ISSN 2278-3091
Volume 9 No.2, March - April 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse177922020.pdf
https://doi.org/10.30534/ijatcse/2020/177922020

Stock Price Prediction Using Tech News Based Soft Computing Approach

R.Suganya[1], Dr.A.Sasi Kumar[2]

[1]Research Scholar, Department of Information Technology, School of Computing Sciences,
Vels Institute of Science, Technology and Advanced Studies (VISTAS), Pallavaram, Chennai, India.
srinisuganya7682@gmail.com
[2]Professor, Department of Information Technology, School of Computing Sciences,
Vels Institute of Science, Technology and Advanced Studies (VISTAS), Pallavaram, Chennai, India.
askmca@yahoo.com

ABSTRACT

In the stock market prediction foremost basic factor is the output stock price. What is more there's the test of demonstrating whether the stock market sectors are sure or not. The foregone behavior conclusion of the market is a difficult task and that has been much mentioned by researchers and data science observers. In our proposed work and by day to day stock price observance and intense research we find out that the future stock price information in accordance with particular stock details may influence the stock price significantly. Even though many adhoc methods are there in real time, this concept significantly improves the performance of the stock price prediction with respect to that of environmental factors. Our methodology reads the appropriate news channel and in convergence with the soft computing results a good predictive development. Additionally they accept that the stock trade follows a stochastic procedure, which recommends that the best forecast depends on the present stock information. The suggestions of news will be cross checked with the experts group.

Key words: Genetic Algorithm, Proficient Market Hypothesis, Soft Computing, Stochastic process, Stock market prediction

1. INTRODUCTION

Presently days, an average recognition is that a lot of capital ventures are made around the Stock Markets all through the world. National economies and strategies joined together powerfully and highly influence of the operations of the Stock Market sector. In the most recent advancement varying indicators[1] incorporate moving normal, intermingling, difference parameters, irregular list and so on. Moreover, recently the Markets have become a lot of accessible investment tool, that include computing associated techniques using intelligent agents[2]. In another development neural network[3] can be used to analyse the twitter sentiment score for a given stock. In Case based reasoning[4] based both the fundamental and technical predictors are used to find the future stock value. The forecasting of stock price[5] is not only for long term bulk investors but also for common people interested to do trading account.

Stock price information's are not solely associated with economic parameters, however they influence lifestyle in a very a lot of direct means. So they constitute a mechanism that has vital and direct social impacts. The characteristic that each one Stock Markets[6] have in common is that the uncertainty, which is related with their short and long state for further. This feature is undesirable for the investor however it's additionally ineluctable whenever the exchange is chosen because the investment tool. The simplest that one will do is to undertake to scale back this uncertainty. Stock Market Prediction (or forecasting) is one among the instruments during this method [7]. Many researchers provide many methods that motivates others to make an attempt to predict the securities market.

There is no doubt that no one will do profit at all the times or losing frequently and it is not the individual work. It is the involvement of various parameters and other technological executions associated with particular stock to be profitable. Earning profit from an investment in stocks needs various considerations related to that stock. To be honest, it is highly volatile in nature. There are numerous endeavors to achieve a route or in other to foresee the fate of the stock cost. The question now is which factors determines the future stock price?. The approach is quiet convincing that it may need a hybrid technique for bot buying and selling strategies[8]. The data belongs to the stock market exchange comes from the study of related data.

2. LITERATURE REVIEW

Based on others evaluation to the literature with a purpose defines a few fundamental traits to examine the review of previously done works. The characteristic nature of all parameters is a real concern in finding the precise definition of all prediction technique[9] in accordance with the stock market. It needs a individual stock approach rather than a global technique.

The survey we did as such far gives us various ends that prompted the estimation of future stock cost. The research work of Hesieh and Tsibouries et al gives more informations about the stock behavior. It concludes that the market is not random and it may be a non linear.It also says that by applying some high technological methods it is surely predictable. According to White's[10] study either the linear model or the Neural Network can do similarities in patterns within the records and its impact.

