International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

Summarize the Work Related to Prediction of Stock Market's Option Prices with Artificial Intelligence using Standard Dataset

Ashish Pathak, Parmalik Kumar

Abstract*: Forecasting and prediction are based on pattern recognition. It may be a human energy potential increase day today when he grownup a young guy, but afterward, his energy potential going downwards. So, we observed the pattern with the help of neural network models; these are radical bias function (RBP) and back-propagation (BP). Utilizing the neural network model, it also has many classification parts like a deep neural network, feedforward neural network, recurrent neural network, convolutional neural network and many more. In the forecasting or prediction, we have a large amount of data to manage. We trained the data with algorithm and here we also use the neural network models. We used optimization techniques that are inspired by biological swarm. Nowadays, lots of data generate day by day like market, medical, education, automobile, etc. we need recognition of the pattern for prediction of future expectations. That expectation of prediction very helpful and needy to gain profit of human beings. In this work, we use SOM (self-Organized Map), RBF (Radical Bias Function), DNN (Deep Neural Network) and PGO (Plant Grow Optimization).

The total data point for the processing used 27500. The evaluation of the performance used standard parameters such as ET, MAE, MSE, RMSE and MI. The proposed algorithm implemented in MATLAB software. The cascaded neural network classifier is the combination of the SOM and RBF neural network models. The SOM neural network model proceeds the task of clustering and RBF neural network model used for prediction.

***Keywords*: Prediction, Stock market, option price, NSE CNX dataset, classifier, neural network, radial bias function, plant grow optimization.**

I. INTRODUCTION

Financial exchange costs and its patterns structure the most unpredictable division of finance. It tends to be the main purpose of some significant examines that target catching its unpredictability and anticipating its best courses of actin. The historical information created by stock exchanges is gigantic. Since the measure of information to be examined is so immense, it is extremely hard for any person to consider every one of the information esteems while anticipating future stock trends[1]. The basic and specialized investigations are two principle systems pervasive in money that is utilized for financial exchange expectations. The principal investigation considers reasons and factors that influence an organization's or economy's standard business and its future possibilities. On another hand, specialized investigation endeavors to think about varieties in costs of stock and uses this verifiable information to foresee its future costs [2]. In this way, it considers past patterns in the stock cost to predicted stock value changes later on. There is the prominence of virtual financial exchange games as of late, where news and occasions are made utilizing on the web stages and scored physically [3, 4]. These occasions, thusly highly affect the estimation of future stock costs. The creators have obtained the examination thought from this idea of virtual shares games. The target of this work is to discover the link between the authentic stock information of any organization and news/occasions identified with a similar organization and to predicted a model that can observe the stock costs precisely thinking about the effect of news and occasions on the stock cost. The writers have planned and actualized a conclusion model that chooses the extremity of the latest stories. At that point, the news vector generation technique is created to change over content articles into numeric vectors. At long last, a regression model is based on understanding the effect of news/events on changes in financial exchange costs [5-7].

In the remaining part of this paper, section-II defined the research study, in the next section-III, described the problem domain & motivation, in the section IV methodology introduce the proposed model, algorithm and used performance parameters. In the section V simulation and result analysis, we discussed the dataset of NSE, simulation process, results from analysis and performance analysis and finally concluded the conclusion & future direction of the work.

II. RESEARCH STUDY

Manuscript published on November 30, 2019.
* Correspondence Author
Mr. Ashish Pathak*, Computer Science & Engineering, Bhopal, India. Email: ashishpathakinfo@gmail.com
Prof. Parmalik Kumar, Computer Science & Engineering, Bhopal, India. Email: parmalik83@gmail.com

