Financial Trends prediction using the Back Propagation Neural Network and YQL

Trapti Mishra
Institute of Engineering & Technology,
Devi Ahilya University, Indore

V. Kapoor
Institute of Engineering & Technology,
Devi Ahilya University, Indore

ABSTRACT
In this world a amount of things and procedures are impulsive, among the stock market value is an entity which is impulsive. The stock market and their prices are haphazard in impulsive manner that can be depends on amount of external associations such as politics, terror, social and others. But a number of researchers are building efforts to forecast the stock market values. To forecast the upcoming stock market trends the data mining and machine intelligence based techniques are employed to forecast precisely the worth. It works too precisely estimate the stock value a back propagation neural network based system is used. The proposed system is an enhanced version of the traditionally available back propagation neural network. The back propagation neural network is a supervised learning advance therefore the training and testing both are required. To prepare the proposed data model the financial records of the companies are required. Thus for finding the historical financial market trends the YQL (yahoo query language) is proposed for use. This YQL make available the historical records of the target company between two dates. Using the acquired data from YQL is used to train the BPN network. After training the implemented classifier is used for forecasting the upcoming stock value. The implementation of the proposed system is performed using the JAVA technology. Additionally to justify the obtained results the performance of the proposed system is compared with the traditional BPN model. The relative performance study is performed using the accuracy, error rate, and memory and time consumption. According to the obtained results the performance of the proposed system is found improved and adoptable.

Keywords
Stock market prediction, Neural network, YQL, data mining, machine learning.

1. INTRODUCTION
In current financial world, Stock Market forecasting is considered as one of the most exigent tasks. So a lot of consideration has been given to study and forecast upcoming values and behavior of financial time series. Diverse issues intermingle in stock market such as business cycles, interest rates, monitory policies, general economic conditions, traders' expectations, political events, etc. According to scholastic investigations, activities in market value are not haphazard rather they act in a highly non-linear, dynamic manner [1].

Expertise to forecast direction and correct value of upcoming stock market values is the most important issue in financial market to make money. These days for the reason that of online trading, stock market has turn out to be one of the burning targets where someone can make profits. So forecasting the exact value and performance of stock market has become the area of interest. Though, because of high explosive nature of fundamental laws behind the financial time series, it is not any easy task to assemble such a forecasting system. As mentioned earlier, stock market forecasting has been one of the burning

topics in the middle of researchers over the years. As a consequence, a lot of researches are carried out and many forecasting systems have been projected. The most common advance in use so far is artificial neural networks (ANNs). But using genetic encoding in this ground is pretty new notion as most studies showed that ANN has some boundaries in learning the patterns because stock market data has remarkable noise and intricate dimensionality [2].

Moreover, ANN has unsurpassed learning ability while it is often confronted with contradictory and impulsive performance for noisy data. In addition, sometimes the quantity of data is so large that the learning of pattern possibly will not effort as well. In meticulous, the survival of nonstop data and large quantity of data may pretence a exigent task to explicit concepts extraction from raw data due to enormous sum of data space determined by continuous features [3].

2. PROPOSED TECHNIQUE
This section provides the understanding about the proposed working model for accurate stock market price forecasting. Therefore the detailed system implementation is described in this chapter with the implementable algorithm steps.

2.1 Domain overview
Stock market prediction or forecasting is an endeavour for estimation of future worth of an organization's or company stock or financial exchange. The forecast of a future stock's value can produces a profit but that can also be unpredictable. The stock market proficient who knows about the market trends is bothered about the forecasted values of market. According to the assessment of the stock market movement the market is spontaneous and can be fluctuate according to the politics, nature and other momentous social and commercial aspects. But in prose there are amount of practices are available who claims to forecast the stock market value more specifically. Among most of the system exploit the notion of data mining learning algorithms.

In this offered work the data mining methods are used for forecasting the stock market values. In order to forecast precisely the value the supervised classification method is used. Previous to the selection of the proposed classification method five different techniques are examined for verdict their accuracy and among them the back propagation neural network is found most frequently used and valuable technique for forecast system development. On the other hand the supervised learning techniques are required to be train with the historical data sample. Therefore the Yahoo financial database based historical stock trends are confined and used for training the selected classifier. After training of the proposed classifier capable to accept the current market values and forecast the definite or nearer values of the stock prices.