There are lot of studies on stock price prediction using Soft Computing techniques and text mining[11] were performed regularly. These examinations utilized different sorts of fuzzy rationale, Genetic Algorithms (GAs) and semantic content mining to foresee precisely the stock cost and the course of its change. The principle inspiration for the utilization of hereditary calculations for grouping is to choose best guidelines at the expectation of stock cost. The significant bit of leeway of utilizing GAs for this assignment is the development of worldwide hunt performed by this strategy, which improves the probability of acquiring a lot of rules with high prescient exactness.

GAs are computational methods[12] dependent on a Darwinian hypothesis, whose objective is to streamline a given wellness work .Genetic Algorithm was initially developed in 1960 by John Holland. The initial idea is to select best species for mutation and these concepts were introduced to the computers. Numerous investigations have been done in the zone of GAs so as to actualize new venture systems for the securities exchange. An ongoing article proposed another methodology in anticipating the stock using genetic algorithmic[13] approach to bring out good accuracy in prediction. Recently, Brian et al. investigated that the correlation of sentiment analysis of stock price in news channel with stock increase and decreases using Pearson correlation coefficient for stocks. In this work, we took a novel approach of predicting rise and fall in stock prices based on the sentiments[14] extracted from news channel to find the correlation. The core contribution of our work is the development of a sentiment analyzer which works better than the existing approach.

Sentiment analysis or opinion mining is the process of determining whether the derived text had developed an effect of positive or negative or neutral. Sentiment analysis is extremely useful in news analysis and social media monitoring as it allows us to view the public opinion. Earlier , Wiebe, Hoffman [15] [16] presented a new approach that automatically identify the contextual polarity for a large set of sentence expressions, and produces useful results.Gabriele Ranco [17] made an analysis of twitter information and concluded that it provides a valuable insight into public sentiment in a geographic region. Sandner and Welpe [18] in his research foretell the election results using Twitter data, and developed a model to predict the result usin g the tweets posted by various people.J. Yi, T. Nasukawa, R.B., Niblackhad found a new method of Extracting sentiment about a specific topic using NLP techniques.

La Sheng Yu [19] in his research addressed opining mining in the social media that gives an opportunity to explore more on news and social media and developed various useful patterns that will influence the results. According to Sasikumar et al[20], insists the importance of parameter selection in real time data networks. Generation of new test conditions [21] by combining machine learning with data mining gives a very good accuracy rate for real time data. Tawarish, M. & Satyanarayana[22] in his review emphasized about the decision tree approach and other versatile methods to find the price band. Haur Teng & king Kouk[23] improved the whale optimization technique for dedicated prediction techniques.

3. STOCK MARKET PREDICTION AND ITS ANALYSIS

A number of different ways are applied in order to foresee Stock Market returns. These ways are regularly arranged in four significant classes: Technical Analysis ways, Elementary Analysis ways, Ancient statistic forecasting and Machine Learning ways. Technical analysts, called chartists, attempt to predict the market by tracing patterns that come back from the study of charts which describe historic information of the market. Elementary analysts study the intrinsic value of associate stock and that they invest on that if they estimate that its current worth is lower that it's intrinsic worth.

The ancient statistic predicting models uses linear prediction technique to find useful patterns in historic data. It is separated in to two classes: the univariate and the bivariate. The implementation of the variable regression models, proves to some extent but it lags in time sequence and accuracy. The capturing of the data in the real time is also a constraint.

There are lot ot other way in which the researchers were focused. The usage of Machine Learning techniques and the samples made a impact in the field but it needs more sampling and learning practices. Data formatting is also a constraint.

The level of prediction varies with the methods and the data set they implemented. Still may tools are available in online for prediction none of them has been tested to be the consistent in producing the outcome. Our approach takes this issue as a predominant one and proceeds further with more learning approach.

The prediction of the stock market is a challenging criterion with no doubt that it is quiet complex in nature. The usage of diverse technique leads to data containment and hence the prediction of such type may doubt the outcome and it proves to be faulty. Hence both linear and non linear regression lags accuracy.

