IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-5, Issue-2 E-ISSN: 2347-2693

Artificial Neural Network Based Trend Analysis and Forecasting Model for Course Selection

Chowdhury D.R.[1*] and Sen D.[2]

[1*]Department of Computer Science and Application, University of North Bengal, Darjeeling, India
[2] Department of Computer Science and Application, University of North Bengal, Darjeeling, India

[*]Corresponding Author: diliproychowdhury@gmail.com,

Available online at: www.ijcseonline.org

Received: 13/Jan/2017 Revised: 22/Jan/2017 Accepted: 12/Feb/2017 Published: 28/Feb/2017

***Abstract*—Selection of the proper higher educational courses is absolutely necessary for the prospective students. Selecting appropriate courses are really cumbersome job for the students who are having less information about present trend of education relating to get placements or jobs and for better development in future. In this paper, trend analysis and forecasting has proposed to predict the prospects of the selected higher educational courses in the field of computer science/technology. An online survey has done to get the dataset for analysis and there were altogether 151 data selected for the study. A Feed Forward Artificial Neural Network model has proposed and the best network architecture has been selected among the top five NN considering the parameters like fitness value, AIC (Akaike's Information Criterion) value, training, validation, test error values. The best network architecture is further analyzed using Levenberg-Marquardt (LM) and Conjugate Gradient Descent (CGD) algorithms for finding the accuracy of the trend. The study focuses on important input parameters during training of network architecture. Correct Classification Rate (CCR) for training and validation has been prepared to find the best network after a number of iterations. A comparative study between the LM and CGD algorithm has primed with a focus on confusion matrix. This study recommends and predicts the future trends of the selected higher educational computer science/technology courses by using ANN.**

Keywords-Artificial Neural Network (ANN); Conjugate Gradient Descent (CGD); Confusion Matrix; Feed-Forward Artificial Neural Network (FFANN); Levenberg- Marquardt (LM); Multi- Layer Preceptron (MLP); Trend Analysis.

I. INTRODUCTION

In today's world of technologies, Computer has infused into mostly all fields. A substantial swing has seen with the uses of computer technology by the general people of modern society which was mainly practiced by the experts earlier. Extensive improvements have seen in the design, architecture, usability and effectiveness of computing technology these days. As the increasing popularity and adoption of computer in different arenas is becoming more widespread and affordable, profound knowledge has become indispensable. Multi- disciplinary education fields are often attracted by the grown-up students. Choosing of any course is literally a cumbersome job for those who don't have much idea about the future opportunities in education. As a result, different under-graduate, post-graduate, and engineering courses related to the field of Computer Science/Technology have emerged. Different under-graduate course of three years' time span related to the field of computer are - B.Sc.(CS), BCA, B.Sc.(IT). Engineering Courses generally are of four years duration which includes B.Tech.(CSE) and B.Tech.(IT). Different post-graduate courses includes M.Sc.(IT), M.Sc.(CS), MCA, M.Tech(IT) and M.Tech(CSE). Each of which are two years durations except the MCA course, which is of three years duration.

This is basically a study for the students for selecting and preferring the higher education courses depend upon some parameters mentioned in later sections. Trend analysis has been proposed to forecast the prospects of the selected higher educational courses. So, for this particular study, following four courses have considered - M.Sc.(CS), MCA, B.Tech (CSE) and B.Tech(IT) to analyse and predict the future trend. The erstwhile methodologies for forecasting are usually established on historical data and different statistical models. Based on the comparison between the Neural Network and statistical models prepared by Faber and Lapedes, it has been concluded that Neural Network is more powerful forecasting method [1].

In the evolution of Intelligent Computing, ANN technique has found the most efficient and immensely used techniques for future prediction in different domains like Stock Market Analysis and others [2]. This technique can also be merged with different emerging fields like Green Computing [3]. Non-linear statistical modelling of data can be related with neural network. Neural network technique is being noticed working better in trend detection and pattern extraction from complicated or vague data [4].

II. MULTI- LAYER PERCEPTRON OF ANN

Due to their structural flexibility, good representational proficiencies and availability, MLPs are the simplest, fully connected and commonly used feed forward neural network architecture programs that are able to transform input data into a predictable solution with a massive number of programmable algorithms [5].