This section provides the overview of the proposed stock market forecasting system and the next section involves the detailed description of the proposed methodology.

2.2 Methodology

The proposed methodology for the proposed solution development can be defined using the figure 1. This diagram contains the different components that are used to design the entire forecasting system.
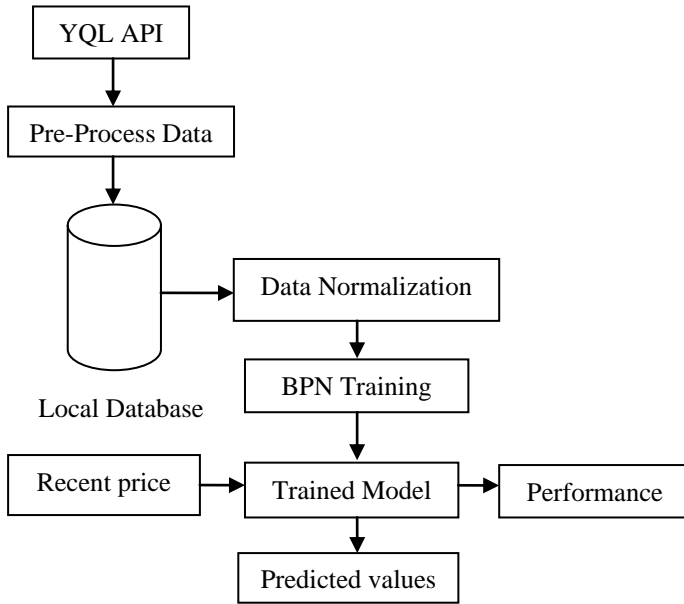

Fig 1: Proposed system architecture

YQL API: Every supervised learning algorithm requires being train earlier than use of the classifier for forecasting. Therefore, historical data required to learn the algorithm. In proposed system the financial historical data is essential therefore Yahoo financial records are takeout using the YQL (yahoo query language). In this system the custom query is fired on the Yahoo data based and that provides the relevant stock's data between two dates.

Pre-processing: The pre-processing is a vital step of the machine learning and data mining. In the pre-processing procedure the data is making clean and proper to adopt with the algorithms. The data extracted straight from the Yahoo server contains a number of useless attributes also therefore the pre-processing practice are utilized for making it clean for extraction of the requisite data attributes.

Local Database: After pre-processing of the data, it is cleaned and requisite attributes are extracted form dataset. These attributes are preserved in a local database for further use with the learning algorithms.

Data Normalization: The received data is in different scales of commerce. Therefore its necessary to scale in a similar scale for better understanding about the hidden patterns of the participating attributes. In the proposed system the data is normalized for scaling the attribute values between 0-1. Therefore the following formula is used for normalization process.

$$v^{'} = \frac{v - min}{max - min}$$

Where $v^{'}$ is the new scaled value between 0- 1

V= the actual dataset value

Min = the minimum value that contains by a given attributes

Max = the maximum value in the defined attribute

BPN Training: The execution of neural network is defined in two phases' first training and second prediction: training process utilizes data and designs the data model. By this data model next stage forecasting values is performed [23].

Training steps:

2.2.1 Prepare two arrays, one is for input and hidden unit and the second is for output unit.

2.2.2 Here first is a two dimensional array W_{ij} and output is a one dimensional array Y_i.

2.2.3 Original weights are random values placed inside the arrays after that the output is given as**.**

$$x_j = \sum_{i=0} y_i W_{ij}$$

Where, y_i is the activity level of the jth unit in the previous layer and W_{ij} is the weight of the connection between the ith and the jth unit.

