IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.301

Volume 3, Issue 8, April 2023

National Bombay Stock Trade Market Expectation Utilizing KNN Calculation

V. Hemanth Kumar[1] , K Vandana[2], M Vikas[3], K Deepak[4], M Hemanth Sai[5]

Assistant Professor, Department of Computer Science and Engineering[1]

U.G Scholars, Department of Computer Science and Engineering[2,3,4,5]

Raghu Institute of Technology, Dakamarri, Visakhapatnam, A.P, India

Abstract: *Predicting stock prices is a fascinating and difficult area of study. The economies of developed nations are compared to their power economies. As a result of their easy profits and low risk rate of return, stock markets are currently regarded as an illustrious trading field. For researchers in data mining and business, the stock market is thought to be a good place to work because of its numerous and constantly changing information sources. In order to assist investors, management, decision-makers, and users in making correct and informed investment decisions, we used the k-nearest neighbour algorithm and a non-linear regression approach in this paper to predict stock prices for a sample of six major companies listed on the Jordanian stock exchange. The results indicate that the KNN algorithm is reliable with a low error ratio; Consequently, the outcomes were reasonable and rational. Additionally, the prediction results were nearly identical to the actual stock prices, depending on the data on the stock price.*

Keywords: Stock Price Prediction, Listed Companies, Data Mining, K-Nearest Neighbour, Non-Linear Regression.

I. INTRODUCTION

Recent interests in business research have focused on challenging and demanding areas of future stock price predictions. The stock price prediction of movements in stock markets is a topic of intense interest for researchers, business communities, and interested users who believe that future occurrence is dependent on current and previous data (Kim, 2003). Notwithstanding, monetary information is considered as intricate information to estimate or potentially anticipate. According to the efficient market hypotheses (EMH) that Fama proposed, predicting market prices is seen as problematic. The financial market and financial information are thought to be separated by the EMH; Additionally, it demonstrates that the price swings are only the result of newly acquired information; and that market prices reflect all available information. The EMH declare that stocks are consistently in balance and are challenging for designers to hypothesize. In addition, in addition to buying and selling stocks and shares in stock markets, each stock is characterized not only by its price but also by other variables, such as the closing price, which represents the most important variable for predicting the next day's price for a specific stock. In addition, it has been confirmed that stock prices do not follow a random walk and that stock prediction requires additional evidence. Every variable that has an effect on stock movements over time has a relationship and specific behavior. Stock price predictions have taken into account a variety of economic factors, including political stability and other unforeseeable circumstances. Data mining technology is used to analyze a large volume of business and financial data and is used to determine stock movements. When existing data and their interactions need to be observed through the time dimension, temporal stock market mining is needed to provide additional capabilities. In stock expectations, a bunch of unadulterated specialized information, essential information, and determined information are utilized in forecast of future upsides of stocks. The unadulterated specialized information depends on past stock information while the key information addresses the organizations' action and the circumstance of market. When used in stock prediction, data mining classification methods combined with historical data provide a future value for each unknown entity of a company's stock value. Neural networks, regression, genetic algorithms, decision tree induction, and k-Nearest Neighbors (kNN) are among the classification techniques utilized in this prediction. In order draws near, an informational collection is separated into preparing informational index and testing set. kNN utilizes similitude measurements to contrast a given test element

DOI: 10.48175/IJARSCT-9599

ISSN 2581-9429 IJARSCT

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 8, April 2023

and the preparation informational collection. A record with n features is represented by each data entity. kNN selects k recodes from the training data set that are closest to the unknown records in order to predict a class label for that record.A Recommendation System Based on Trust:The hereditary calculation had been taken on by Shin et al. (2005); In Sweden, Hellestrom and Homlstrom (1998) used a statistical analysis based on a modified kNN to determine where correlated areas fall in the input space to improve the performance of prediction for the period 1987-1996. The number of trading rules was generated for the Korea Stock Price Index 200 (KOSPI 200). The Zimbabwe stock exchange offered both of the aforementioned models, the Weightless Neural Network (WNN) model and the Single Exponential Smoothing (SES) model Mpofu, for the purpose of predicting stock prices. Gavrilov et al.'s clustering stocks approach was used (2004) to group 500 Standard & Poor's stocks. The data consisted of 252 numbers, the opening stock price included. A fluffy hereditary calculation was introduced by Cao (1977) to find pair relationship in stock information in light of client inclinations. Markets, stock-trading regulations, and potential guidelines for mining pairs of stocks were the subject of the study; it likewise showed that such methodology is helpful for genuine exchanging. Besides, different investigations embraced kNN as forecast procedures, for example, (Subha et al., 2012; Liao and co. 2010; 2010 by Tsai and Hsiao; 2007 (Qian and Rasheed)