The techniques can be classified as: fundamental approach, Technical analysis and technological approach. The time series Prediction is quiet straight forward but it never always true in values. The criterion to this categorization is the type of gear and the kind of statistics that each technique is using that allows you to predict the marketplace. What is common to these strategies is that they're used to are expecting and Ahence benefit from the market's destiny conduct.

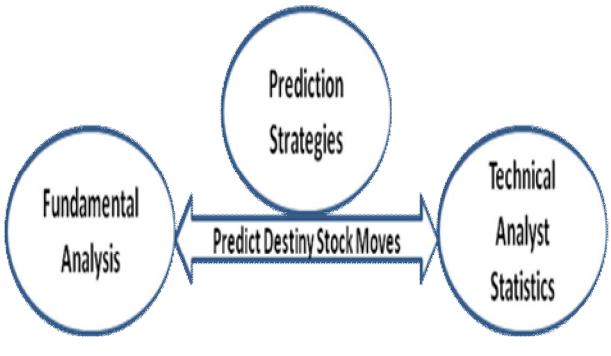

Figure 1: Fundamental stock market prediction analysis

The basic prediction strategy is shown in Figure 1. It's a bridge between the fundamental analysis and stock value. The historical rule strategy fails with inadequate explanation and it also lags accuracy. It is also found that there is no device using fundamental technique to predict the stock price frequently.

Technical analysis evaluation uses is the method of predicting the stock price based on time frequent formulas. It then gives the appropriate time to buy or sell a script that yield to be profitable. As there are other external parameters that affect the price this methods also lags in accuracy. It is said to be a partially acceptable condition. Technical statistics along with capital fee, volume, of trade, maximum value and minimum value builds the formula.

As a result, on applying the formula it produces a chart with an interpretation. Value diagrams are utilized to hit upon qualities, these characteristics are thought to be fundamentally founded on past time consumables and henceforth it won't go long.

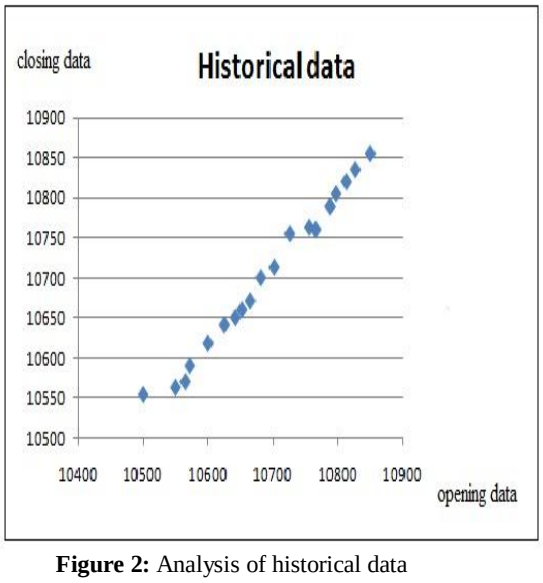

Figure 2: Analysis of historical data

The Figure 2 shows the variations of the closing data and opening data for a given stock. It's a regular approach used to predict the market trend and it fails considerably. The major setback of this method is the buying and selling are not linear and therefore the approximation of non linearity results in an error. Hence this method too be not substantially success.

4. NEURAL NETWORK FORECASTING USING GENETIC ALGORITHM

The manner in which this idea logy arises is really awesome. Today the people belong to neural community keeps on growing as there is a persistent dement in this versatile field. The neural community plays a key role in the field of prediction it's a clear benchmark standard. Neural network combine the feature of training and testing and gives immense power to the predictive environment. it is mainly used in the field of education and research to explore the skill of the methodology. It gains over ran down walk and prove way for the right prediction. The entire outcome depends on the data set that we are going to use. So the selection of data set and the allocation of learning and testing phenomena is vital. The reason for the deviation is sometimes the real values are varying much than the experimental values. In that the percentage of learning and testing should be adjusted to get an optimal value. Consequently Neural Network plays a vital role in real fact problems.