2.2.4 Next, action level of y_i is projected by sigmoidal function of the total weighted input.

$$y_i = \left[\frac{e^x - e^{-x}}{e^x + e^{-x}} \right]$$

When occurrence of the all output units have been determined, the network calculates the error (E) given in equation**.**

$$E = \frac{1}{2} \sum_i (y_i - d_i)^2$$

Where, y_i is the event level of the jth unit in the top layer and d_i is the preferred output of the j_i unit.

Calculation of error for the back propagation algorithm is as follows:

- Error Derivative (EA_j) is the alteration among the real and desired target:

$$EA_j = \frac{\partial E}{\partial y_j} = y_j - d_j$$

- Error Variations is total input received by an output changed

$$EI_j = \frac{\partial E}{\partial X_j} = \frac{\partial E}{\partial y_j} X \frac{dy_j}{dx_j} = EA_j y_j (1 - y_i)$$

- In Error Fluctuations computation connection into output unit is required:

$$EW_{ij} = \frac{\partial E}{\partial W_{ij}} = \frac{\partial E}{\partial X_j} = \frac{\partial X_j}{\partial W_{ij}} = EI_j y_i$$

- Overall Influence of the error:

$$EA_i = \frac{\partial E}{\partial y_i} = \sum_j \frac{\partial E}{\partial x_j} X \frac{\partial x_j}{\partial y_i} = \sum_j EI_j W_{ij}$$

Trained Model: After learning with the specified financial data set the back propagation neural network turn out to be trained and now this system can be used for forecasting the upcoming financial trends.

Recent Price: Forecasting the new values need an input the recent or just last value of financial pattern, based on the patterns calculated in the trained system the back propagation neural network forecast upcoming values.

Performance: During estimation of the new data pattern for forecasting the performance of classifier is cross validated to find the forecast precision. The obtained results of the classifier's performance are stock up in database for upcoming results analysis. Therefore in this pace four parameters are computed namely accuracy, error rate, memory used and the time consumption of the system.

Predicted values: These values are new values that are forecasted by the trained data model, and are the final outcome of the proposed system.

3. PROPOSED ALGORITHM

The proposed system of forecasting the stock market price can be summarized using the suitable steps for understanding. The table 1 shows the summarize algorithm for the proposed concept.

Table 1. Proposed Algorithm

Input: Dataset D, learning cycle $epoch$, Recent pattern R_p

Output: predicted value P

Process:

1. $R_d = ReadDataset(D)$

2. $[row, col] = normalizeData(R_d)$

3. $Activation = \frac{e^x - e^{-x}}{e^x + e^{-x}}$

4. $T_{model} = BPN.Train(R_d, Activation, epoch)$

5. $P = BPN.Predict(T_{model}, R_p)$

6. $Return\ P$

4. RESULT ANALYSIS

The given chapter provides the results expansion and the performance analysis of the proposed and base algorithm. Therefore to compare the performance essential performance factors are evaluated and their results are reported.

4.1 Accuracy

The accuracy of a extent system is the unit of proximity of amounts of a quantity to that quantity's accurate value i.e. we can say that an Artificial Neural Network have classifier to classify the sets of data that most of the data are perfect in output pattern representing correctness of the neural network. The accuracy of the proposed technique is evaluated using the following formula.

$$Accuracy = \frac{Total\ correctly\ identified\ patterns}{Total\ Patterns\ to\ classify} X100$$
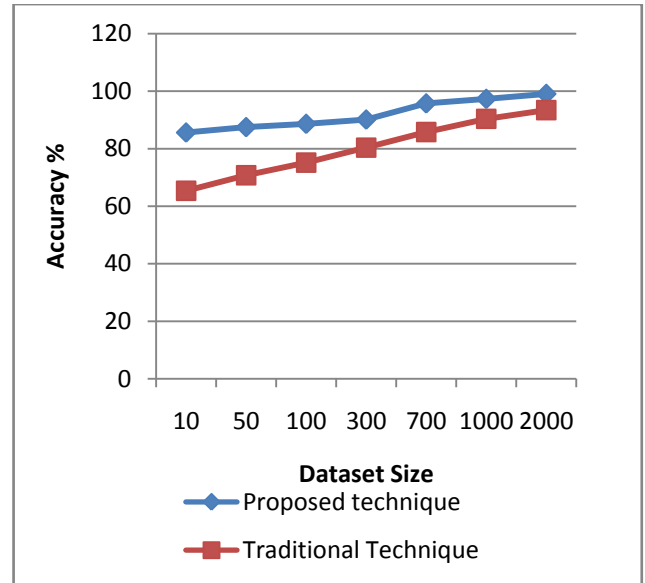