II. LITERATURE SURVEY

A Recommendation System Based on Trust:The hereditary calculation had been taken on by Shin et al. (2005); In Sweden, Hellestrom and Homlstrom (1998) used a statistical analysis based on a modified kNN to determine where correlated areas fall in the input space to improve the performance of prediction for the period 1987-1996. The number of trading rules was generated for the Korea Stock Price Index 200 (KOSPI 200). The Zimbabwe stock exchange offered both of the aforementioned models, the Weightless Neural Network (WNN) model and the Single Exponential Smoothing (SES) model Mpofu, for the purpose of predicting stock prices. Gavrilov et al.'s clustering stocks approach was used (2004) to group 500 Standard & Poor's stocks. The data consisted of 252 numbers, the opening stock price included. A fluffy hereditary calculation was introduced by Cao (1977) to find pair relationship in stock information in light of client inclinations. Markets, stock-trading regulations, and potential guidelines for mining pairs of stocks were the subject of the study; it likewise showed that such methodology is helpful for genuine exchanging. Besides, different investigations embraced kNN as forecast procedures, for example, (Subha et al., 2012; Liao and co. 2010; 2010 by Tsai and Hsiao; 2007 (Qian and Rasheed)

III. TECHNOLOGICAL INNOVATION THEORY

3.1 Existing Analysis

Shin et al. had utilized the genetic algorithm (2005); In Sweden, Hellestrom and Homlstrom (1998) used a statistical analysis based on a modified kNN to determine where correlated areas fall in the input space to improve the performance of prediction for the period 1987-1996. The number of trading rules was generated for the Korea Stock Price Index 200 (KOSPI 200).The two models referenced were given in the Zimbabwe stock trade to foresee the stock costs which included Weightless Brain Organization (WNN) model and single outstanding smoothing (SES) model Mpofu (2004). Gavrilov et al.'s clustering stocks approach was used (2004) to group 500 Standard & Poor's stocks. The data consisted of 252 numbers, the opening stock price included.

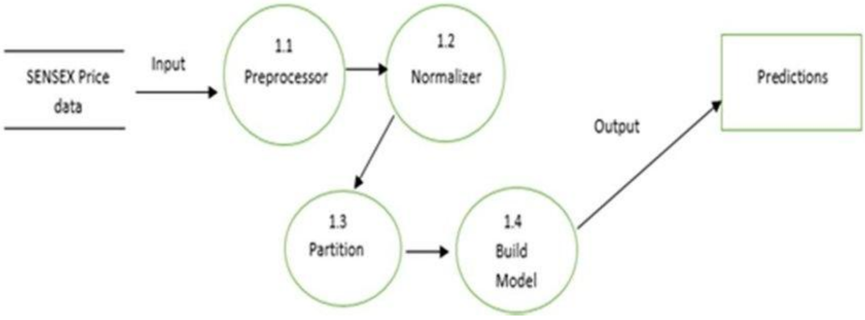
3.2 Proposed Analysis

To overcome these challenges, we propose the KNN algorithm for stock price prediction in this project. Using the proposed method, it is simple to forecast stock market trends and estate prices. The historical dataset is used to calculate the k-weighted nearest neighbor for KNN, which is then used to predict future stock market indices.