Ref. no.	Author	Index term	Technique	Type	Publication
[1]	Ülke V, Sahin A, Subasi A.	• Inflation forecasting • Time series models • Machine learning models	• KNN • Naïve • MLP • SVR	Comparative Analysis	Neural Computing and Applications. 2018
[2]	Čeperić E, Žiković S, Čeperić V.	• Natural gas • Henry hub • AI • Feature selection techniques • Regression-based SVM • NN	• SVR • NN • Naïve • AR • ARIMA	Review	Energy. 2017
[3]	Barboza F, Kimura H, Altman E.	• Bankruptcy pattern find • ML • SVM • Boosting • Bagging • Random forest	• SVMLin • • SVM-RBF • Boosting • Bagging • Random Forest • NeuralNet • Logit • LDA	Result Analysis	Expert Systems with Applications. 2017
[4]	Bzdok D, Meyer-Lindenberg A.	• AI • Endophenotypes • ML • Null-hypothesis testing • Personalized medicine • RDoC • Single-object observation	• RDoC • Machine Learning	Review	Biological Psychiatry: Cognitive Neuroscience and Neuroimaging. 2018
[5]	Dimitroff G, Röder D, Fries CP.	• The Heston models • Risk Neutral Model Calibration using Neural Networks • The neural network architecture	• Neural Networks	Result Analysis	SSRN 3252432. 2018

[6]	Hajizadeh E, Mahootchi M.	• RBF • NN • Approximate versatile coding • PSO	• Binomial Tree • ADP method • RBFNN • RBFNN-PSO	Result Analysis	Neural Computing and Applications. 2018
[7]	Yi H, Jung H, Bae S.	• Traffic prediction • Deep learning neural networks • Transportation big data	• DNN	Result Analysis	IEEE. 2017
[8]	Hu H, Tang L, Zhang S, Wang H.	• BP-NN • Sine technique • Stock cost	• GA feature discretization Classification model • SVM • BPNN • GA-ANN hybrid model • ISCA-BPNN hybrid model	Result Analysis	Neurocomputing. 2018
[9]	Karaoglan AD, Karademir O.	• Production System • Data and techniques	• FF-BPN • ANN • SVM • GA • LVQ	Result Analysis	The Engineering Economist. 2017
[10]	Kim HY, Won CH.	• LSTM • Deep learning • GARCH • Volatility prediction • Hybrid model	• DFN • LSTM • W-DFN • G-DFN • E-DFN • W-LSTM	Comparative Analysis	Expert Systems with Applications. 2018
[11]	Lago J, De Ridder F, De Schutter B.	• Electricity price forecasting • Deep learning • Benchmark study	• DNN • GRU • LSTM • MLP • SVR. • SOM-SVR • SVR-ARIMA • CNN • RBF	Result Analysis	Applied Energy. 2018

[12]	Long W, Lu Z, Cui L.	• Stock Price Prediction • Feature Engineering • Deep Learning	• MFNN • RNN • CNN • LSTM • SVM • LR • RF • Line-based regression	Result Analysis	Knowledge-Based Systems. 2019
[13]	McNally S, Roche J, Caton S.	• Bitcoin • AI • GPU • RNN • LSTM	• LSTM • RNN • ARIMA	Result Analysis	IEEE. 2018
[14]	Pan WT, Han SZ, Yang HL, Chen XY.	• BP-NN • GABPN • Decision tree • Data envelopment analysis • Mutual fund	• DEA • BPN • GABPN	Result Analysis	Cluster Computing. 201
[15]	Pradeepkumar D, Ravi V.	• FTSV Prediction • GARCH • NN • Particle swarm optimization	• GARCH • MLP • GRNN • GMDH • RF • QRRF • PSOQRNN • QRNN	Result Analysis	Applied Soft Computing. 2017
[16]	Ramakrishnan S, Butt S, Chohan MA, Ahmad H.	• Exchange Rate • Commodity Price • Machine Learning • Neural Network • SVM • RF	• NN • SVM • RF	Result Analysis	IEEE. 2017
[17]	Yang Z	• Forecasting • ARMA • KELM • Wavelet transform • SAPSO	• KELM • ARMA • W-ARMA-BP • W-ARMA-LSSVM		Applied Energy. 2017

III. PROBLEM DOMAIN & MOTIVATION

In today's era, the stock market has lots of data with several attributes like option price, high price, low price, open price and close price, etc. here, we focus on option price pattern day today. The option price prediction finds some research gaps in the working of machine learning-based techniques. When we use machine learning, selected the variable input like the learning rate, hidden layer, nodes in every layer to get effective results[17, 18, 19]. The data size of the stock market minimizes. For the utilization forecasting or prediction is chosen a better attribute. How attribute extracted in this process similarly mapping of stock market option price data. Need to improve the learning rate with the help of trained the option price data[19, 20].