Fig 2: Accuracy Performance

In above figure 2shows the comparison of the proposed and base system. X-axis obtain amount of data files and Y-axis shows accuracy in terms of percentage. Additionally to represent the performance of both the implemented classifiers are reported thus the blue line shows the performance of classical approach and the red line shows the performance of the proposed approach. By obtaining performance of the graph proposed system gives high accuracy and precise data efficiently while base method less accurate for the final results.

Table 2. Accuracy Performance

Dataset Size	Proposed Method	Base method
10	85.63	65.37
50	87.51	70.77
100	88.65	75.16
300	90.13	80.38
700	95.74	85.81
1000	97.33	90.39
2000	99.05	93.44

4.2 Error Rate

Error rate is the measurement of the classifier performance in terms of misinterpreted rate. In other words the amount of data which is not properly recognized using the trained classifier is termed as the error rate of the system. That can be evaluated using the following formula:

$$Error\ Rate = 100 - accuracy$$

Or

$$Error\ Rate$$
$$= \frac{Total\ Incorrectly\ Classified\ pattern}{Total\ Pattern\ to\ classify} X100$$
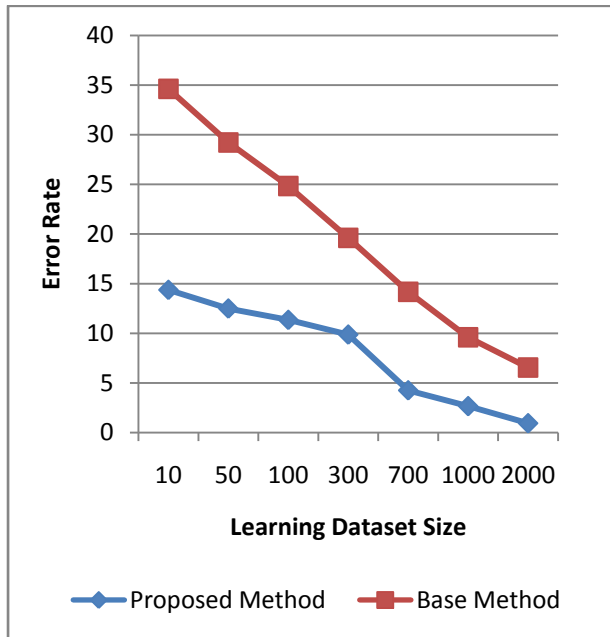

Fig 3: Error Rate

The graph 3 represents the Error rate of implemented Artificial Neural Network systems and traditional artificial neural network. In this graph the blue line represents the error rate of the proposed technique and red lines of traditional algorithms. The X-axis of the given graph covers the data stored in the database for evaluation and concluding error rate that means amount of patterns not executed appropriately and that produces error data and the Y-axis shows how much error rates are obtained by implementing both scenarios. According to the obtained result error in percentage is very fewer to base neural method and the proposed system conveys the minimization of data error.

Table 3. Error rate

Dataset Size	Proposed Method	Base method
10	14.37	34.63
50	12.49	29.23
100	11.35	24.84
300	9.87	19.62
700	4.26	14.19
1000	2.67	9.61
2000	0.95	6.56