In econometrics many forecasting models such as time constraint framework, linear probabilistic were used. These models are used to find the linear relationships between various parameters involved in price estimation. As per Jenkins(1976), the rule of choice states among the models the one with less number of parameters are chosen for experimentation.

There is no trade off between the accuracy and the number of parameters chosen. Many methods are formulated to cut short the parameters but obviously it affects the performance. The concept of multivariate regression is also applied and the results are studied against the real time entities. Many tools to predict the price fails at the time of rapid economic decision during nation's emergency and natural disaster and other external factors that are triggered by an event. The various experiments in neural network to predict the stock price shows that the accuracy can be widely improved by using some associate algorithms in conjunction with neural network without disturbing the core data set.

The essential etymological standards of Genetic Algorithms are propelled by the component of regular assessment where influential people can most presumably win the challenging condition. This strategy is applied accomplishment to different spaces, for example, frontier streamlining, modernized expectation, gadget execution higher, financial matters, human resistant frameworks, pre-biology, control hereditary qualities, advancement and self-learning. The primary thing is connected with the presentation of a fundamental populace of m arbitrarily by carefully choosing individuals.

The underlying people shapes the essential time. The subsequent one issue m individual's inputs and gives yield with an assessment for every one of them dependent on a target. This assessment depicts how near our requests are for each such m individuals. Eventually the 33% variables are chargeable for the parts of the ensuing innovation. Another trouble that is related with a GA is the idea of the end measurement. This might be overcome by the assessment of some of advancement cycles (ages).

When the essential period has been portrayed a wellness component has for use so as to assess all of the m chromosomes of a time. The current GA permits us to apply smash of explicit strategies to appraise the ideal wellness of a chromosome.

5. METHODOLOGY AND ITS CALCULATION

There are two most significant classifications of system preparing the steady and the batch processing. Both using the gradual preparing of the loads of the network and are balanced each time when every single one of the info tests are provided to the system. In group mode the preparing of the loads are balanced handiest when the entirety of the preparation tests have been introduced to the system. The genetic approach is shown in Figure 3.