Figure 1 depicts the input layer with a vector of predictor variable values ($x_1 \dots x_p$), that ranges from -1 to 1, there is one hidden layer consisting of three neurons. The output from the hidden layer is fed to the output layer of the model. The output values are ($y_1 \dots y_m$) [6].

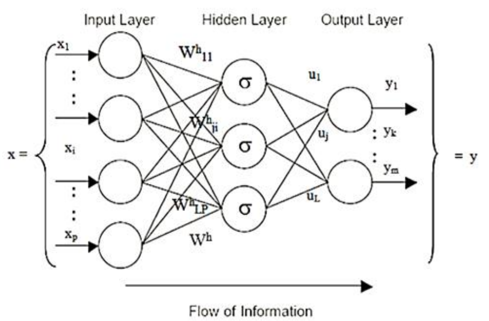

Fig 1: Multilayer Perceptron Network

III. NEURAL NETWORK TRAINING ALGORITHM

Different alternative learning methods and variants exist for neural networks. The classical back propagation algorithm, the first successful algorithm in the context of feed forward multilayer networks, is very useful for the learning process. It has two major shortcomings [7]:

- Convergence to local minima.
- Slow learning speed.

To resolve these obstacles several variants of the initial algorithm as well as new methods focusing on the problem of slow learning speed have been projected to accelerate it [8]. Several methods have proposed for the learning of various algorithms. Among them, second order methods are the fastest learning algorithms [9]. For incremental growth of the convergence speed, second derivatives have proposed in several works [10][11]. In case of learning speed, it has been recognized that these methods are the more proficient ones than those methods based only on the gradient descent technique.

This study basically focused on two most relevant algorithms of second order methods. They are:

A. Levenberg-Marquardt (LM) Algorithm

The Levenberg-Marquardt algorithm is an iterative procedure, one of the extensively used non-linear optimization algorithm and reputably fast algorithm. It is appropriate for exceptional results while functioning with small training set and also for regression networks as it is defined only for sum squared error function and will be ignored if other error function is selected for the network during LM training [12]. The LM algorithm can be seen as an interpolation between Gradient Descent and Gauss-Newton method. The following equation depicts the algorithm:

$$[J^T J + \lambda diag(J^T)]\delta = J^T[y - f(\beta)] \quad (1)$$

where, J is the Jacobian matrix whose ith row equals to J_i , λ is non-negative damping factor which is adjusted in each iteration, f and β are vectors with ith component $f(x_i, \beta)$ and y_i, respectively, diag is the diagonal matrix containing the diagonal elements of $J^T J$, given the increment δ to the estimated parameter vector β.

B. Conjugate Gradient Descent (CGD) Algorithm

The Conjugate Gradient Descent algorithm is batch based or batch update algorithm. It also works better for advanced method of training multilayer perceptron network model. It is based on the linear search usage in the line of an optimum network weights' change. Network which contains a huge number of weights (>few hundreds) and/or multiple output units, it is considered as a recommended practice [12][13]. The correction of weights is conducted once per iteration.

Comparing the performance of Conjugate Gradient Descent with that of Back Propagation, the result of the first one is significantly superior; also much easier to use, and can be used whereas Back Propagation is apt, but a Conjugate Gradient Descent period is significantly more time-consuming (usually 3-10 times lengthier) than a back propagation period [13]. Mostly, this method provides more precise forecasting results. Figure 2 depicts a comparative analysis on the above mentioned algorithms based on computational speed and memory usage.

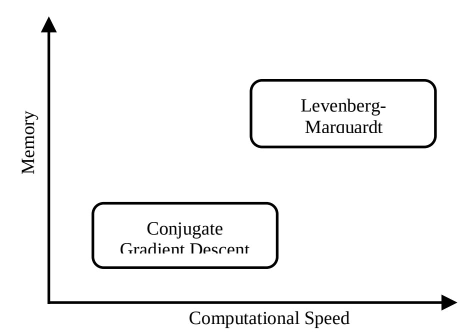

Fig 2: CGD method Vs. LM method

CGD algorithm is considerably slower whereas the requirement for memory is much lesser than LM algorithm.