IV. APPLICATION ARCHITECTURE

```
from tkinter import *
import tkinter
from tkinter import filedialog
import numpy as np
```

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 8, April 2023




```
from tkinter import simpledialog
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.svm import SVC
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_squared_error
from sklearn.metrics import accuracy_score
from sklearn.metrics import f1_score
from sklearn.metrics import roc_auc_score
from sklearn import metrics
from sklearn.neighbors import KNeighborsClassifier
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import RandomForestClassifier
from sklearn.linear_model import LogisticRegression
from xgboost import XGBClassifier
from sklearn.ensemble import AdaBoostClassifier
from sklearn.naive_bayes import GaussianNB
from keras.utils.np_utils import to_categorical
from keras.models import Sequential
from keras.layers import Dense, Dropout, Flatten, LSTM, Activation
from keras.utils.np_utils import to_categorical
import webbrowser

main = tkinter.Tk()
main.title("Predicting Stock Market Trends Using Machine Learning and Deep Learning Algorithms Via Continuous and Binary Data a Comparative Analysis") #designing main screen
main.geometry("1000x650")

global filename
global dataset
global trainXX, trainY, scalerX, original_data,testX
c_accuracy = []
c_roc = []
c_fscore = []
b_accuracy = []
b_roc = []
b_fscore = []

global plist,tlist
global plist1,tlist1
```

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
IJARSCT
Impact Factor: 7.301
Volume 3, Issue 8, April 2023

```
def difference1(datasets, intervals=1):
 difference = list()
 for i in range(intervals, len(datasets)):
 values = datasets[i] - datasets[i - intervals]
difference.append(values)
 return pd.Series(difference)

def convertDataToTimeseries1(dataset, lagvalue=1):
dframe = pd.DataFrame(dataset)
 cols = [dframe.shift(i) for i in range(1, lagvalue+1)]
cols.append(dframe)
dframe = pd.concat(cols, axis=1)
dframe.fillna(0, inplace=True)
 return dframe

def scaleDataset1(trainX, testX):
scalerValue = MinMaxScaler(feature_range=(-1, 1))
scalerValue = scalerValue.fit(trainX)
trainX = trainX.reshape(trainX.shape[0], trainX.shape[1])
trainX = scalerValue.transform(trainX)
testX = testX.reshape(testX.shape[0], testX.shape[1])
testX = scalerValue.transform(testX)
 return scalerValue, trainX, testX

def forecastRNN1(model, batchSize, testX):
testX = testX.reshape(1, 1, len(testX))
 forecast = model.predict(testX, batch_size=batchSize)
 return forecast[0,0]

def inverseDifference1(history_data, yhat_data, intervals=1):
 return yhat_data + history_data[-intervals]

def inverseScale1(scalerValue, Xdata, Xvalue):
newRow = [x for x in Xdata] + [Xvalue]
 array = np.array(newRow)
 array = array.reshape(1, len(array))
 inverse = scalerValue.inverse_transform(array)
 return inverse[0, -1]


def difference(datasets, intervals=1):
 difference = list()
 for i in range(intervals, len(datasets)):
 values = datasets[i] - datasets[i - intervals]
difference.append(values)
 return pd.Series(difference)

def convertDataToTimeseries(dataset, lagvalue=1):
```

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 8, April 2023

```
dframe = pd.DataFrame(dataset)
 cols = [dframe.shift(i) for i in range(1, lagvalue+1)]
cols.append(dframe)
dframe = pd.concat(cols, axis=1)
dframe.fillna(0, inplace=True)
 return dframe

def scaleDataset(trainX, testX):
scalerValue = MinMaxScaler(feature_range=(-1, 1))
scalerValue = scalerValue.fit(trainX)
trainX = trainX.reshape(trainX.shape[0], trainX.shape[1])
trainX = scalerValue.transform(trainX)
testX = testX.reshape(testX.shape[0], testX.shape[1])
testX = scalerValue.transform(testX)
 return scalerValue, trainX, testX

def forecastRNN(model, batchSize, testX):
testX = testX.reshape(1, len(testX))
 forecast = model.predict(testX)
 return forecast[0]

def inverseDifference(history_data, yhat_data, intervals=1):
 return yhat_data + history_data[-intervals]

def inverseScale(scalerValue, Xdata, Xvalue):
newRow = [x for x in Xdata] + [Xvalue]
 array = np.array(newRow)
 array = array.reshape(1, len(array))
 inverse = scalerValue.inverse_transform(array)
 return inverse[0, -1]
```