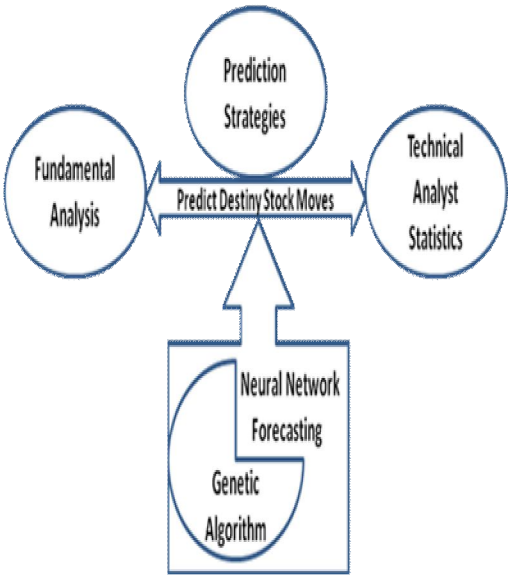

Figure 3: The genetic approach

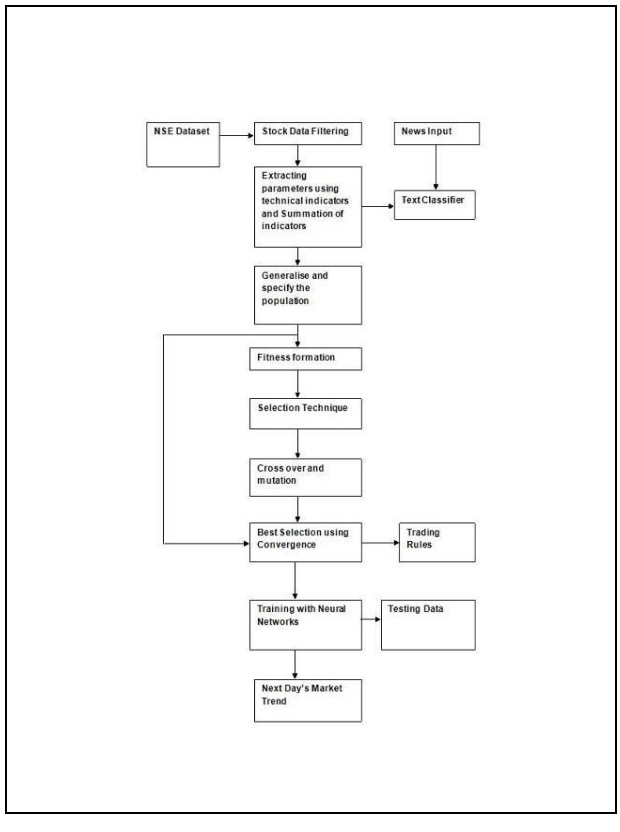

Figure 4: Prediction Methodology

The neurons of a next network are involved throughout all the layers. Each system has got one or more layer, at least 1 shrouded layers and one yield layer. Aside from the allotted layer, it likewise incorporates extra neurons. Even from the existing allotted layer, it also includes additional neurons. The methodology for predicting the stock price is shown in Figure 4 and the algorithm is given below. In our work the output layer has only one neuron. The value of the soft computing ranges from 0 to 1. In that the 80% are used for learning and the remaining 20% are used for testing.

The iterations are repeated for different inputs and the network was trained The or all the possible scenarios. The next step is the test phase in which the network is abruptly tested for all the samples in various time domain and in the occurrence of various parametric situations. The correlation between the learning and the testing pattern are studied and it is further used to produce the predicted pattern. The process is repeated until we get the required output that matches the real value. Effective care should be taken in doing the iterations as it is proportional to the time domain.

6. TECH NEWS-INTENSITY BASED NEURO-GENETIC ALGORITHM

The input to the algorithm is the three months data from the National Stock Exchange website. It's a free downloadable. It selects the optimum value with respect to the intensity of the threshold executed by the previous transactions and it fix the rise band and stop loss criteria limits..If the intensity value increases the threshold, the stock price may increase whereas if it goes down the stop loss criteria then the stock price may drop.

Algorithm : Intensity based data filtering algorithm

Aim: To find future stock price

Input: Open price, close price, max price, min price, vol of trans(in no's).

Output: Future stock price

Step-1 Data sample collection
Step 2 Fixing training and testing metric data
Step 3 Encoding of ANFIS
Step 4 Initialize a genuine population
Step 5 Decoding of data to ANFIS
Step 6 Perform RMSE
Step 7 Update the value of light intensity
Step 8 Compare the intensity and move the respective Position.
Step 9 Update the new value.
Step 10 Rank the iterated values and find the current best
Step 11 If Iterative value< Max then Goto Step 5
Step 12 Return the best value
Step 13 Stop

The sample iterative loop is given below

```
while (result not optimum)
 {
 select the best chromosomes;
 do crossover operation;
 do mutation;
 compute fitness of each chromosome in the population;
 }
```

The following graph (Figure 5) shows the comparison of iterated value of various prediction methods. It's a graph between evaluation and the training datasets of various prediction strategies. It is found that the algorithmic approach is more important for the iterations,

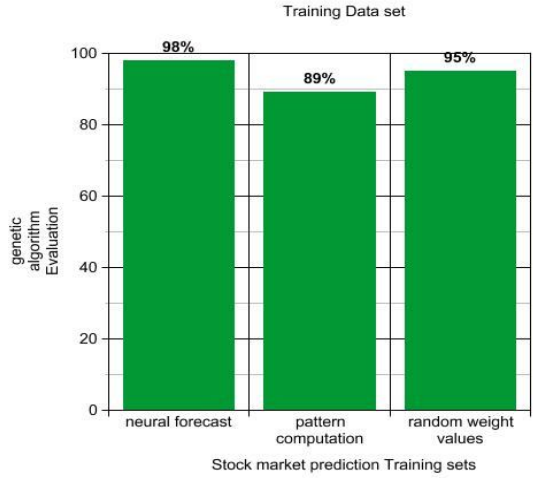

Figure 5: Comparison of various prediction methods

The mass of neural layers are nothing but the mass of neurons in each concealed layer and it is corresponding to the ability of the network to react to increasingly entangled capacities. The chance of error is overcome by comparing with pervious iterated value and intern it gives the new real value ahead of the existing value and hence it is placed in higher position. The idea behind the technique is the greater complex of a network is the extra sensitive towards the noise or other signal.