IV. PROPOSED MODEL

A. Input Data

The data has been collected from an online survey by interviewing various people of different domains. Based on the survey, 151 data has been taken into observations for this study. The data has been standardized so as to be error free in nature. Table I depicts the input parameters of our study:

TABLE I
INPUT PARAMETERS

Sl. No.	Parameters	Column type
1	Equivalence MCA and B. Tech CSE	Categorical
2	Time span of MCA	Numerical
3	IT Industry Preferences	Categorical
4	Academic Field Preferences	Categorical
5	Expected Salary	Numerical
6	MCA obsolescence Time Span	Categorical
7	MCA obsolescence Limited Seats	Categorical
8	MCA obsolescence Less application oriented	Categorical
9	MCA obsolescence Less Valued than B. Tech	Categorical
10	MCA obsolescence Less demand in IT industry	Categorical
11	Course Preferences	Categorical

B. Data Analysis

The input dataset is taken for analysis through the process of careful training, validation and testing using Neuro-Intelligence Tool. This dataset is taken for training. Data analysis information is desired for correct data pre-processing. Here 11 columns and 151 rows are accepted for neural network training. Data partitioning has done at random manner. Table II depicts the partition result:

TABLE II
DATA PARTITION SET

Partition Set	Records	Percentage (%)
Total	151	100%
Training Set	103	68.21%
Validation Set	24	15.89%
Testing Set	24	15.89%
Ignore set	0	0

C. Pre-processing of Analysed Data

Data pre-processing is essential before designing the network architecture. Here the analysed data is pre-processed. Before pre-processing 11 columns are identified and after pre-processing 23 columns are identified. The input columns are ranged from -1 to 1 and the scaling range of output column(s) is from 0 to 1. The encoding parameters of numeric and categorical columns are depicted on Table III.

TABLE III
INPUT SCALING PARAMETERS

Code	Name of the Input Column	Encoding Parameters
C2	Equivalence MCA and BtechCSE	Two-state
C2	Time span of MCA	Two-state
C4	IT Industry Preferences	One-of-4
C4	Academic Field Preferences	One-of-4
C4	Expected Salary	One-of-4
C2	MCA obsolescence Time Span	Two-state
C2	MCA obsolescence Limited Seats	Two-state
C2	MCA obsolescence Less application oriented	Two-state
C2	MCA obsolescence Less Valued than BTech	Two-state
C2	MCA obsolescence Less demand in IT industry	Two-state
C4	Course Preferences	One-of-4

D. Selection of Hidden Layer

This study finds that one hidden layer is appropriate for design and development of the proposed network architecture, as two hidden layers are preferred for data modelling with discontinuities. It has been found that using two hidden layers improves the model rarely in this scenario, and it may leads to a greater risk of converging to a local minimum. Hence, a network model of three layers with one hidden layer has been designated.

E. Selection of Neurons in Hidden Layer

The numbers of neurons in the hidden layer are decided by the MLP network, which is a significant characteristic. The network properties for determining the number of neurons to be used in the hidden layers are as follows:

Input activation FX: Logistic
Output name: Course Preferences
Output error FX: Sum-of-squares
Output activation FX: Logistic
Classification model: Winner-takes-all.

Table IV depicts the top five network architectures to determine the appropriate NN model as well as the neurons in the hidden layer.

TABLE IV
TOP 5 NETWORK ARCHITECTURE

ID	Architecture	# of Weights	Fitness	Train Error	Validation Error	Test Error	AIC
1	[19-3-4]	76	3	0.89320	0.458333	0.66666	-555.7
3	[19-30-4]	724	2.6666	0.66990	0.541667	0.625	856.4
7	[19-16-4]	388	2.4	0.96115	0.583333	0.5833	-35.96
8	[19-22-4]	532	2.4	0.99029	0.583333	0.5833	109.2
5	[19-12-4]	292	2.1818	1	0.541667	0.6666	-421.0

F. Development of FFANN architecture

In this study, the Multi-layered Feed-Forward Artificial Neural Network architecture with 19 input nodes, 3 hidden nodes, and 4 output nodes have been implemented as the fitness of this architecture is comparatively better which is illustrated in Table IV and the best network graph is depicted in Figure 3.

