ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

Stock Credibility Prediction Using Multilayer Perceptron and Statistical Computational Methodologies

A. Rao[1], S. Hule[2], H. Shaikh[3], E. Nirwan[4], Prof. P. M. Daflapurkar[5]

UG Student, Department of Computer Engineering, MMIT, Pune, Maharashtra, India[1, 2, 3, 4]

Head of Department, Department of Computer Engineering, MMIT, Pune, Maharashtra, India[5]

ABSTRACT: This paper proposes a system aimed at forecasting stock values and deducing their credibility using a Stock Market Prediction Model. The system is modeled on the efficacy of Feed-Forward Artificial Neural Network Algorithm, namely, Multilayer Perceptron. The system additionally uses Forecasting Algorithm, Moving Averages Algorithm and Linear Regression Algorithm. The results of each algorithm are compared with Real Time Data, sourced from online financial repositories, in order to obtain a prediction price with the highest accuracy.

KEYWORDS: Artificial Neural Networks; Multilayer Perceptron; Linear Regression; Forecasting; Moving Averages

I. Introduction

Stock Market Prediction deals with the various approaches that are used to forecast the value of a company's stock. Although, the Efficient Market Hypothesis considers stock prices to be unpredictable as they are a function of revealed information and rational expectations, targeted studies have refuted this claim. Auto-Regressive Integrated Moving Averages (ARIMA) Model, for forecasting the future stock prices, is the key approach that has reinvigorated the research in the applications of Statistical Computational Methodologies to build financial forecasters [1]. On the other hand, there is a growing interest in Artificial Neural Networks (ANN) as the preferred choice for the development of stock prediction models. The dynamic complexity of stock prediction that arises due to the non-linear relationship between the market parameters and the closing price is compensated well by utilizing a Back Propagation algorithm to train a Multilayer Feed-Forward ANN for prediction. Both the Statistical Computational Approach and the ANN-driven Approach provide efficient results, although, with varying levels of implementation complexity and prediction accuracy. Some may infer that if the two approaches are used interchangeably, the chances of accurately predicting the credibility of the stock values may increase [6]. Drawing from this inference, we propose a system, namely, *Efficiency Filter Prediction*. This Stock Market Prediction Model utilizes the two approaches in the form of their algorithmic implementations, namely – *Multilayer Perceptron (MLP)*, *Linear Regression*, *Moving Averages*, and *Forecasting Algorithm*. The system filters an algorithm with the highest efficiency which, thereafter, determines the credibility of the stocks owned by the end-user by making a call (decision). The system routes the result to the end-user through a web-based Graphical User Interface (GUI).

II. Related Work

We have carried out an extensive survey of the research work, the proposed systems and the implemented predictive models that have employed ANN and conventional statistical computational algorithms to forecast financial data, specifically, in the realm of stock market analysis [1-15]. Our survey is divided into two parts. The first part of the survey comprehensively outlines the methodologies using which financial forecaster models have been developed and tested. It covers the two standard approaches as mentioned earlier. It also investigates the attempts that have gone into creating, deploying and testing hybrid prediction models that incorporate both the approaches in an integrated capacity to lessen the forecast error rates. The second part of the survey concentrates on establishing the reasons behind the characteristic efficacy of MLP in forecasting market trends and stock movements by reviewing the architectural setups for data generation, data processing and fiscal trend analysis of the implemented systems.