The noise level inside the system is kept minimum by comparing it with the values from 0 to 1. It is further fine tuned to avoid the level switch off impulsement. For our experiment we had consider the script name CUB.

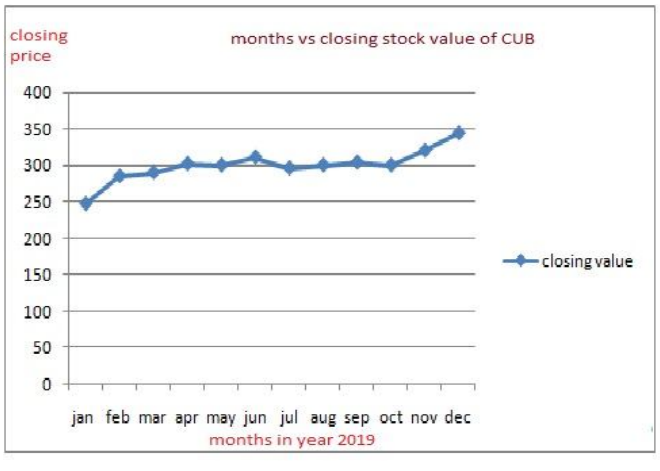

Figure 6: Graph representing the results of training sets using genetic approach

In Figure 6 we will find the variations of closing stock data with respect to stock price variations. The way that a network behaves is depicted by using the following patterns. Every pattern consists of two elements the input and the output part (target). Initially the weights of the networks are assigned to random values. At that point the info some portion of the main example is provided to the neural system. The framework processes a yield which depends on the estimations of its info, its loads, and the addictive part of the mass of neurons in each layer. The residuals of the AR models are then compared with the known set of data and it is further studied with its variation from the real value.

7. CONCLUSION

The work we proposed uses the novice feature of the neural network with that of the genetic algorithm. The parameters we had chosen here are more rational. Our rapid iterative algorithm with the genetic approach proves to be a modest one to predict the stock price. Furthermore, the models can be improved by adding some optimization techniques that can further more increase the accuracy of the prediction.

REFERENCES

1. Yauheniya Shynkevich,T M, McGinnitty, **Forecasting price movements using technical indicators: Investigating the impact of varying input window length**, *Neurocomputing, Elsevier*, Vol. 264, pp. 71-88, 2017. https://doi.org/10.1016/j.neucom.2016.11.095
2. Mohammad Hossein Fazel Zarandi, Esmaeil Hadavandi**, A hybrid fuzzy intelligent agent-based system for stock price prediction**, *International Journal of Intelligent systems*, Vol.27, No.11, pp.947-969, 2012. https://doi.org/10.1002/int.21554
3. Pakdaman Naeini, Mahdi & Taremian, Hamidreza & B. Hashemi, Homa. **Stock market value prediction using neural networks**, *Computer Information Systems and Industrial Management Applications (CISIM)*,Vol. 4, pp. 132 – 136, 2010.
4. Yaojun Wang, Yaoqing Wang, **A Case-Based Reasoning-Decision Tree Hybrid System for Stock Selection**, *International Journal of Computer and Information Engineering, World Academy of Science, Engineering and Technology*, Vol.10, No:6, 2016.
5. Clive W.J. G-anger, **Forecasting stock market prices: Lessons for forecasters**, *International Journal of Forecasting*, Vol.8, pp. 3-13, 1992. https://doi.org/10.1016/0169-2070(92)90003-R
6. H. Luo and M. Wang. **Statistical Analysis of the Characteristics of Stock Returns in China's Securities Market,** *Second International Conference on Networking and Distributed Computing*, Vol.4, No.2, pp. 197-200, 2011. https://doi.org/10.1109/ICNDC.2011.47
7. Ding, G., Qin, L. **Study on the prediction of stock price based on the associated network model of LSTM**, *International Journal of Machine Learning & Cybernetics*,Vol.30, No..11, 2019.
8. Lai, Kin Keung. **Generating Profit Using Option Selling Strategies.** *Proceedings of the fifth International Conference on Business intelligence and financial engineering*, Vol. 11, No.9, pp. 56-64, 2012.
9. Dattatray P, Gandhmal, Kumar K. **Systematic analysis and review of stock market prediction techniques**, Computer science review, *Elsevier*, Vol. 34, 2019.