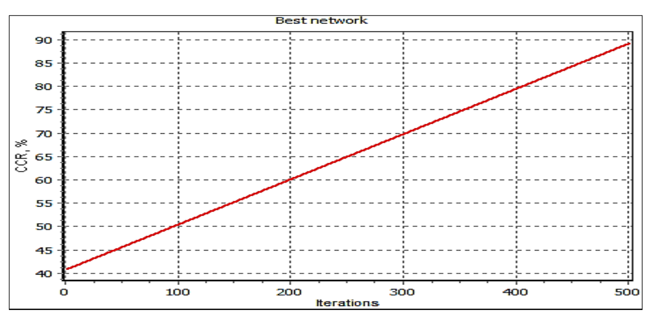

Fig. 3: Best Network Graph

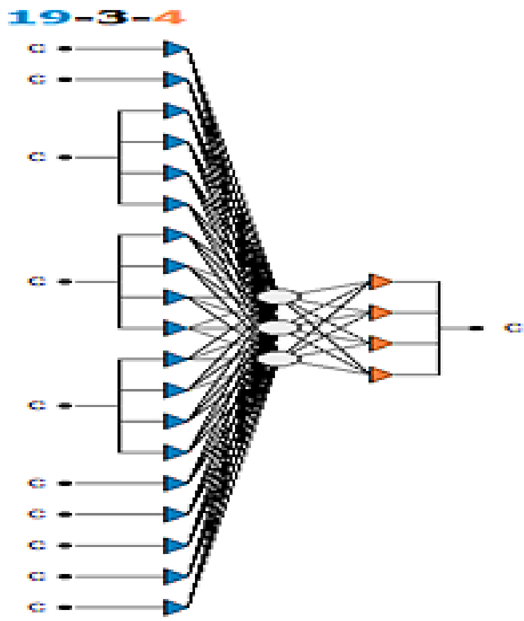

Fig 4: Multi-layered FFANN Architecture

The finalized data determines the numbers of input nodes; the numbers of hidden nodes are determined through trial and error method; and the numbers of output nodes are represented as a range. So [19-3-4] architecture is selected for training. Figure 4 shows the network architecture.

The weight of this network is 76. The train error, validation error and test error is generated which are 0.89320, 0.458333 and 0.66666, respectively. The AIC (Akaike's Information Criterion) value for the network is -555.7. Figure 3 demonstrates the CCR% (Correct Classification Rate) graph of the selected FFANN architecture after 501 iterations.

G. Training of FFANN using LM & CGD Algorithm

In this section our objective of the training process is to find the set of weight values which will produce the output from the neural network to match the actual target values as its closets. Neural network training and network weights adjustment of the input dataset are deliberate for training set. The validation set comprises of the portions of the data that are used to adjust network topology or network parameters other than weights and to choose the best network. Portion of the input data set is considered as test set, which are used to test how well the neural network performs on new data and also the errors that will occur during future network application.

Levenberg-Marquardt Algorithm and Conjugate Gradient Descent Algorithm both have been deployed for training the FFANN to show a comparative trend analysis. Figures 5 and 7 shows dataset errors, based on training set, validation set and the best network of both the training algorithms. The network errors have been shown in figure 6 and figure 8, respectively.