With the help of prediction or forecasting, we get the future of stock market prices. It widely impacts the trading of the share market; these markets belong to banking, finance, statics and computer science, etc. The motivation for which is normally to predict the method for future option costs of the stock exchange with the end goal that stocks can be purchased and sold at beneficial positions.

Proficient brokers normally use crucial and additionally specialized examination to analyze stocks and settle on investment choices. The basic examination is the recent decade methodology, including an investigation of organization essentials, for example, incomes and costs, advertise position, yearly development rates, and so on. The specialized investigation, then again, is exclusively founded on the investigation of recorded value variations.

IV. METHODOLOGY

A. Proposed Model

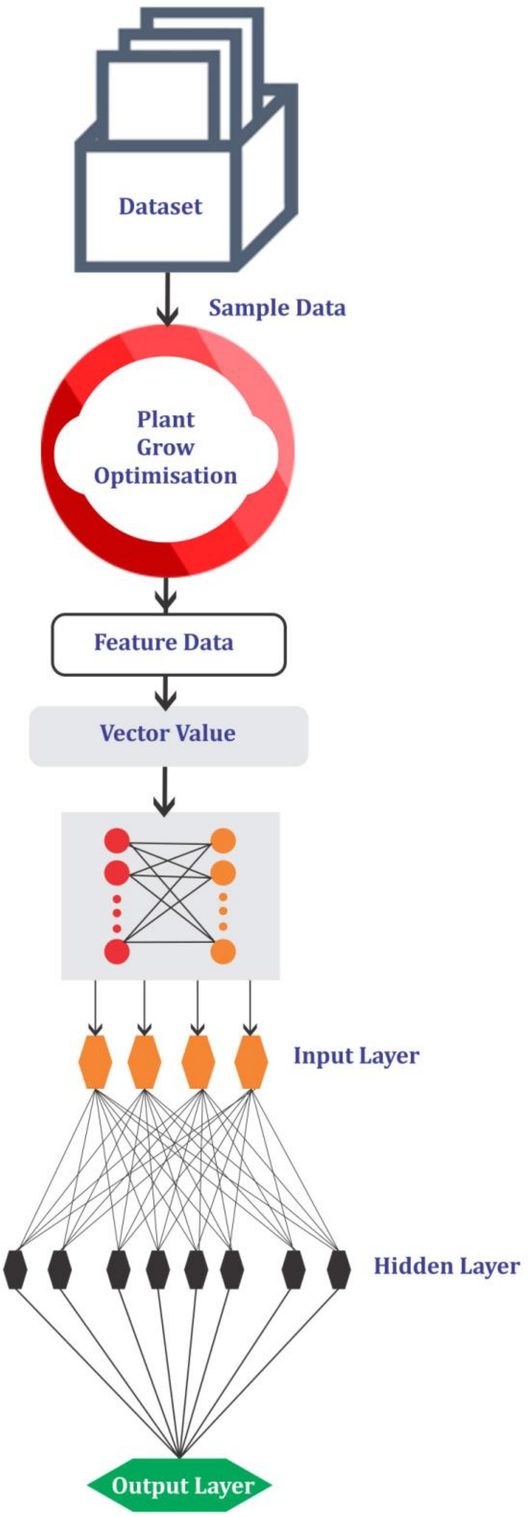

Fig. 1.Process block diagram of price prediction with attribute optimization.

B. Proposed Algorithm

The output of optimal data of the PGO map as $X = [x_1, x_2, \ldots, x_{N_l}, x_{N_l+1}, \ldots, x_N] \in R^{D\times N}$ that is corresponding to a cascaded, where each$x_i \in R^D$ represents the ith index of vector, D is the number of data samples and$N = (N_l + N_u)$ is the number of training samples. The first N_l samples $X_l = [X_1^l, X_2^l, \ldots .. X_{N_l}^l]$ are labeled with the corresponding labels $Y_l = \{y_i\}_{i=1}^{N_l}$ the remaining N_u samples $X_u = [X_{N_l+1}^u, \ldots \ldots \ldots, X_N^u]$.