The author of this paper evaluates the KNN (K-Nearest Neighbor) supervised machine learning algorithm's performance. In the money world stock exchanging is quite possibly of the main action. An attempt to predict the future value of a stock or other financial instrument traded on a financial exchange is known as stock market prediction. The programming language is utilized to anticipate the financial exchange utilizing AI is Python. In this paper, we propose a Machine Learning (ML) method that will learn from the stock data and gain intelligence before making an accurate prediction with the learned information. In this setting this study utilizes an AI strategy called K-Closest Neighbor to foresee stock costs for the huge and little capitalizations and in the three unique business sectors, utilizing costs with both everyday and expert frequencies.Since the stock market's inception, investors have tried and failed to predict it. On the exchange, billions of dollars are traded each day, with investors hoping to profit in one way or another behind each dollar. Every day, the market's behavior determines the rise and fall of entire businesses. Should a financial backer have the option to precisely foresee market developments, it offers a tempting commitments of riches and impact. Therefore, it should come as no surprise that each time the Stock Market behaves badly, the difficulties it poses are brought to the attention of the general public. The flood of films and documentaries based on the 2008 financial crisis shows that it was no different. Those productions all had one thing in common: very few people were aware of how the market worked or reacted. In the event that similar events occur in the future, perhaps having a better comprehension of stock market prediction will be of assistance.Stock market prediction is still a secretive and empirical art, despite its prevalence. Hardly any individuals, if any, will share what fruitful methodologies they have. This project's primary objective is to improve academic knowledge of stock market prediction. The expectation is that with a more prominent comprehension of how the market moves, financial backers will be better prepared to forestall another monetary

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 8, April 2023

emergency. The project will provide a quantitative evaluation of new strategies and a rigorous scientific evaluation of some existing strategies.There are a number of data mining algorithms that can be used to predict financial outcomes. The naive Bayes classifier, the KNN algorithm, and the classification and regression tree algorithms are a few examples (Wu et al.). 2007). The paper could use any of the aforementioned algorithms, but the focus will be on the kNN algorithm and the MA formula as methods for predicting stock market movements. A well-estimated forecast will be made by looking at a large amount of historical data and identifying patterns that indicate the movements. According to Berson et al., this particular algorithm was chosen because it is easy to use and very effective when analyzing large amounts of data. 1999). This is all that the KNN algorithm says: Prediction values for objects that are "near" one another will also be similar. Therefore, you can predict the prediction value of one object for its closest neighbors if you know its prediction value" (Berson et al.). 1999). As a correlation with the KNN calculation, the Mama equation was picked. The Mama recipe has its straightforwardness as a typical component with the KNN calculation, yet it is a measurable technique utilized as often as possible by dealers (Intuitive Information Corp, 2014).

Multispectral prediction, distortion-controlled prediction, and lempel-ziv based prediction are all methods that are currently in use for stock prediction. These are based on the fact that the essential information is kept in an easily accessible format while redundancy is eliminated, making the data representation smaller (Azhar et al.). 1994). Despite the aforementioned methods, the KNN algorithm and the MA formula proved to be the most suitable tools for the project due to its scope.The author used the Yahoo Finance stock Dataset for the experiment. The following are some example records from that dataset that contain request signatures. I have additionally utilized same dataset and this dataset is accessible inside 'dataset' organizer.Dataset model

['High', 'Low', 'Open', 'Close', 'Volume', 'Adj Close']Above list are the segments of Yippee finance

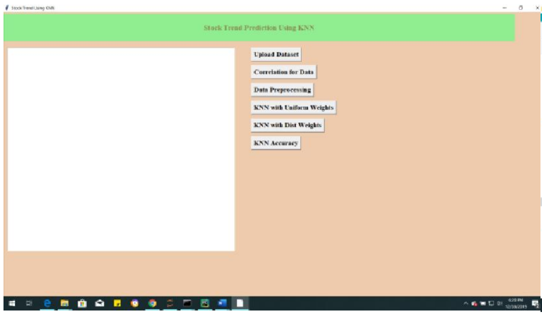
2017-01-05	116.860001	115.809998	115.919998	116.610001	22193600.0	111.393303
2017-01-09	119.430000	117.940002	117.949997	118.989998	33561900.0	113.666824