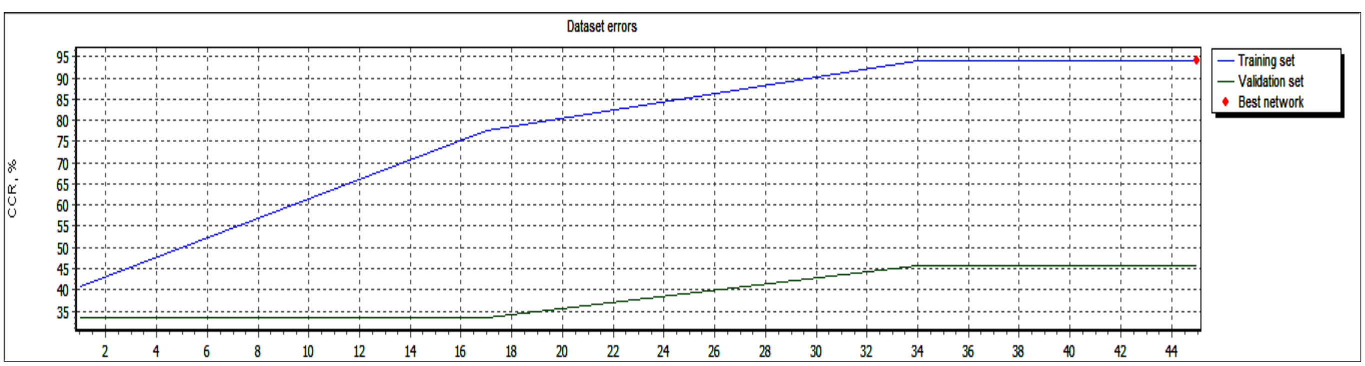

Fig. 5: Dataset Error in LM Method

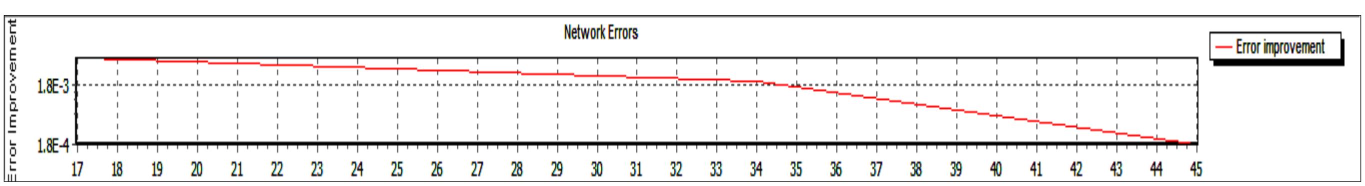

Fig. 6: Network Error in LM Method

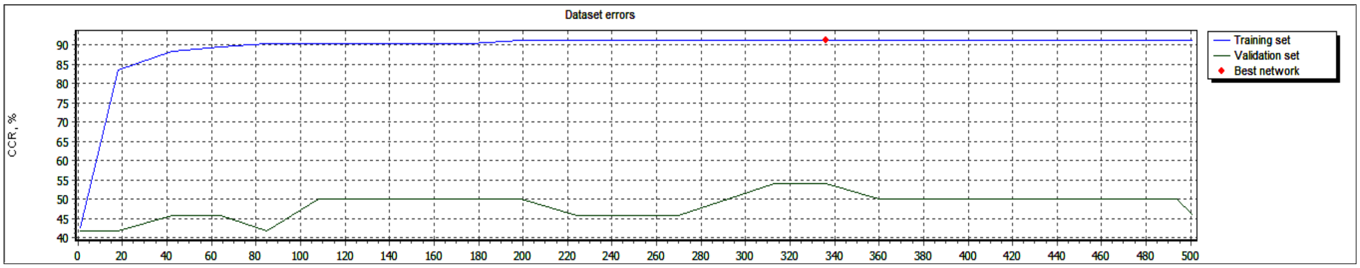

Fig. 7: Dataset Error in CGD Method

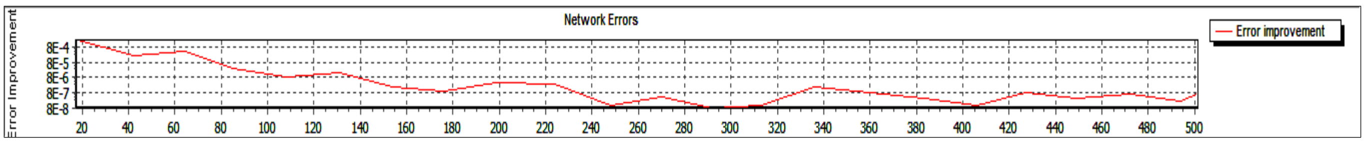

Fig. 8: Network Error in CGD Method

V. RESULT AND DISCUSSION

During training, importance that has been quantified by both the algorithms to the input parameters are depicted in Table V.

TABLE V
NETWORK STATISTICS

Parameters	LM Training Algorithm Importance (%)	CGD Training Algorithm Importance (%)
Equivalence MCA and B. Tech CSE	1.372376	3.932689
Time span of MCA	7.935179	6.411255
IT Industry Preferences	23.988147	31.240502
Academic Field Preferences	20.819233	21.771213
Expected Salary	14.186499	10.288175
MCA obsolescence Time Span	4.213254	5.148287
MCA obsolescence Limited Seats	5.545682	5.116603
MCA obsolescence Less application oriented	6.158504	3.183197
MCA obsolescence Less Valued than B. Tech	8.202358	5.828416
MCA obsolescence Less demand in IT industry	7.578768	7.079662

After training, the level of best network is selected through repeated iterations. Correct Classification Rate (CCR) for training and validation has computed to find the best network after a number of iterations.

