# Dogo Rangsang Research JournalUGC Care Group I JournalISSN : 2347-7180Vol-11 Issue-01 - 2021STOCK MARKET TREND PREDICTIONUSING K-NEAREST NEIGHBOUR(KNN) ALGORITHM

# M. CHAITANYA KISHORE REDDY

#### M.S.K.P.SRI PRADA,K.SUDHAKAR, A.SUKEERTHI

Professor &HOD, Dept. of INFORMATION TECHNOLOGY, NRI INSTITUTE OF TECHNOLOGY, A.P., India.

Student, B.Tech (IT), NRI INSTITUTE OF TECHNOLOGY, A.P., India.

Abstract — Stock prediction has always been a challenging problem for statistics experts and finance. The main reason behind this prediction is buying stocks that are likely to increase in price and then selling stocks that are probably to fall. Generally, there are two ways for stock market prediction. Fundamental analysis is one of them and relies on a company's technique and fundamental information. In this paper author is performance of KNN(K-Nearest evaluating Neighbor) supervised machine learning algorithm. 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. 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 K-Nearest Neighbor to predict stock prices large small for the and capitalizations and in the three different markets, employing prices with both daily and up-to-the-minute frequencies.

#### **INTRODUCTION**

Stock market movement has always been ambiguous for investors because of various influential factors. This study aims to significantly reduce the risk of trend prediction with machine learning and deep learning algorithms. Four stock market groups, namely diversified financials, petroleum, non-metallic minerals and basic metals from Tehran stock exchange, chosen for experimental are

evaluations. Stock Market prediction remains a secretive and empirical art. Few people, if any, are willing to share what successful strategies they have. A chief goal of this project is to add to the academic understanding of stock market prediction. The hope is that with a greater understanding of how the market moves, investors will be better equipped to prevent another financial crisis. The project will evaluate some existing strategies from a rigorous scientific perspective and provide a quantitative evaluation of new strategies.

There are several data mining algorithms that can be used for prediction purposes in the field of finance. Some examples would be the naive Bayes classifier, the k nearest neighbour (KNN) algorithm and the classification and the regression tree algorithm (Wu et al. 2007). All the mentioned algorithms could fill the purpose of the paper but it will center around the kNN algorithm as a method of predicting stock market movements as well as the MA formula. The movements will be detected by looking at a large amount of historical data and finding patterns to establish a well estimated forecast. This specific algorithm was chosen as it is a simple but a very effective algorithm to implement when looking at large amounts of data (Berson et al. 1999). The KNN algorithm simply states: "Objects that are 'near' to each

#### UGC Care Group I Journal Vol-11 Issue-01 - 2021

other will have similar prediction values as well. Thus if you know the prediction value of one of the objects you can predict it for its nearest neighbours" (Berson et al. 1999). As a comparison with the KNN algorithm, the MA formula was chosen. The MA formula has its simplicity as a common factor with the KNN algorithm, but it is a statistical method used frequently by traders (Interactive Data Corp, 2014).

#### LITERATURE SURVEY

Aditya Menon et al. in 2019, this paper is focused on a review of neural model for forecast the stock tread after reviewing on a neural model they think that The long short term memory algorithm for predicting the economic information in confluence into the trendy era, this would be prioritized algorithm for forecasting.

**Arash Negahdari kia et al. in 2018**, as the stock prediction so many experiments and models, have been developed for prediction purpose on historical data like as in this paper the author present HyS3 graph-based semi-supervised model and through a network views Kruskal

based graph algorithm called ConKruG. In the future they think social media data, Twitter data could be used for the prediction of stock for better results using these algorithms.

Nuno Oliveira et al. in 2016, in this paper the author purposed a methodology by which they can access the value of stock prediction and microblogging data they used, for stock prices and return indices and some more like a portfolio. For this experiment, they have used huge data of Twitter, for all this experimental work they use Kalman filter to merge the microblogging data and some external sources and as a result, they found twitter data and blogging data were relevant for the purpose of forecasting these datasets were very useful. This result can be improved by using some more and different data such as social media datasets and others.