10. White H. **Economic prediction using neural networks: The case of IBM daily stock returns**. Neural networks in finance and investing. *Chicago, Illinois, Cambridge, England: Probus Publishing Company*, Vol. 2, pp.315-329, 1993.
11. Nikfarjam, Azadeh & Emadzadeh**. Text mining approaches for stock market prediction**. *The 2nd International Conference on Computer and Automation Engineering, ICCAE*, Vol.4, pp 256-260, 2010. https://doi.org/10.1109/ICCAE.2010.5451705
12. Huang W., Wang S., Yu L., Bao Y., Wang L. **A New Computational Method of Input Selection for Stock Market Forecasting with Neural Networks.** *Computational Science – ICCS 2006. Lecture Notes in Computer Science, Springer Berlin, Heidelberg*, Vol. 3994, pp.308-315, 2006.
13. S. Sable, A. Porwal and U. Singh. **Stock price prediction using genetic algorithms and evolution strategie**s, International conference of Electronics, Communication and Aerospace Technology (ICECA), *Coimbatore*, pp. 549-553. 2017. https://doi.org/10.1109/ICECA.2017.8212724
14. B. Pang and L. Lee, **Opinion mining and sentiment analysis,** *Foundations and Trends in Information Retrieval*, Vol. 2, pp. 1-135, 2008.
15. Wiebe, Hoffmann. **Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis**, *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pp,347-354, 2005.
16. Wiebe, Hoffmann. **Recognizing Contextual Polarity: A Exploration of Features for Phrase-Level Sentiment Analysis**, *Journal of Computational Linguistics*, Vol. 35, No.3, 2009.
17. Ranco, Gabriele & Aleksovski, Darko. **The Effects of Twitter Sentiment on Stock Price Returns**. (PLOS ONE), 2015. https://doi.org/10.1371/journal.pone.0138441
18. Sandner, Phillipp & Welpe, **predicting elections with twitter**, Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, 2010.
19. La sheng yu. **A sentiment analysis approach based on Arabic social media platforms**, *International conference on mechanical, electronic and information technology (ICMEIT*), Vol. 8, No.6, pp.345-352, 2018.
20. A.Sasi Kumar and S.P.Tamizhselvi. **Qos-Mindful Channel Allocation in Directional Wireless Sensor Networks**, *International Journal of Engineering and Advanced Technology (IJEAT)*, Vol. 8, No.5, June 2019.
21. A.Sasi Kumar**,** G.V.Sriramakrishnan. **Internet of Things Based Clinical Decision Support System Using Data Mining Techniques**, *Journal of Advanced Research in Dynamical and Control Systems*, Vol. 10, No. 4, pp.132-139, 2018.
22. Tawarish, M. & Satyanarayana, **An enabling technique analysis in Data Mining for Stock Market trend by Approaching Genetic Algorithm,** *International Journal of Advanced Trends in Computer Science and Engineering Vol.8(1) pp.27-33,2019.* https://doi.org/10.30534/ijatcse/2019/06812019
23. Haur Teng,king Kouk, **Development of Whale Optimization Neural Network for Daily Water Level Forecasting,** *International Journal of Advanced Trends in Computer Science and Engineering. Vol.8(3) pp.354-362,2019.* https://doi.org/10.30534/ijatcse/2019/04832019