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

III. Proposed System Architecture

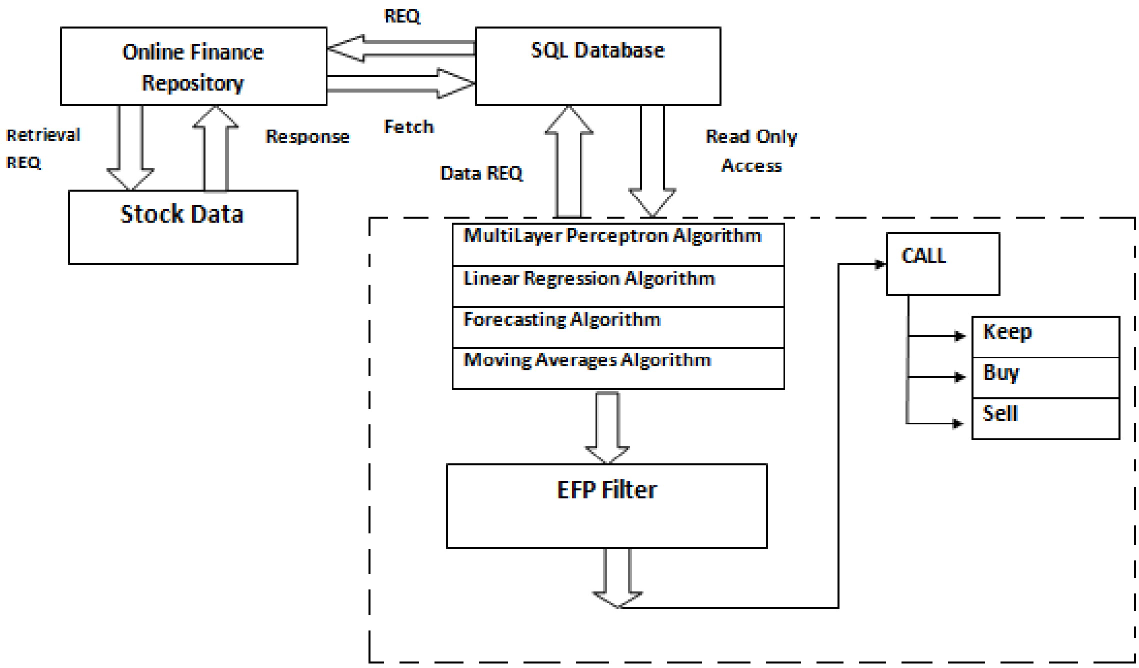

Fig.1. System Architecture

The GUI has been developed using Java Enterprise Edition (J2EE). The System can be primarily divided into three distinct but inter-related modules namely: Real Time Data Fetch, Data Analysis, and Stock Credibility Prediction. The output of the preceding module acts as input for the succeeding module. The modules can be described as follows:

A. *Real Time Data Fetch:* This module fetches the data in Real Time from an Online Financial Repository e.g., Yahoo Finance. The data is highly scalable. Its size can be customized as per the developer's requirements. In our network configuration, 70% of the data is used as the training set whereas the remaining part is used as the test set. It is to be noted that the percentage allotted for the training set and the test set can also be customized as per requirement.

B. *Data Analysis:* The implementations of the four previously mentioned system algorithms are granted read-only access to the training set. The algorithms function concurrently and access the database in a mutually exclusive manner with the help of high-level synchronization primitives, thereby, avoiding resource contention. The system architecture is built on the idea of *Efficiency Filter Prediction*. Each algorithm predicts the stock price attributes for each and every record of the training set by taking into consideration a pre-defined range of the time-series data set. The predicted value is compared with the original value of the stock attributes for the respective records and points are allotted to an algorithm for every correct prediction. In the end, the algorithm with the highest hit rate of correct predictions has the highest accumulated points, thereby, naturally having the highest efficiency for that range of training data. Thus, this module filters an algorithm based on its efficiency and hands over the control to it.

C. *Stock Credibility Prediction:* The filtered algorithm gets mutually exclusive read-only access to the test set and makes a call in terms of Stock Credibility. Specifically, it decides whether the user should *Buy*, *Sell* or *Keep* the stocks. The results are routed to the end-user in the form of tabular as well as graphical representations.

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

IV. MATHEMATICAL MODEL

Let S be the Stock Prediction System such that S = {i, e, U, C, O, D, F_1, F_2, F_3, F_4, F_5, F_6, F_7, DD, NDD} where,

- i = Initial State
- e = End State
- Set of System Users (U) = {U_1, U_2, U_3…. U_n}
- Set of User Credentials (C) = {C_1, C_2}
- Set of Outputs (O) = {Buy, Sell, Keep}
- Set of Fetched Stock Data (D) = {D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8}
- F_1 = Function for User Validation
- F_2 = Function for Stock Data Fetch
- F_3 = Function for Stock Market Prediction using Multilayer Perceptron Algorithm
- F_4 = Function for Stock Market Prediction using Linear Regression Algorithm
- F_5 = Function for Stock Market Prediction using Forecasting Algorithm
- F_6 = Function for Stock Market Prediction using Moving Averages Algorithm
- F_7 = Efficiency Filtered Algorithm that computes outcome O
- DD = Deterministic Data – Static System Identifiable Data
- NDD = Non-Deterministic Data – Dynamic Runtime Generated Data

D, being the set of fetched stock data, has eight parameters, D_1 – D_8, which are consistent with the eight attributes of the stock data fetched from the online finance repository. The Main Control Flow Function (F_{main}) is concerned with the sequential execution of the system from start to end. The various functional dependencies relevant to the system are shown in the following figures:

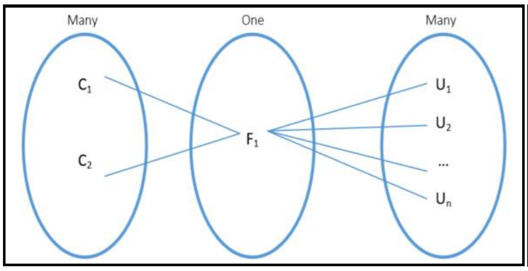

Fig.2. Login Validation Module

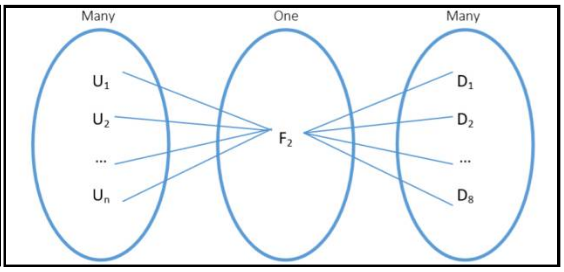

Fig.3. Data Fetch Module

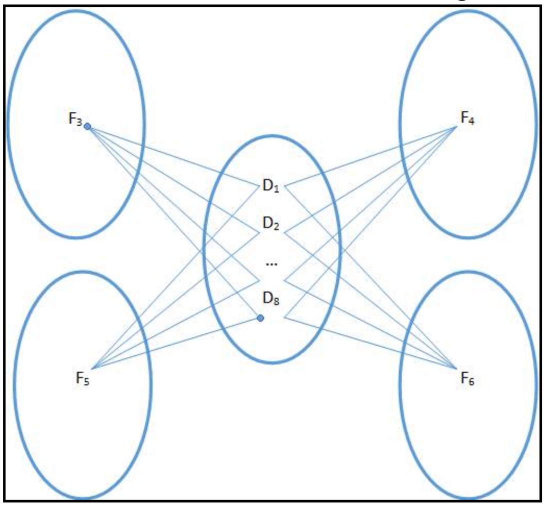

Fig.4. Data Analysis Module

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

V. DATASET AGGREGATION

The dataset has been sourced from Yahoo Finance which is an online financial repository. It targets 16 companies listed under the Nifty Stock Index. The data is fetched in real time after each successful user validation and is stored in MySQL Relational Database. The dataset is accessed using MySQL QueryBrowser. The fetched data is categorized by eight attributes which are as follows: Date, Open Price, High Price, Low Price, Close Price, Volume, Adjusted Close Price and Symbol ID. These parameters, taken one tuple at a time, define the stock data for the corresponding stock ID. The tuple attributes and their corresponding column attribute headings are shown in Tab.1.

Parameters	Date	Open Price	High Price	Low Price	Close Price	Volume	Adjusted Close Price	Symbol ID
Corresponding Attribute Headings	sdate	sopen	shigh	slow	sclose	svolume	sadjclose	symbid

Tab.1. Sample Stock Data Tuple Attributes

VI. SYSTEM RESULTS

The decision call outputs and the prediction efficiency of each algorithm, for the 16 companies, shown in Tab.2., are observed and recorded for 27th April, 2016. The variations between the actual and the predicted stock prices of Google for the same day are shown graphically.