Smruti rekha das et al. in 2019, in this paper authors, used firefly method for forecasting the stock prices as an input dataset author collect from four different websites name as NSE-India, BSE, S&P 500 and FTSE, and all collected dataset transformed is well bv using proper mathematical formulas using by the backpropagation, neural network and more two

#### UGC Care Group I Journal Vol-11 Issue-01 - 2021

methods used for prediction, forecasting according to the time horizon of alternate days 1 day, 3 days, 5 days and so on. For future work, there may be some chance to get more accurate results by giving more parameters to the implemented algorithms.

#### **PROPOSED SYSTEM**

Predicting the Stock Market has been the bane and goal of investors since its existence. Everyday billions of dollars are traded on the exchange, and behind each dollar is an investor hoping to profit in one way or another. Entire companies rise and fall daily based on the behaviour of the market. Should an investor be able to accurately predict market movements, it offers a tantalizing promises of wealth and influence. It is no wonder then that the Stock Market and its associated challenges find their way into the public imagination every time it misbehaves. The 2008 financial crisis was no different, as evidenced by the flood of films and documentaries based on the crash. If there was a common theme among those productions, it was that few people knew how the market worked or reacted. Perhaps a better understanding of stock market prediction might help in the case of similar events in the future.

#### **RELATED WORK**

Tehran's stock market has been greatly popular lately due to the remarkable growth of the main index in the last decade. The important reason behind that is privatizing most of the stateowned in the Iranian constitution firms under the general policies of article 44. The shares of lately privatized firms can be bought by ordinary people under particular conditions. The market has some special features compared to other country's stock markets; for example, dealing price limitation that is  $\pm 5\%$  of opening price for every index in each trading day. This matter hampers scatter market shocks and irregular market fluctuations, political matters, etc. over a particular time and could form the market smoother. However, the effect of fundamental parameters on the market is considerable and the prediction task of future movements is not easy [23]. This study employed stock market groups (that are important for traders) to investigate the task of predicting future trends. In spite of remarkable progress in Tehran stock market in the recent decade, there has been not adequate papers on the stock price predictions and trends via novel machine learning algorithms. However, a paper has been published recently by Nabipour et al. [23] where they employed tree based models and deep learning algorithms to estimate future stock prices from 1 day ahead to 30 days ahead as a regression problem. The experimental results

#### UGC Care Group I Journal Vol-11 Issue-01 - 2021

indicated that LSTM (as the superior model) could successfully predict values (from Tehran Stock Exchange) with the lowest error. In this research. we concentrate on comparing prediction performance of nine machine learning models (Decision Tree, Random Forest, Adaboost, XGBoost, SVC, Naïve Bayes, KNN, Logistic Regression and ANN) and two deep learning methods (RNN and LSTM) to predict stock market movement. Ten technical indicators are utilized as inputs to our models. Our study includes two different approaches for inputs, continuous data and binary data, to investigate the effect of preprocessing; the former uses stock trading data (open, close, high and low values) while the latter employs preprocessing step to convert continuous data to binary one. Each technical indicator has its specific possibility of up or down movement based on market inherent properties. The performance of the mentioned models is compared for the both approaches with three classification metrics, and the best tuning parameter for each model (except Naïve Bayes and Logistic Regression) is reported. All experimental tests are done with ten years of historical data of four stock market groups (petroleum, diversified financials, basic metals and non-metallic minerals), that are totally crucial for investors, from Tehran stock exchange. We believe that this study is a new

research paper that incorporates multiple machine learning and deep learning methods to improve the prediction task of stock groups' trend and movement.

# **IMPLEMENTATION**

The Moving Average makes the line smooth and showcase the increasing or decreasing trend of stocks price.



# Return Deviation — to determine risk and return

Expected **Return** measures the mean, or expected value, of the probability distribution of investment **returns**. The expected **return** of a portfolio is calculated by multiplying the weight of each asset by its expected **return** and adding the values for each investment — Investopedia. Following is the formula you could refer to:





Before running code execute below two commands

Screen shots

Double click on 'run.bat' file to get below screen



In above screen click on 'Download Button' download the Apple Stock and competitors data from Yahoo Finance Dataset



In above screen I am Downloading of Apple Stock and Apple competitor Stock Data from Yahoo Finance Dataset.

Now click on 'Correlation Data' Button to find the correlation between Apple and Competitor Stock market Dataset. show the trend in the technology industry rather than show how competing stocks affect each other.