Table VI describes the best LM network on iteration. Though 500 iterations have taken for training the network, the training was stopped after 45 iterations as it generated no error improvement reason.

TABLE VI
BEST LM NETWORK ON ITERATION

Iteration	CCR (training)	CCR (validation)
17	77.669907	33.333332
34	94.174759	45.833332
45	94.174759	45.833332

The study of identifying the best CGD network based on iteration is depicted in Table VII, where 500 iterations have taken and after completing of all iterations, the training stopped.

TABLE VII
BEST CGD NETWORK ON ITERATION

Iteration	CCR (training)	CCR (validation)
18	83.49515	41.66667
42	88.34952	45.83333
64	89.32039	45.83333
85	90.29126	41.66667
108	90.29126	50
131	90.29126	50
154	90.29126	50
176	90.29126	50
199	91.262138	50
224	91.262138	45.83333
248	91.262138	45.83333
270	91.262138	45.83333
292	91.262138	50
313	91.262138	54.16667
336	91.262138	54.16667
383	91.262138	50
406	91.262138	50
427	91.262138	50
449	91.262138	50
472	91.262138	50
494	91.262138	50

In this study, Confusion Matrix has been used to analyze the performance of neural network classification. It displays a square matrix, whose rows and columns are represented by the target column categories for classification problems or sub-ranges for regression problems for the real world target and network outputs, respectively. The practice of using the Confusion Matrix for experimenting with several values is recommended to find the best performance ratio of the projected result of proposed dataset. This study focuses on accuracy, which is one of the parameters of Confusion Matrix in this case. Table VIII and Table IX depict the results of Confusion Matrix for LM algorithm and CGD algorithm, respectively.

TABLE VIII
CONFUSION MATRIX FOR LM ALGORITHM WITH ACCURACY

	MCA	BTech CSE	MScCS	BTechIT	Accuracy
MCA	30	5	5	0	75%
BTechCSE	3	56	1	0	93.33%
MScCS	1	4	26	0	83.871%
BTechIT	1	8	0	10	52.632%

TABLE IX
CONFUSION MATRIX FOR CGD ALGORITHM WITH ACCURACY

	MCA	BTech CSE	MScCS	BTech IT	Accuracy
MCA	33	4	2	1	82.5%
BTechCSE	3	52	3	2	86.667%
MScCS	1	4	27	0	84.375%
BTechIT	1	6	3	9	47.368%

VI. CONCLUSION AND FUTURE PROSPECTS

Predicting the future trends of the above mentioned courses is the central objective of this study. Comparative analysis has also done to get better result on the Levenberg-Marquardt Algorithm and Conjugate Gradient Descent Algorithm, the two most relevant algorithms of Second Order methods to attain the primary goal. By surveying 151 people of different domains, our proposed FFANN model has analyzed all the projected scenarios and it concludes that the future trends of B.Tech (CSE) course is the highest among the stated four courses.

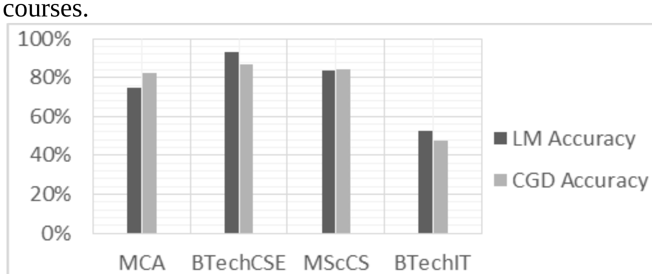

Fig 9: Comparison of accuracy of LM and CGD method

The outcome of the comparisons between two projected algorithms has been depicted in figure 9, which determines that the accuracy of the Levenberg-Marquardt Algorithm is better in this particular study.

The appropriate technique to study the future trends in terms of alternative scenarios, trend analysing and forecasting methods plays a crucial role. The above mentioned algorithms for predicting and analyzing can be further implemented or related in different areas like Green Computing, Stock Market Analysis and so on. Among all forecasting methods, NN models are proficient to deliver the best results if they are correctly configured.