In the process of cascaded classifier, the value of matrix W map SOM output in RBF input vector and minimized the attribute variance value as

$$W_p = arg_w \min\left(\sum_{s=1}^{N}\sum_{t=1}^{k_s}\left\|W_{X_a} - W_{X_b}\right\|^2 rf\right) \quad \ldots\ldots\ldots\ldots\ldots(1)$$

where X_b is the neighbor of X_a and k_s is the number of neighbors of X_s. rf is the relative feature difference values of sample X_a and Xb , mapping of each sample $X_i^l (i = 1, 2, \ldots, N_l)$ can be as vector $a_i \in R^{N_l\times 1}$ under a samples of data $X_l = \left[X_1^l, X_2^l, \ldots .. X_{N_l}^l\right] \in R^{N_l\times 1}$ that is composed as

$$\min_{a_i} \|a_i\|_1, s.t. \frac{1}{2} \|X_i^l - X_l a_i\|^2 < MAR \ldots\ldots\ldots\ldots..(2)$$

where MAR is the mean absolute error

define the value of W_{dif}^i as the distance between the winner X_i^l and its successor data

$$w_{dif}^i = \sum_{k=1}^{k_s2} \|X_i^l - X_{ik}^l\|^2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

where X_{ik}^l is the successor data of points X_i^l and k_{s2} is the number of selected winners.

The learning rate in should be time varying. This requirement can be satisfied by choosing an exponential decay for $\eta(t)$.

$$\eta(t) = \eta o exp\left(-\frac{t}{w1}\right), t = 0,1,2 \ldots\ldots\ldots, \ldots\ldots\ldots\ldots. (4)$$

The update matrix value process after learning

$$updateWij = \sum_{k=1}^{k_s2}\left\|WX_i^l - WX_{ik}^l\right\|^2 - \sum_{j=1}^{k_s1}\left\|WX_i^l - WX_{ij}^l\right\|^2 \quad \ldots\ldots\ldots\ldots(5)$$

The updated winner's matrix data passes through the RBF interconnected input layers

$$Rfb = \sum_{i=1}^{N_l} updateWij,$$

$$= \sum_{i=1}^{N_l}\sum_{k=1}^{k_s2}\left\|WX_i^l - WX_{ik}^l\right\|^2 - \sum_{i=1}^{N_l}\sum_{j=1}^{k_s1}\left\|WX_i^l - WX_{ij}^l\right\|^2 \quad \ldots\ldots\ldots\ldots\ldots..(6)$$

Calculating the deviation value of sample data points to measure the error rate

$$MSE = \frac{1}{n}\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m}(Rfbij - yij)^2} \ldots\ldots\ldots\ldots\ldots\ldots\ldots..(7)$$

C. Performance Parameter

Normalized Mean Squared Error (NMSE)[21]

$$NSME = \frac{1}{N}\frac{\sum_{n=1}^{N}(r_{t+1}^{n} - r^{\wedge}{}_{t+1}^{n})^2}{var(r_{t+1}^{n})}$$

Root Mean Squared Error (RMSE)[22]

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(r_{t+1}^{n} - r^{\wedge}{}_{t+1}^{n})^2}$$

Mean Absolute Error (MAE)[22]

$$MAE = \frac{1}{N}\sum_{n=1}^{N}|r_{t+1}^{n} - r^{\wedge}{}_{t+1}^{n}|$$

Mutual Information (MI)[22]

$$MI(r_{t+1}; u_t) = \sum_{r_{t+1}; u_t} p(r_{t+1}, u_t) \log\frac{p(r_{t+1}, u_t)}{p(r_{t+1})P(u_t)} \approx \frac{1}{N}\sum_{n=1}^{N}\log\frac{p(r_{t+1}^{n} \mid u_t^{n})}{p(r_{t+1}^{n})}$$

V. SIMULATION & RESULT ANALYSIS

A. Dataset

A stock's market prediction performance can, to some extent, be evaluated based on financial indicators presented in the company's yearly report[17]. The implementation analysis used the dataset of eight-thousands of data of different selection points recent five year. Source of dataset given below:

https://www.quandl.com/data/NSE/CNX_NIFTY-National-Stock-Exchange-of-India-NSE-Stock-Index-CNX-Nifty

