algorithms are available for carrying out different projects in the field of ML and AI. But their accuracy and performance depends upon different elements. Similarly, for projects based on prediction, not all algorithms turn out to be super accurate. So we've done an analysis of algorithms ourselves, to decide which algorithm among all would be most accurate and effective.

Long Short-Term Memory (LSTM):

LSTM's are very powerful in predicting the future price. LSTM stores past prices efficiently. In this model predicting the future price is important, hence, having a set of previous prices is very necessary.

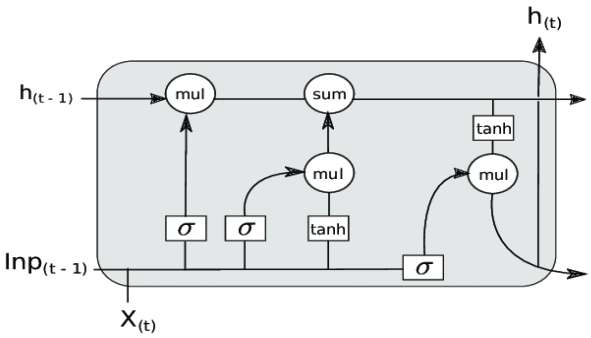

Fig-1: LSTM

Linear Regression (LR):

Linear regression is the analysis of two separate variables to define a single relationship and is a useful measure for technical and quantitative analysis in financial markets.

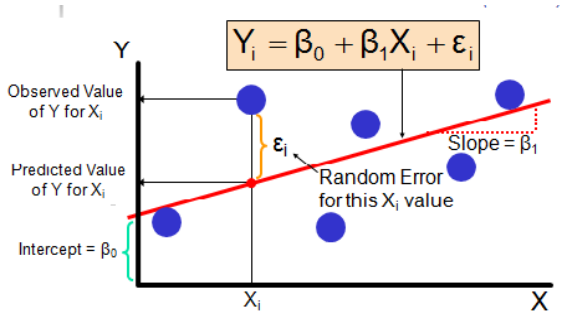

Fig-2: Linear Regression

K-Nearest Neighbours:

K-NN algorithm assumes the similarity between the new case/data and available cases and put the new case into the category that is most similar to the available categories.

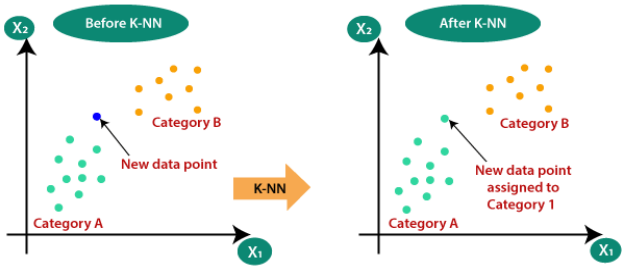

Fig-3: K-Nearest Neighbours

Auto Arima:

ARIMA stands for Auto-Regressive Integrated Moving Averages. ARIMA models work on the following assumptions – The data series is stationary, which means that the mean and variance should not vary with time.

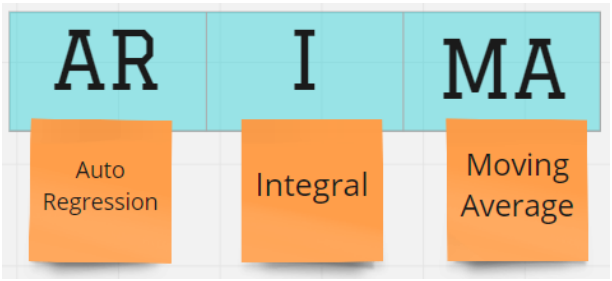

Fig-4: Auto Arima

5.1 Analysis Of Algorithms

Here, in this paper, we tried to compare different, efficient machine learning algorithms using datasets of a particular stock having historical data set. We conducted experiments with several algorithms on chosen data set and determined the mean square error to foretell accuracy . Algorithms viz. LSTM, Linear regression, k-Nearest neighbor, Auto Arima and Prophet are used. Python libraries such as pandas, numpy are used to load the dataset and to perform numerical calculations we utilized sklearn to model different machine learning algorithms. The results look as follows:

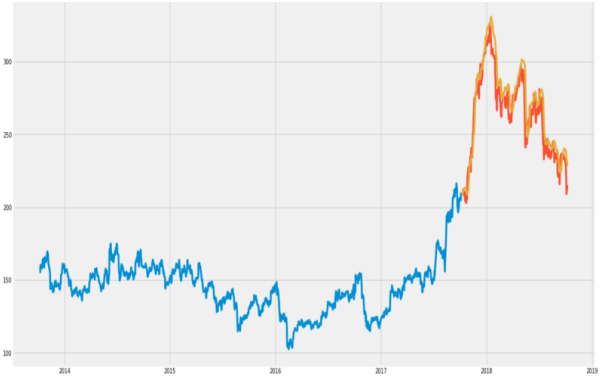

Chart -1: LSTM Graph

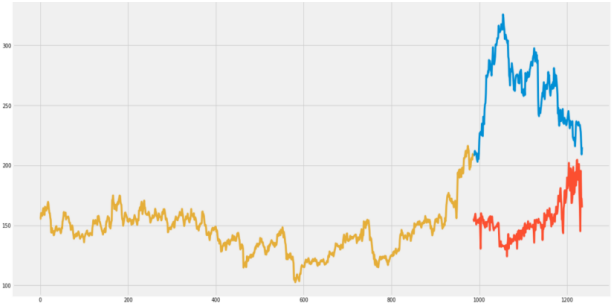

Chart -2: KNN Graph

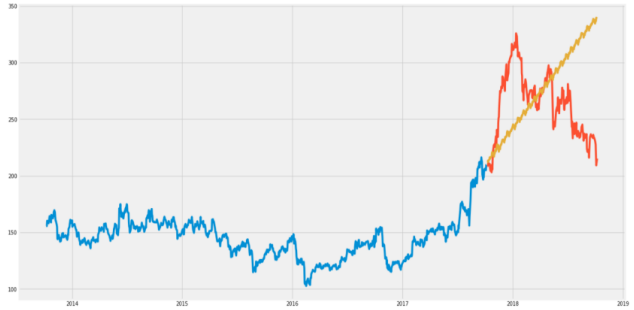

Chart -3: Auto Arima

Chart -4: Prophet Graph

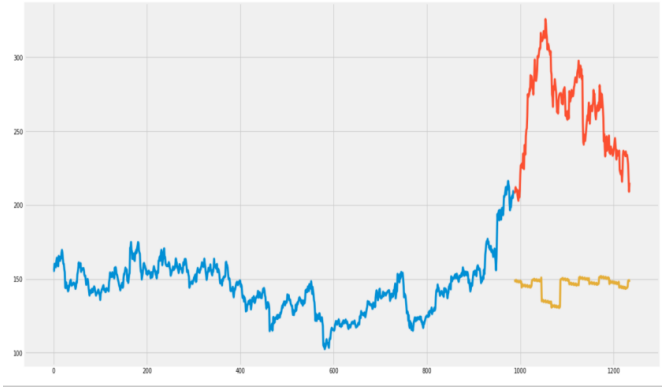

Chart -5: Linear Regression Graph

Table -1: Mean Square Error for Data set

Algorithms	Mean square error
Long Short-Term Memory	11.967
Linear Regression	121.16
Auto Arima	54.54
k-Near Neighbors	115.17
Prophet	57.31

The above analysis makes us analyse that LSTM proves to be more accurate with prediction of actual value and predictions made. Also, LSTM has the least mean square error. Hence, LSTM outer performs among all the algorithms considered.

SYSTEM OVERVIEW

Our methodology comprises six steps:

Collection of raw, historical data relevant to the aim of the project.

Pre-process the data in a such a way that only necessary parameters are considered w.r.t to the problem statement.

Distribute data into study sections and allocate it to the training and the testing data sections for further analysis.

Study the algorithm and apply it for the considered data, analyse the predictions and focus on the desired output.

Sentiment analysis sector is used for determining the positive and negative scores from live news of the particular stock.

Merge the outcomes, evaluate the results and predict the output of the stock via developed system.

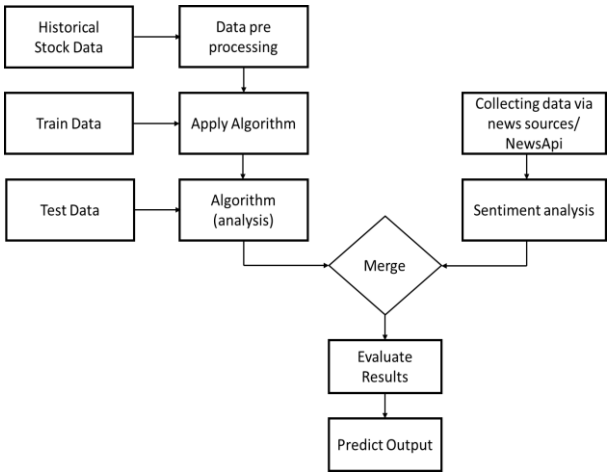

Fig-5: Flow of the proposed system

6.1 Historical Stock Data:

The initial phase of our project, collection of historical stock data, is the base of the entire system.