4.3 Time Consumption

The total time required to learn the given patterns from the storage using the selected classifier is known as the time consumption of the system. The graph 4 contains the time consumption of the base algorithm and the proposed ANN based algorithm in terms of seconds. To represent the performance of the classifier the proposed technique is established using the red line graph and the blue line shows the performance of the traditional approach of learning. According

to the given results the proposed technique consume less time for training of input of stock market as compared to the base method. Therefore the proposed system is less time consuming for prediction of the data patterns.
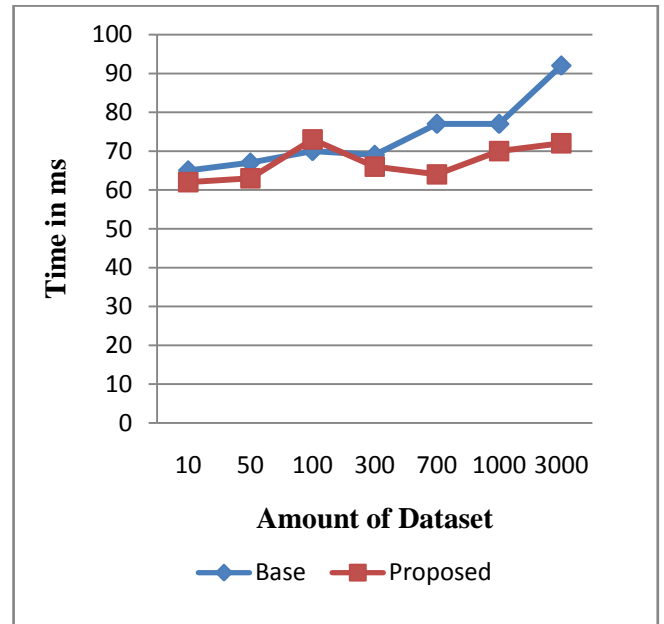

Fig 4: Time Consumption

4.4 Memory Consumption

The amount of main memory required to execute the algorithm is termed as the space complexity of the system. That is occasionally also called the memory consumption of the algorithms. The comparative memory consumption of both the algorithms namely Base method and Artificial Neural network based Proposed Algorithm is given using figure 5. In this diagram the X-axis includes amount of dataset stored in database, and the Y-axis shows the amount of memory utilized during processing of the data. The given memory utilization is provided here in terms of kilobytes. According to the evaluated results the performance of the traditional technique is less efficient as long as slow performance as compared to the proposed system.
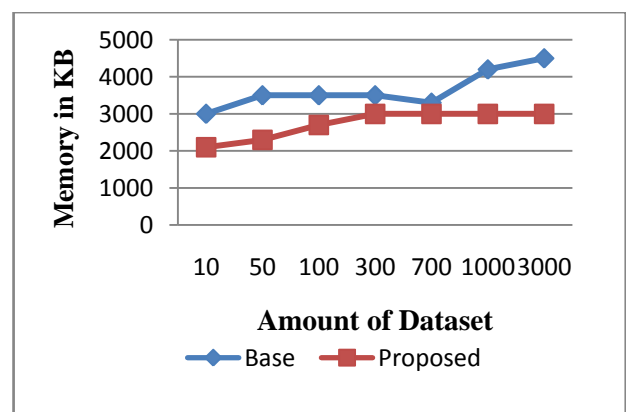

Fig 5: Memory Consumption

5. CONCLUSION S & FUTURE WORK

The proposed work is provoked to design an improved supervised learning system for forecasting the stock market values. This section provides the summary of the entire hard

work made for designing of the proposed system. In addition of that the future extension of the proposed work is also provided.

5.1 Conclusion

The data mining and machine learning practice are used in a number of purposes for automatic data analysis and recovery of essential patterns. These patterns are used to approximate the future trends of the business and other functional area of purposes. The data mining practice supports the analysis of both kinds' i.e. supervised and unsupervised learning techniques. The key advantage to employ the supervised learning approaches for the predictive data analysis is their accuracy as compared to unsupervised learning approaches. In the proposed work the supervised learning technique is employed and the small modification is performed for improving the existing classifiers performance. Additionally their functional area is reported in stock market price forecasting application.

The proposed system involves the artificial neural network technique for pattern learning of stock market trends. Therefore a considerable amount of learning data is required to perform learning of the proposed classifier. To train the historical patterns of the proposed classifier the YQL (yahoo query language) for financial data base is used. This is basically the yahoo based API (application programming interface) for finding the historical records of the stock exchange. This data is further pre-processed and the requisite patterns are stock up on the local data base. Further the proposed neural network is applied for training. In the proposed neural network system two main modifications are projected first change in the activation function and secondly the use of normalized data to scale up the training values into a predefined range. Through which the learning capability of the learning algorithm is enhanced.