B. Simulation Process

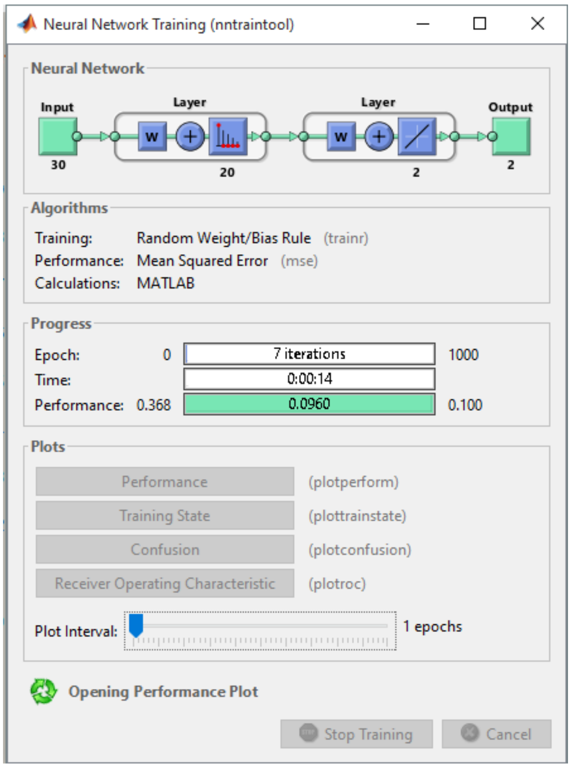

Fig. 2.Neural network training wizard with complete process performance in stock market prediction.

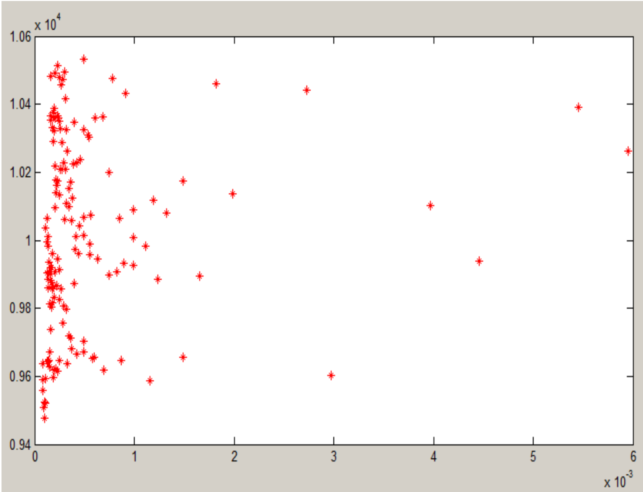

Fig. 3.Output view with complete process performance in stock market prediction.

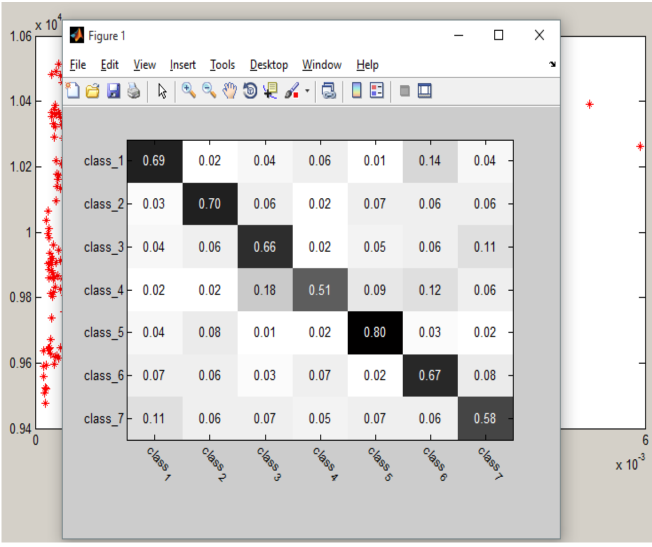

Fig. 4.Output view with complete process performance in stock market prediction.