Company Details		**Algorithm Key Statistics**								**Efficiency Call**
Company Name	**Company ID**	**Linear Regression**		**Moving Averages**		**Forecasting**		**Multilayer Perceptron**		
		E	C	E	C	E	C	E	C	
ACC LTD.	ACC	27.27	Wait	40.91	Wait	13.64	Sell	27.27	Buy	Wait
SPICEJET LTD.	SPICEJET.BO	39.13	Sell	43.48	Sell	39.13	Buy	39.13	Buy	Sell
TCS	TCS	50	Wait	54.55	Wait	22.73	Sell	22.73	Buy	Wait
TATA MOTORS LTD.	TATAMOTORS.BO	30.43	Buy	34.78	Buy	17.39	Sell	56.52	Buy	Buy
UNITECH LTD.	UNITECH.BO	43.48	Buy	43.48	Buy	8.7	Sell	13.04	Buy	Buy
TVS MOTOR COMPANY LTD.	TVSMOTOR.BO	34.78	Sell	34.78	Sell	26.09	Buy	26.09	Sell	Sell
JK TYRE LTD.	JKTYRE.BO	30.43	Buy	30.43	Wait	27.14	Sell	39.13	Buy	Buy
UNION BOI	UNIONBANK.BO	21.74	Buy	30.43	Buy	13.04	Sell	43.48	Buy	Buy
ONGC	ONGC.BO	39.13	Buy	39.13	Buy	26.09	Sell	39.13	Buy	Buy
JINDAL STEEL & POWER LTD.	JINDALSTEEL.BO	39.13	Buy	39.13	Wait	26.09	Sell	43.48	Buy	Buy
ICICI BANK	ICICIBANK.BO	30.43	Buy	43.48	Buy	34.78	Sell	43.48	Buy	Buy
BAJAJ HINDUSTHAN SUGAR LTD.	BAJAJHIND.BO	21.74	Wait	21.74	Wait	4.35	Buy	21.74	Buy	Buy
AXIS BANK	AXISBANK.BO	43.48	Buy	47.83	Buy	30.43	Sell	30.43	Sell	Buy
GOOGLE	GOOG	50	Buy	54.55	Buy	45.45	Buy	45.45	Sell	Buy
IDBI BANK LTD.	IDBI.BO	34.78	Buy	34.78	Buy	21.74	Sell	43.48	Buy	Buy
RELIANCE	RELIANCE.BO	26.09	Buy	30.43	Buy	21.74	Buy	21.74	Sell	Buy