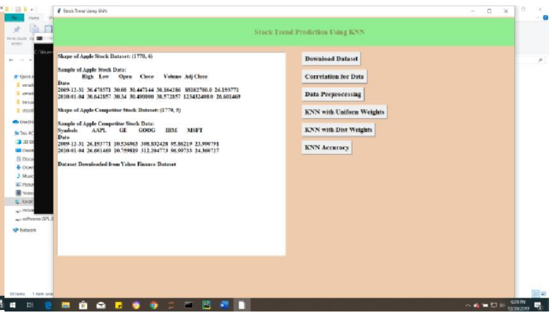
Return Deviation — to determine risk and return

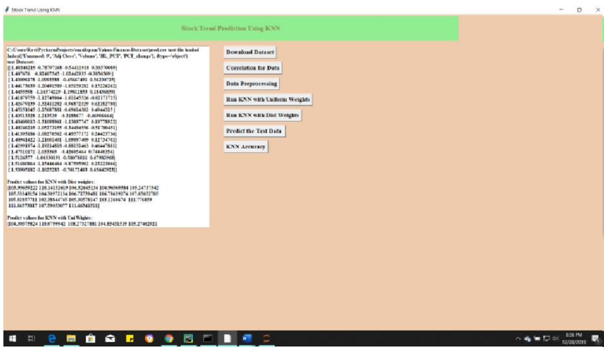
Following is the formula you could refer to:

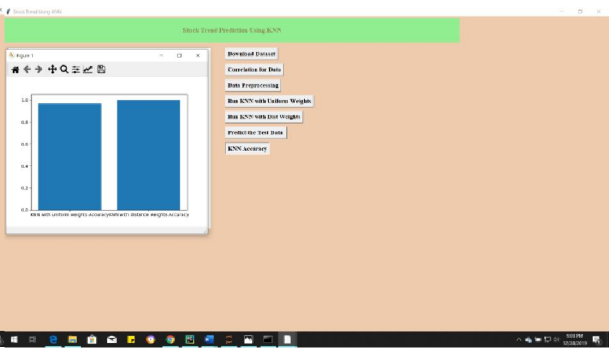
$$r_t = \frac{p_t - p_{t-1}}{p_{t-1}} = \frac{p_t}{p_{t-1}} - 1$$

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 8, April 2023

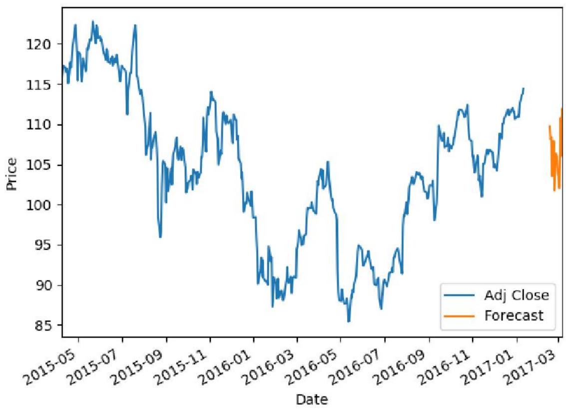








From above graph we can see that distance weights has little bit better better accuracy compare to Uniform weights, in above graph x-axis contains algorithm name and y-axis represents accuracy of that algorithms



Plotting the Prediction for KNN with Distance Weights

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.301

Volume 3, Issue 8, April 2023

V. CONCLUSION

K-Nearest Neighbor algorithms, which are supervised algorithms, were used in our implementation in this paper. By measuring the accuracy of the various algorithms, we discovered that the random forest algorithm is the most suitable algorithm for predicting the market price of a stock based on various data points from the historical data. As we have calculated the result using this algorithm, this algorithm is the best for prediction and will be very effective and profitable for those who invest their money in the stock market. As we have calculated this result through different kinds of data points from lateral data, this algorithm is the best for prediction. As we have calculated this result through different kinds of data Because it has been tested on sample data and is trained on a large amount of historical data, the algorithm will be a great asset for brokers and investors investing in the stock market. In contrast to previous machine learning models, the project demonstrates a machine learning model that accurately predicts stock value.