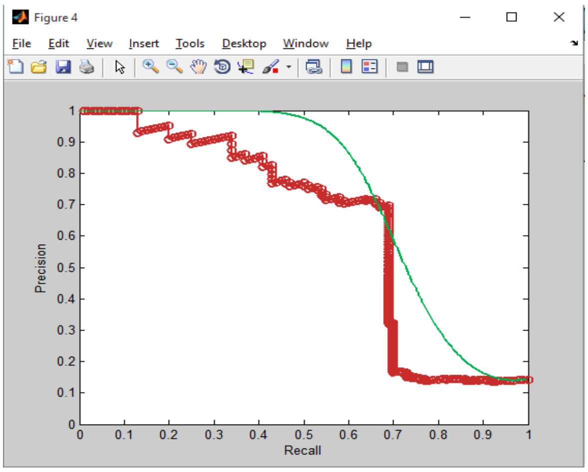

Fig. 5.Output view between recall and precision within stock market prediction.

Retrieval Number: D8192118419/2019©BEIESP
DOI:10.35940/ijrte.D8192.118419
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

C. Result Analysis

Table- I: Comparative result of elapsed time using DNNAR and CSR technique.

Input Data Points	DNNAR Method	CSR Method
2	103.660632	99.835782
4	98.950054	98.024044
6	97.863901	97.264640
8	98.376154	97.695583
10	97.815638	88.568070

Table- II: Comparative result of root mean squared error for DNNAR and CSR technique.

Input Data Points	DNNAR Method	CSR Method
2	0.932100	0.433210
4	0.766000	0.416600
6	0.710700	0.411070
8	0.683000	0.408300
10	0.666400	0.406640

Table- III: Comparative result of normalized mean squared error for DNNAR and CSR technique.

Input Data Points	DNNAR Method	CSR Method
2	0.388320	0.651260
4	0.438710	0.643800
6	0.443800	0.643700
8	0.462950	0.634700
10	0.446800	0.651260

Table- VI: Comparative result of mean absolute error for DNNAR and CSR technique.

Input Data Points	DNNAR Method	CSR Method
2	0.772100	0.363210
4	0.606000	0.346600
6	0.550700	0.341070
8	0.523000	0.338300
10	0.506400	0.336640

Table- V: Comparative result of mutual information for DNNAR and CSR technique.

Input Data Points	DNNAR Method	CSR Method
2	0.932100	0.913210
4	0.766000	0.916600
6	0.710700	0.911070
8	0.683000	0.908300
10	0.666400	0.906640

D. Performance Analysis

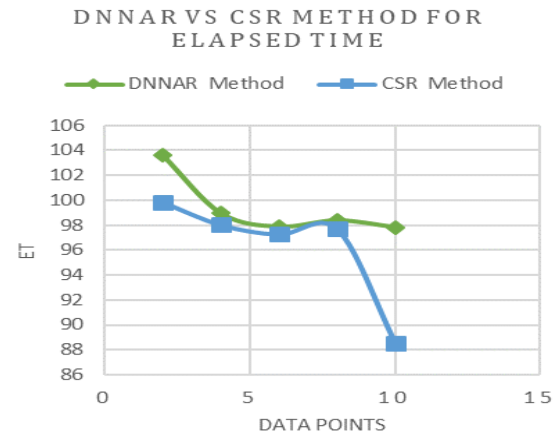

Fig. 6.Comparative performance between DNNAR and CSR method for elapsed time (ET).

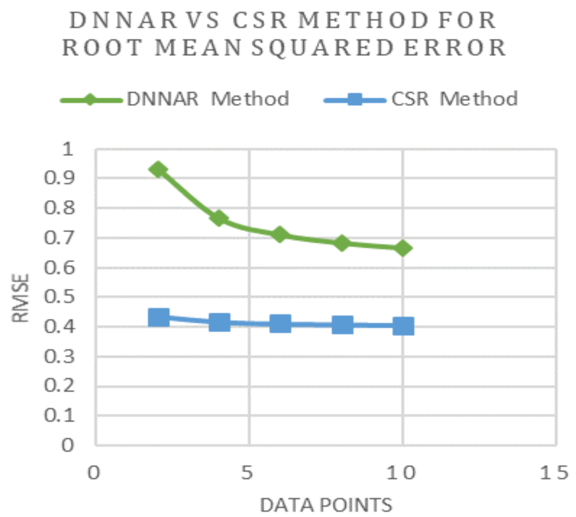

Fig. 7.Comparative performance between DNNAR and CSR method for mean squared error.