The implementation of the proposed system is performed using the JAVA technology and with the help of YQL API. After implementation of the proposed technique a traditional back propagation algorithm is also implemented to compare the performance with the proposed modified system. The comparative analysis of both the implemented classifiers is performed with the help of accuracy, error rate, time consumption and space complexity. The obtained performance summary of the both the techniques are demonstrated below using table 2:

Table 4. Performance Summary

S. No.	Parameters	Proposed	Traditional
1	Accuracy	High	Low
2	Error rate	Low	High
3	Time complexity	Low	High
4	Space complexity	Similar	Similar

The proposed system is adoptable due to high precise outcomes and efficient learning potential. Additionally consumes less time and space for training and forecast. Thus the proposed data model is just right and efficient for stock market forecasting.

5.2 Future extension

The proposed system is a task of a precise stock market forecasting technique execution. Therefore a back propagation neural network based technique is employed and their performance analysis is performed. According to the obtained performance the proposed technique produces high accurate outcomes with less resource consumption. On the other hand

the proposed technique is an offline prediction technique, therefore in near future the following modification proposed to incorporate with the proposed technique.

5.2.1 Stock market depends on diverse real world factors also therefore in near future the different real world factors are also suggested to include for getting better predictive accuracy of the system.

5.2.2 Stock market also varies with the news and their impact on the active economy therefore that is must to include the opinion mining concepts for enhancing the prediction of the existing system.

6. REFERENCES

[1] Choudhary, Rohit&Garg, Kumkum, "A Hybrid Machine Learning System for Stock Market Forecasting", World Academy of Science, Engineering and Technology, Vol. 39, PP. 315-318, 2008

[2] Han, Ingoo& Kim, Kyoung-jae, "Genetic algorithms Approach to feature Discretization in artificial neural networks for the prediction of stock price index", Expert Systems with Applications, Vol. 10, No. 5, PP. 120-122., 2000.

[3] Mahesh S. Khadka, K. M. George, N. Park and J. B. Kim, "Performance Analysis of Hybrid Forecasting Model in Stock Market Forecasting", International Journal of Managing Information Technology (IJMIT), Vol.4, No.3, August 2012.

[4] ZabirHaider Khan, TasnimSharminAlin, Md. AkterHussain, "Price Prediction of Share Market using Artificial Neural Network (ANN)", International Journal of Computer Applications (0975 – 8887) Volume 22–No.2, May 2011

[5] KainazBomiSheriwala, "Data Mining Techniques in Stock Market", INDIAN JOURNAL OF APPLIED RESEARCH, Volume 4, August 2014.

[6] S. L. Pandhripande and Aasheesh Dixit, "Prediction of 2 Scrip Listed in NSE using Artificial Neural Network", International Journal of Computer Applications (IJCA), Volume 134, No.2, January 2016.

[7] Data Mining - Classification & Prediction, available online: http://www.tutorialspoint.com/data_mining/dm_classification_prediction.htm, [accessed 27 April 2016].

[8] Samuel OdeiDanso, "An Exploration of Classification Prediction Techniques in Data Mining: The insurance domain", Master's Degree in Advanced Software Engineering, School of Design, Engineering, and Computing. Bournemouth University, September, 2006.

[9] Dr. B. Srinivasan and K. Pavya, "A STUDY ON DATA MINING PREDICTION TECHNIQUES IN HEALTHCARE SECTOR", International Research Journal of Engineering and Technology (IRJET), PP. 552-556, Volume 3, Mar-2016

[10] Vipin Kumar, JoydeepGhosh and • David J. Hand, "Top 10 algorithms in data mining", Knowledge and Information System, PP. 1–37, (2008).

[11] Vapnik V (1995), the nature of statistical learning theory. Springer, New York.