VI. FUTURE WORK

The addition of additional parameters and factors, such as financial ratios, multiple instances, and so forth, will expand the project's scope in the future. The accuracy will increase as more parameters are considered. The calculations can likewise be applied for breaking down the items openly remarks and in this way decide designs/connections between the client and the corporate worker. The utilization of conventional calculations and information mining procedures can likewise help foresee the corporation‟s execution structure in general

REFERENCES

[1]. R. Iacomin, “Stock Market Prediction,”2015 19th International Conference on System Theory, Control and Computing(ICSTCC0), October 14-16, CheileGradistei, Romania.

[2]. S., Kamath, “Stock Market Analysis”, Master’s Projects, pp. 326, 2012.

[3]. P.Domingos, “A Few Useful Things to Know about Machine Learning,”Communications of the ACM, Vol. 55 No. 10, October 2012.

[4]. Gharehchopogh, F.S., &Khaze, S.R (2012), Data Mining Application for Cyber Space Users Tendency in Blog Writing: A Case Study. International Journal of Computer Applications, 47(18), 40-46.

[5]. Berry, M. J., &Linoff, G. S. (2004). Data Mining Techniques: for marketing, sales, and customer relationship management. Wiley.com.

[6]. S AbdulsalamSulaimanOlaniyi, Adewole, Kayode S., Jimoh, R. G, “Stock Trend Prediction Using Regression Analysis- A Data Mining Approach”, ARPN Journal of Systems and Software Volume 1 No. 4, JULY 2011, Brisbane, Australia.

[7]. M. Ballings, D. VandenPoel, N. Hespeels, R. GrypEvaluating multiple classifiers for stock price direction prediction Expert SystAppl, 42 (20) (2015), pp. 7046-7056. [8]R. Dash, P.K. DashA hybrid stock trading framework integrating technical analysis with machine learning techniques J Finance Data Sci, 2 (1) (2016).

[8]. E.A. Gerlein, M. McGinnity, A. Belatreche, S. ColemanEvaluating machine learning classification for financial trading: an empirical approach Expert SystAppl, 54 (1) (2016), pp. 193-207

[9]. J. Patel, S. Shah, P. Thakkar, K. KotechaPredicting stock market index using fusion of machine learning techniques Expert SystAppl, 42 (4) (2015), pp. 2162-2172

[10]. K. KimFinancial time series forecasting using support vector machinesNeurocomputing, 55 (1–2) (2003), pp. 307-319 [12]R.K. Nayak, D. Mishra, A.K. RathA Naïve SVM-KNN based stock market trend reversal analysis for Indian benchmark indicesAppl Soft Comput, 35 (1) (2015), pp. 670-680

[11]. S. Barak, M. ModarresDeveloping an approach to evaluate stocks by forecasting effective features with datamining methods Expert SystAppl, 42 (3) (2015), pp. 1325-1339

[12]. C.-Y. Yeh, C.-W. Huang, S.-J. LeeA multiple-kernel support vector regression approach for stock market price forecasting Expert SystAppl, 38 (3) (2011), pp. 2177-2186

[13]. Y. Xiao, J. Xiao, F. Lu, S. WangEnsembleANNs-PSO-GA approach for day-ahead stock e-exchange prices forecastin

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.301

Volume 3, Issue 8, April 2023

[14]. Vasireddi Hemant Kumar, K Suganya Devi, G N V RajaReddy.,(2021), "Deep feed forward neural network-based screening system for diabetic retinopathy severity classification using the lion optimization algorithm", Graefe's Arch Clin Exp Ophthalmology. 2021 Sep 10. doi: 10.1007/s00417-021-05375-x

[15]. Hemanth Kumar Vasireddi, Suganya Devi. K., (2020)"An ideal big data architectural analysis for medical image data classification or clustering using the map-reduce framework", Lecture notes in electrical engineering, ICCCE-2020, 978-981-15-7960-8, vol:698, Pages. 1481-1494, Springer, Singapore, https://doi.org/10.1007/978-981-15-7961-5_134.