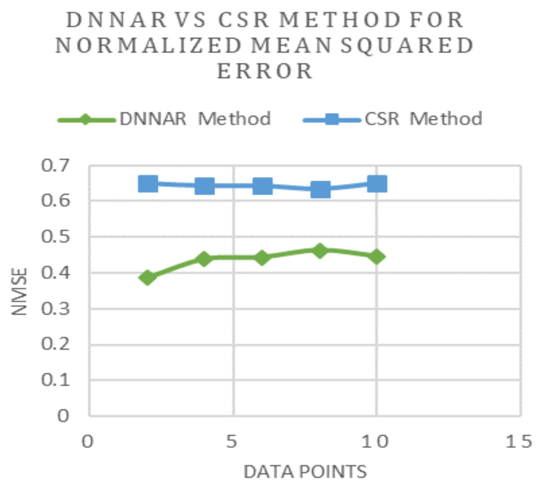

Fig. 8.Comparative performance between DNNAR and CSR method for normalized mean squared error.

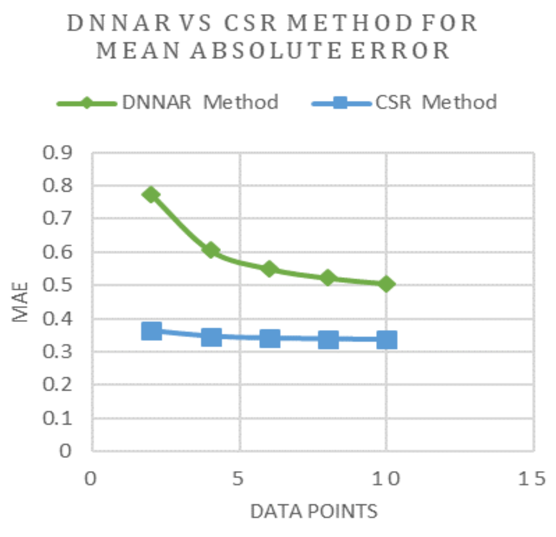

Fig. 9.Comparative performance between DNNAR and CSR method for mean absolute error.

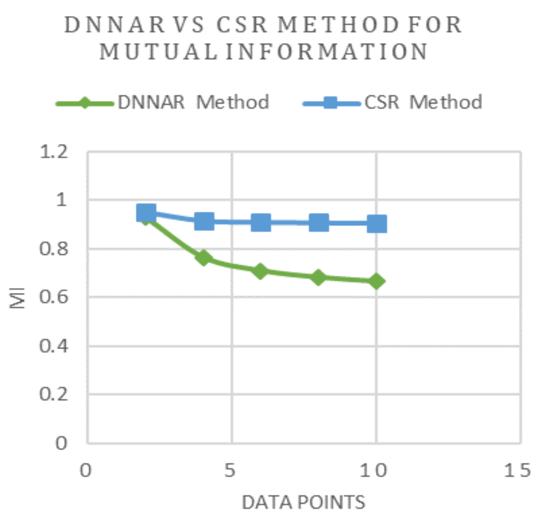

Fig. 10. Comparative performance between DNNAR and CSR method for mutual information.

VI. CONCLUSION & FUTURE DIRECTION

In this paper, we studied the many research papers and got the basic knowledge and all the initial term that help us to understand the research papers. During the study of these research papers, the most important thing is prediction another word we can say to know about the future, but that is mostly described based on the data patterns. We know about the datasets like google-dataset, yahoo dataset, BSE dataset, NSE, CNX and many more to get the pattern and also used to classify or simplify to minimize the data of dataset using neural network models and machine learning algorithms. With the help of machine learning, we trained the data of the dataset. The optimization algorithm reduces the variation of attribute selection for the process of classification. Artificial intelligence is an essential factor in the option price prediction in the stock market. On behalf of the whole study, we will design a predicted model with machine learning that is useful for the stock market and will get the benefit of buying or sell the stock's shares. They evaluated the performance of option price predication using two algorithms; these are optimization algorithms and cascaded neural network-based classification. With the help of the plant grow optimization (PGO) algorithm, the variation of attribute selection reduced. It has 3 steps: 1st one is dynamic to define the population in the form of data loading for the process of optimization. 2nd one is used branch selection process for the new population and the final selection of the optimal value of price for the input processing of cascaded classifier (two neural network models: RBF and SOM). SOM used for clustering and RBF used for trained patterns of data and predicts the value of option price. We evaluate the data on the five performance parameters ET, NMSE, RMSE, MAE and MI. Got a better performance to compare to the deep neural network-based classifier (DNNAR) by modified plant grow optimization algorithm (CSR).