Tab.2. Result Table

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

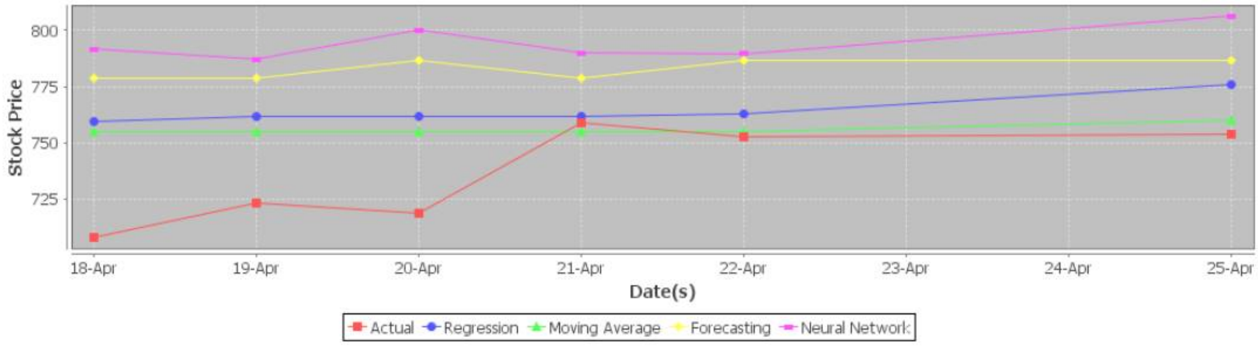

Fig.5. Graphical Visualization of Predicted Values

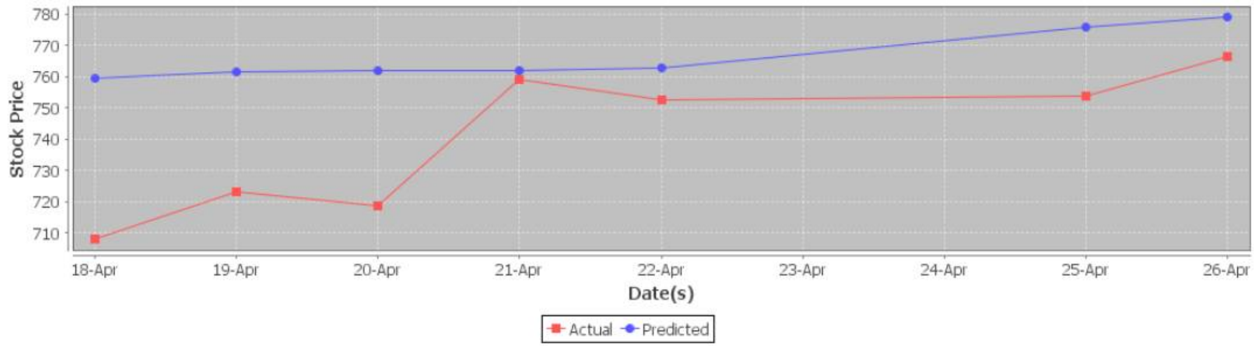

Fig.6. Linear Regression Prediction Graph

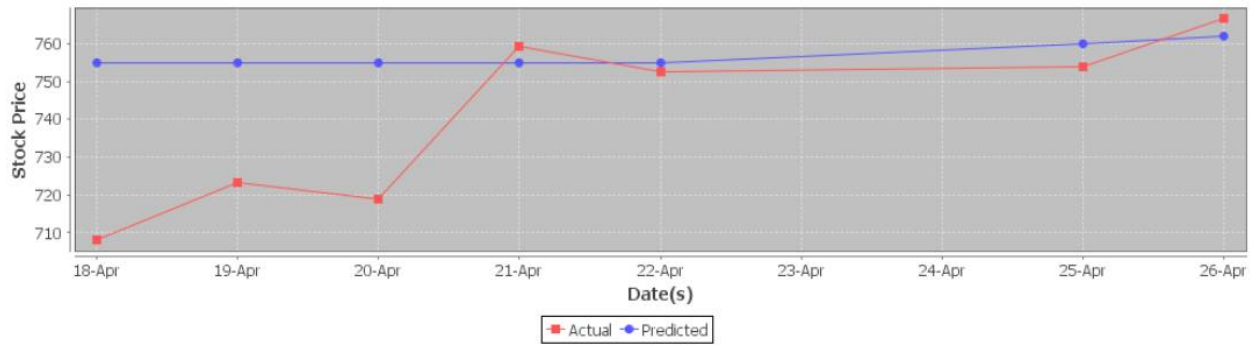

Fig.7. Moving Averages Prediction Graph

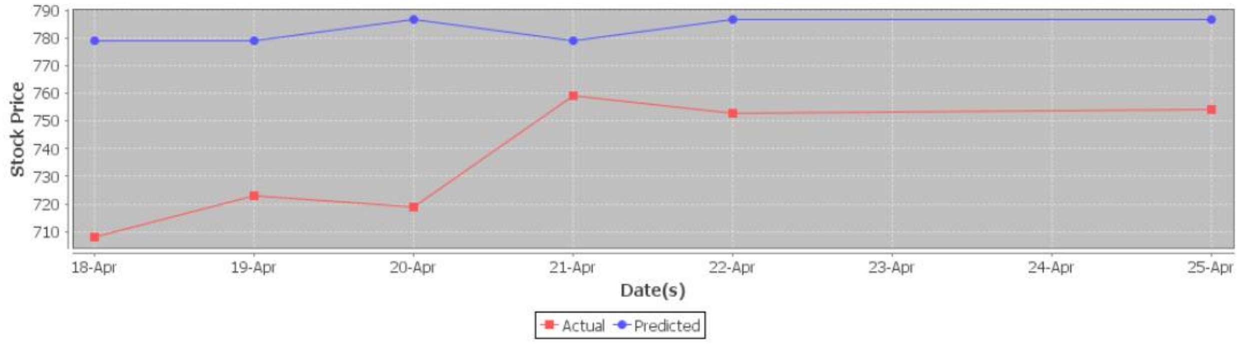

Fig.8. Forecasting Prediction Graph

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

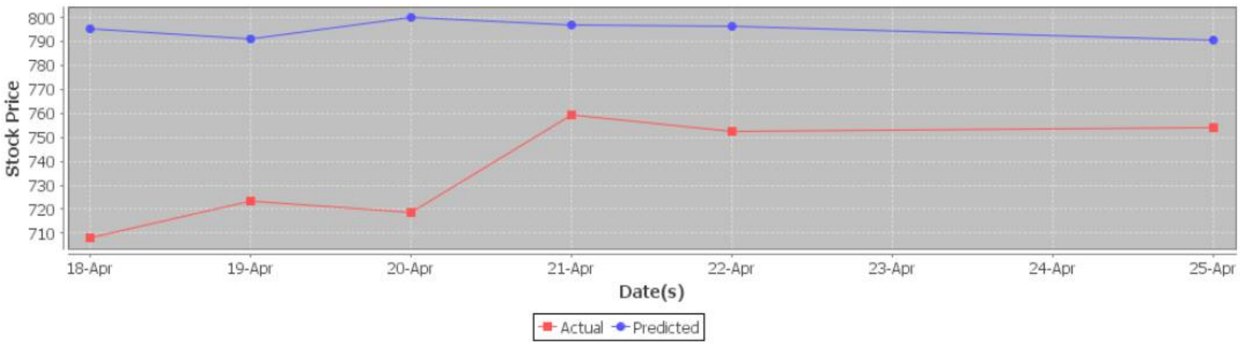

Fig.9. Multilayer Perceptron Prediction Graph

VII. CONCLUSION

The main advantage of our Hybrid Stock Prediction Model system lies in the fact that it integrates the conventional approach of Statistical Computational Methodologies with the ANN approach in implementing a primitive Stock Market Prediction Model. Customizable parameters facilitate higher prediction accuracy. Error-free prediction is achieved as the outcome is filtered on the basis of the efficiency of the underlying algorithms' performance. On the other hand, accuracy of the system fluctuates as per the performance of the algorithms. Computational horsepower can be problematic since the system has a high complexity of implementation due to the underlying algorithms working in parallel. Hence, optimized coding of the modules is essential.