Larger and better the collected data, higher would be the efficiency and accurate would be the predicted output.

The historic data would be relevant to our stocks that we're planning to show and analysis of that data would be done initially, before using it directly with the planned algorithms.

AlphaVantage provides with NSE stock data in json objects, csv files and has larger datasets spread over years.

Through our literature review, we figured out that the datasets have also been obtained via kaggle, yahoo finance and NSE websites of different stocks.

6.2 Data Pre-Processing:

Raw data collected via AlphaVantage, needs to be modified and transformed into a more valuable version of itself.

Removal of null values, repetitive content and considering the parameters that are needed ahead, would be done here.

Hence, reducing data redundancy and efficiency in terms of data storage would happen in this phase of our proposed system.

6.3 Train Data & Test Data:

Training and testing of data predicts the accuracy of the model. Training if done precisely, testing of the model gives fine results.

Many research papers indicated that initially the dataset gets divided into two sets: 80% of it is used for training and 20% for testing.

50-50 % data has been used for testing and training in certain cases.

6.4 Apply Algorithm:

Among all of these algorithms, LSTM proves to be a better algorithm because LSTM is very powerful in sequence prediction problems and they're able to store past information.

This is important in our case because the previous price of a stock is crucial in predicting its future price accurately.

In addition to this, among all research papers that we studied, LSTM's results and efficiency proves to be better among the rest.

6.5 Sentiment Analysis

We will be using NewsApi for having the live news of the stocks and the finance sector.

Using sentiment analysis, we plan to know the current status and condition of the company of a particular stock.

Through this analysis, we'll have certain sentiment scores that will be generated which help to analyze the short term trend of the respective stock.

The sentiment scores could be positive and negative based on its features.

6.6 Evaluation and prediction of the output:

Analysis from the algorithms and sentiment analysis is combined together while determining the output.

The user can hence make an intelligent decision by first getting to know the current status of the stock, secondly prediction done via LSTM and then decide when to buy and for that particular stock.

7. CONCLUSIONS

The system proposed, uses different data mining tools and advanced techniques along with machine learning and state the condition of the sector depending on the sentiments. Among all the algorithms stated, LSTM outer performs based on the predictability criteria. Long Short-term Memory gives the most accurate predictions, while Linear Regression, KNN, Auto Arima and Prophet cannot be used for prediction as they have shown prediction value very far from actual values, from the most of the data. We can infer that LSTM is the most precise algorithm to predict the stock market value.

REFERENCES

1) Dev Shah, Haruna Isah, Farhana Zulkernine, "Predicting effects of news sentiments on stock market," 10.1109/BigData.2018.8621884. M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.
2) Pushpendu Ghosh, Ariel Neufeld, Jajati Keshari Sahoo on 'Forecasting directional movements of stock prices for intraday trading using LSTM and Random Forests', arXiv:2004.10178
3) Strader, Troy J.; Rozycki, John J.; ROOT, THOMAS H.; and Huang, Yu-Hsiang (John) (2020) "Machine Learning Stock Market Prediction Studies: Review and Research Directions," Journal of International
4) Stock Market Prediction using Machine Learning parmar2018.
5) V kranthi sai Reddy on " Stock Market Prediction using Machine Learning"https://www.researchgate.net/publication/328930285_Stock_Market_Prediction_Using_Machine_Learning
6) H. L. Siew and M. J. Nordin, "Regression techniques for the prediction of stock price trend," 2012 International Conference on Statistics in Science, Business and Engineering (ICSSBE), Langkawi, 2012, pp. 1-5.
7) Stock Market Prediction term paper https://www.scribd.com/document/489443204/StockMarketPredictionTermPaper-Final1-pdf
8) Automated stock price prediction using machine learning https://www.aclweb.org/anthology/W19-6403
9) Mojtaba Nabipour, Pooyan Nayyeri, Hamed Jabani, Shahab S., Amir Mosavi on 'Predicting Stock Market Trends Using Machine Learning and Deep Learning Algorithms Via Continuous and Binary Data; a Comparative Analysis'