REFERENCES

1. Ülke V. A comparison of time series and machine learning models for inflation forecasting: empirical evidence from the USA. NCA. 2018 Sep 1;30(5):1519-27.
2. Čeperić E. Short-term forecasting of natural gas prices using machine learning and feature selection algorithms. Energy. 2017 Dec 1; 140:893-900.
3. Barboza F. Machine learning models and bankruptcy prediction. ESsAs. 2017 Oct 15; 83:405-17.
4. Bzdok D, Meyer-Lindenberg A. Machine learning for precision psychiatry: opportunities and challenges. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging. 2018 Mar 1;3(3):223-30.
5. Dimitroff G. Volatility model calibration with CNN. Available at SSRN 3252432. 2018 Sep 20.
6. Hajizadeh E. Optimized radial basis function neural network for improving approximate dynamic programming in pricing high dimensional options. NCA. 2018 Sep 1;30(6):1783-94.
7. Yi H. DNN for traffic flow prediction. IEEE. 2017 Feb 13 (pp. 328-331).
8. Hu H. Predicting the direction of stock markets using optimized neural networks with Google Trends. Neurocomputing. 2018 Apr 12; 285:188-95.
9. Karaoglan AD. Flow time and product cost estimation by using an ANN: A case study for transformer orders. The Engg. Economist. 2017 Jul 3;62(3):272-92.
10. Kim HY. Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. ESsAs. 2018 Aug 1; 103:25-37.
11. Lago J. Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. Applied Energy. 2018 Jul 1; 221:386-405.
12. Long W. Deep learning-based feature engineering for stock price movement prediction. KBSs. 2019 Jan 15; 164:163-73.
13. McNally S. Predicting the price of Bitcoin using Machine Learning. IEEE 2018 Mar 21 (pp. 339-343).
14. Pan WT. Prediction of mutual fund net value based on data mining model. CC. 2018:1-6.
15. Pradeepkumar D. Forecasting financial time series volatility using PSO trained quantile RNN. ASC. 2017 Sep 1; 58:35-52.
16. Ramakrishnan S. Forecasting Malaysian exchange rate using machine learning techniques based on commodities prices. IEEE 2017 Jul 16 (pp. 1-5).
17. Yang Z. Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods. Applied Energy. 2017 Mar 15; 190:291-305.
18. Kaushal, Aditi, and Prerit Chaudhary. "News and events aware stock price forecasting technique." In 2017 International Conference on Big Data, IoT and Data Science (BID), pp. 8-13. IEEE, 2017.

19. Navale, G. S., Nishant Dudhwala, Kunal Jadhav, Pawan Gabda, and Brij Kishor Vihangam. "Prediction of stock market using data mining and artificial intelligence." International Journal of Computer Applications 134, no. 12 (2016): 9-11.
20. Dingli, Alexiei, and Karl Sant Fournier. "Financial Time Series Forecasting–A Deep Learning Approach." Int. J. Mach. Learn. Comput 7, no. 5 (2017): 118-122.
21. Vaidya, Ajinkya M., Nikunjkumar H. Waghela, and Sneha S. Yewale. "Decision support system for the stock market using data analytics and artificial intelligence." International Journal of Computer Applications 117, no. 8 (2015).
22. Milosevic, Nikola. "Equity forecast: Predicting long term stock price movement using machine learning." arXiv preprint arXiv:1603.00751 (2016).

AUTHORS PROFILE

Ashish Pathak, Master of Technology-2018 in Computer Science and Engineering from RGPV university, Bhopal, MP, India. Bachelor of Engineering-2010 in information technology from RGPV university, Bhopal, MP, India.