[12] Kavitha G, Udhayakumar A and Nagarajan D, "Stock Market Trend Analysis Using Hidden Markov Models",

available   online:
https://arxiv.org/ftp/arxiv/papers/1311/1311.4771.pdf.

[13] Haiqin Yang, Laiwan Chan, and Irwin King, "Support Vector Machine Regression for Volatile Stock Market Prediction", Intelligent Data Engineering and Automated Learning IDEAL, PP. 391- 396, Springer-Verlag Berlin Heidelberg 2002

[14] Investopedia US, A Division of Value Click, Inc., "Investopedia" "http://www.investopedia.com /articles/05/032905.asp",05 March 2013.

[15] Mrs. Keerti. S. Mahajan 1 & R. V. Kulkarni, "A REVIEW: APPLICATION OF DATAMINING TOOLS FOR STOCK MARKET", International Journal Computer Technology & Applications, Volume 4, PP. 19-27, 2013.

[16] TIM BERRY, "What Is a Market Forecast, available online: http://articles.bplans.com/what-is-a-market-forecast/, [accessed on 27 April 2016].

[17] Selvan Simon and ArunRaoot, "ACCURACY DRIVEN ARTIFICIAL NEURAL NETWORKS IN STOCK MARKET PREDICTION", International Journal on Soft Computing (IJSC), Vol.3, No.2, May 2012.

[18] D. VenugopalSetty, T.M.Rangaswamy and K.N.Subramanya, "A Review on Data Mining Applications to the Performance of Stock Marketing", international Journal of Computer Applications (IJCA), Volume 1, No. 3, PP. 24-34, 2010.

[19] ShradhaTulankar , Dr Rahul Athale and SandeepBhujbal, "Sentiment Analysis of Equities using Data Mining Techniques and Visualizing the Trends", IJCSI International Journal of Computer Science Issues, Vol. 10, No 2, July 2013.

[20] AkinwaleAdio T, Arogundade O.T. and Adekoya Adebayo F., "TRANSLATED NIGERIA STOCK MARKET PRICES USING ARTIFICIAL NEURAL NETWORK FOR EFFECTIVE PREDICTION", Journal of Theoretical and Applied Information Technology, PP. 36-43.

[21] Forecasting of Indian Stock Market Index Using Artificial Neural Network, [Online], http://nseindia.com/content/research/proposal_206.pdf, [accessed: 19 April 2016].

[22] ZabirHaider Khan, TasnimSharminAlin, and Md. AkterHussain, "Share Market Price Prediction Using Artificial Neural Network (ANN)".

[23] ArkaGhosh, "Comparative study of Financial Time Series Prediction By Artificial Neural Network with Gradient Descent Learning".

[24] Kyoung-jae Kim, "Artificial neural networks with evolutionary instance selection for financial forecasting", Expert Systems with Applications, PP. 519 – 526, 2006.

[25] ShwetaTiwari, RekhaPandit and VineetRichhariya, "Predicting future trends in stock market by decision tree rough-set based hybrid system with HHMM", International Journal of Electronics and Computer Science Engineering , PP. 1578 – 1587, 2012.

[26] Selvan Simon and ArunRaoot, "ACCURACY DRIVEN ARTIFICIAL NEURAL NETWORKS IN STOCK MARKET PREDICTION", International Journal on Soft Computing (IJSC), Volume3, No.2, May 2012.

[27] Mr. Amit B. Suthar, Ms. Hiral R. Patel, and Dr. Satyen M. Parikh, "A Comparative Study on Financial Stock Market Prediction Models", the International Journal of Engineering and Science (IJES), Volume 1 PP. 188-191, 2012

[28] Sam Mahfoud and Ganesh Mani, "FINANCIAL FORECASTING USING GENETIC ALGORITHMS", Applied Artificial Intelligence, PP. 543 - 565, 1996.

[29] Wanjawa, Barack Wamkaya, Muchemi and Lawrence, "ANN Model to Predict Stock Prices at Stock Exchange Markets", available online: https://arxiv.org/ftp/arxiv/papers/1502/1502.06434.pdf [accessed 28 April 2016]..