# **UGC Care Group I Journal** Vol-11 Issue-01 - 2021



Now click on 'Data PreProcessing button to drop missing values, split labels split train and test

Stock Trend Using KNN			0	×
Stock Tree	nd Prediction Using KNN			
Due Arbernening for Ageb Nock Discust Xishah 11 Janeson and Argent Arbeiten Xishahan and Xishahan Arbeiten Xishahan Arbeiten Xishahan Arbeiten Xishahan Arbeiten Xishahan Arbeiten Xishahan Arbeiten Xishahan Arbeiten Xishahan Arbeiten Xishahan Arbeiten Xishahan Arbeiten Xishahan Xishahan Arbeiten Xishahan Xishahan Arbeiten Xishahan Xishah	Dovaliad Dataset Correlations for Data Data Proprocessing KAN with Clafform Weights KNN Accuracy			
# # <u>e m</u> m 🛋 🖬 🔹 🧶 🖉 🖾 🛒	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	650 12/28	PM (2019	<b>P</b> 20

After pre-processing all missing values are dropped, Separating the label here, Scalling of X, find Data Series of late X and early X (train) for model generation and evaluation, Separate label and identify it as y and Separation of training and testing of model.

1752 records and 1226 used for training and 526 used for testing.

Now click on 'Run KNN with Uniform Weights' to generate KNN model with uniform weights and calculate its model accuracy



### **CONCLUSION**

In this paper, we proposed by comparing all of these algorithm values by percentage and depicted in the table and also represented in the graphical format we got the highest value for random forest algorithm which is the best algorithm for our purpose work and will be the best fit for prediction purpose. We implemented Random forest, SVM, Linear regression, and K-Nearest Neighbor algorithms which are supervised algorithms. We have implemented all four algorithms which have been functional for the stock prediction, as we have calculated the result for all different algorithms, we got to

#### **UGC Care Group I Journal** Vol-11 Issue-01 - 2021

In above screen we can see dataset contains total know that K-NN algorithms are the best and more suitable for prediction purpose.

# References

[1] J. J. Murphy, Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications. Penguin, 1999. [2] T. Turner, A Beginner's Guide To Day Trading Online, 2nd ed. New York, NY, USA: Simon and Schuster, 2007.

[3] H. Maqsood, I. Mehmood, M. Maqsood, M. Yasir, S. Afzal, F. Aadil, M. M. Selim, and K. Muhammad, "A local and global event sentiment based efficient stock exchange forecasting using deep learning," Int. J. Inf. Manage., vol. 50, pp. 432-451, Feb. 2020.

[4] W. Long, Z. Lu, and L. Cui, "Deep learningbased feature engineering for stock price movement prediction," Knowl.-Based Syst., vol. s164, pp. 163–173, Jan. 2019.

[5] J. B. Duarte Duarte, L. H. Talero Sarmiento, and K. J. Sierra Juárez, "Evaluation of the effect of investor psychology on an artificial stock market through its degree of efficiency," Contaduría y Administración, vol. 62, no. 4, pp. 1361–1376, Oct. 2017.

[6] Lu, Ning, A Machine Learning Approach to Automated Trading. Boston, MA, USA: Boston College Computer Science Senior, 2016.

[7] M. R. Hassan, B. Nath, and M. Kirley, "A fusion model of HMM, ANN and GA for stock market forecasting," Expert Syst. Appl., vol. 33, no. 1, pp. 171–180, Jul. 2007.

[8] W. Huang, Y. Nakamori, and S.-Y. Wang, "Forecasting stock market movement direction with support vector machine," Comput. Oper. Res., vol. 32, no. 10, pp. 2513–2522, Oct. 2005.

[9] Kalra S, Prasad JS. Efficacy of News Sentiment for Stock Market Prediction. Proc Int Conf Mach Learn Big Data, Cloud Parallel Comput Trends, Prespectives Prospect Com 2019. Published online 2019:491-496. doi:10.1109/COMITCon.2019.8862265

[10] Menon A, Singh S, Parekh H. A review of stock market prediction using neural networks.2019 IEEE Int Conf Syst Comput Autom Networking, ICSCAN 2019. Published online2019.