ACKNOWLEDGMENT

We would like to express our heartfelt gratitude to our Project Guide, Prof. P. M. Daflapurkar, and our Project Sponsor, Mr. S. P. Chevuru (Luxorsys Technologies Pvt. Ltd.). Their constant support, domain expertise, technical insights and logistics support, throughout the project tenure, aided us greatly in completing this work successfully.

REFERENCES

1. D. Banerjee, "Forecasting of Indian Stock Market using Time-Series ARIMA Model", In: Proc. of the 2nd Int. Conference on Business and Information Management (ICBIM), IEEE, 2014, pp.131–135.
2. O. Abdalla, A. Elfaki, and Y. Murtadha, "Optimizing the Multilayer Feed-Forward Artificial Neural Networks Architecture and Training Parameters using Genetic Algorithm", Int. Journal of Computer Applications (IJCA), Vol.96, Issue-10, 2014, pp.42–48.
3. B. Chauhan, U. Bidave, A. Gangathade, and S. Kale, "Stock Market Prediction Using Artificial Neural Networks", Int. Journal of Computer Science and Information Technologies (IJCSIT), Vol.5, Issue-1, 2014, pp.204–207.
4. S. Jabin, "Stock Market Prediction Using Feed-Forward Artificial Neural Network", Int. Journal of Computer Applications (IJCA), Vol.99, Issue-9, 2014, pp.4-8.
5. M. B. Patel, and S. R. Yalamalle, "Stock Price Prediction Using Artificial Neural Network", Int. Journal of Innovative Research in Science, Engineering and Technology (IJIRSET), Vol.3, Issue-6, 2014, pp.13755-13762.
6. D. Kumar, and S. Murugan, "Performance Analysis of Indian Stock Market Index using Neural Network Time-Series Model", In: Proc. of the Int. Conference on Pattern Recognition, Informatics and Mobile Engineering (PRIME), IEEE, 2013, pp.72–78.
7. A. V. Devadoss, and T. A. A. Ligori, "Stock Prediction Using Artificial Neural Networks", Int. Journal of Data Mining Techniques and Applications (IJDMTA), Vol.2, 2013, pp.283-291.
8. K. Abhishek, A. Khairwa, T. Pratap, and S. Prakash, "A Stock Market Prediction Model Using Artificial Neural Network", In: Proc. of the 3rd Int. Conference on Computing Communications and Networking Technologies (ICCCNT), 2012.
9. S. Simon, and A. Raoot, "Accuracy Driven Artificial Neural Networks In Stock Market Prediction", Int. Journal on Soft Computing (IJSC), Vol.3, Issue-2, 2012, pp.35-44.
10. Z. Khan, T. Alin, and Md. A. Hussain, "Price Prediction of Share Market using Artificial Neural Network", Vol.22, Issue-2, Int. Journal of Computer Applications (IJCA), Vol.22, Issue-2, 2011, pp.42–47.
11. M. P. Naeini, H. Taremian, and H. B. Hashemi, "Stock Market Value Prediction Using Neural Networks", In: Proc. of the Int. Conference on Computer Information Systems and Industrial Management Applications (CISIM), IEEE, 2010, pp.132-136.
12. R. Arora, "Artificial Neural Networks for Forecasting Stock Price: Statistical Arbitrage using Artificial Neural Networks", IIT Kanpur, 2008, pp.1–15.
13. B. Qian, and K. Rasheed, "Stock Market Prediction with Multiple Classifiers", Springer Science + Business Media, LLC, 2006, pp.25–33.
14. M. Thenmozhi, "Forecasting Stock Index Returns Using Neural Networks", Delhi Business Review, Vol.7, Issue-2, 2006, pp.59-69.
15. Y. Shachmurove, and D. Witkowska, "Utilizing Artificial Neural Network Model to Predict Stock Markets", CARESS Working Paper #00-11, 2000, pp.1-25.