REFERENCES

[1] Makvandi P., Jassbi J., and Khanmohammadi S., "*Application of Genetic Algorithm and Neural Network in Forecasting with Good Data*", In the Proceedings of the 6th WSEAS International Conference on NEURAL NETWORKS, Lisbon, Portugal, pp.56-61, June 2005.

[2] Uma Devi B., Sundar D. and Alli P., "*An Effective Time Series Analysis for Stock Trend Prediction Using ARIMA Model for Nifty Midcap-50*", International Journal of Data Mining & Knowledge Management Process (IJDKP), Vol.3(1), pp.65-78, January 2013.

[3] Sen Deepanjan, Chowdhury Dilip Roy, "*Green Computing: Efficient Practices and Applications*", International Journal of Computer Sciences and Engineering, Vol.4(1),pp.38-47, Feb 2016, E-ISSN: 2347-2693.

[4] Chowdhury D.R., Bhattacharjee D., "*Particle Swarm Optimization and Artificial Neural Network Techniques for Predicting Neonatal Disease*", In the Proceedings of National Conference on Research Trends in Computer Science and Application (NCRTCSA-2014), Technically co-sponsored by IEEE Kolkata and CSI Siliguri, pp.45-49, Feb 2014, ISBN No: 978-93-82338-95-6.

[5] Yadav A and Harit V K., "Fault Identification in Sub-Station by Using Neuro-Fuzzy Technique", International Journal of Scientific Research in Computer Science and Engineering, Vol.4(6), pp.1-7 Dec 2016.

[6] Chowdhury D.R., Chatterjee M., Samanta R. K., "*An Artificial Neural Network Model for Neonatal Disease Diagnosis*", International Journal of Artificial Intelligence and Expert Systems (IJAE), Vol.2(3), pp.96-106, August 2011.

[7] Rumelhart D.E., Hinton G.E., and Williams R.J., "*Learning representations of back-propagation errors*," Nature, Vol.323, pp.533-536, 1986.

[8] Castillo E., Guijarro-Berdiñas B., Fontenla-Romero O. and Alonso-Betanzos A., "*A very fast learning method for neural networks based on sensitivity analysis*", Journal of Machine Learning Research, pp.1159-1182, July 2006.

[9] Buntine W.L., Weigend A.S., "*Computing second derivatives in feed-forward networks: A review*", IEEE Transactions on Neural Networks, Vol.5(3), pp.480–488, 1994.

[10] Battiti R., "*First-and second-order methods for learning: between steepest descent and Newton's method*", Neural Computation, Vol.4(2), pp.141–166, 1992.

[11] Parker D. B., "*Optimal algorithms for adaptive networks: Second order back propagation, second order direct propagation, and second order Hebbian learning*", In the Proceedings of the IEEE Conference on Neural Networks, Vol. 2, pp.593–600, June 1987.

[12] Ghosh R., "*A Novel Hybrid Learning Algorithm for Artificial Neural Network*", Griffith University, 2003.

[13] Hill T., Lewicki P., "*Statistics: Methods and Applications: A Comprehensive Reference for Science, Industry and Data Mining*", StatSoft Publisher, First Edition- 2005,ISBN:1-884233-59-7.

Authors Profile

Dr. Dilip Roy Chowdhury is working as Asst. Professor in the Department of Computer Science and Application at the University of North Bengal. Before that, he has served Dept. of Computer Science & Application, Gyan Jyoti College, as HOD and took other administrative responsibilities along with regular teaching job. Dr. Roy Chowdhury's main research interests lies with the design and implementation of Expert System Development using Artificial Intelligence and Soft Computing techniques. Besides, he is continuing his research in the fields of Artificial Neural Networking, Data Mining, Rough Set Computing, Knowledgebase Design and Information Retrieval. He has more than 40 research publications in various national and international journals of repute. He is associated with various national and international journals of repute as a member of reviewing committee and editorial committee. Email: diliproychowdhury@gmail.com

Deepanjan Sen is a student of MCA of Department of Computer Science and Application, University of North Bengal. He has completed his BCA degree from Maulana Abul Kalam Azad University of Technology (Former WBUT). He has one research publicationin International Journal of repute. His broad area of research interest is Green Computing, Artificial Neural Network and Soft Computing.
e-mail: sen.deepanjan92@gmail.com

