International Journal of Engineering Innovation & Research
Volume 3, Issue 6, ISSN: 2277 – 5668

A Novel Approach for Non-Linear and Linear Data Prediction and Abstraction using Artificial Neural Network

Aman Dalmia
UG Student
amandalmia14@gmail.com

Nitish Kumar
UG Student
kumarnitish2512@gmail.com

M. K. Pandit
Professor
mkpandit.seci@gmail.com

A. K. Jana
Associate Professor
asimkjana@gmail.com

Abstract **– The ability to have prediction about the future helps us in taking measured steps in the present. In this work we have predicted data well ahead in time using a novel approach based on Artificial Neural Networks (ANNs). This paper focuses on predicting both non-linear and linear data. For linear data we design a system which can predict the Grade Point Average (GPA) of a student based on his past track record. Neural network learns using the Back Propagation Algorithm. After learning from training sets, it accurately predicts the GPA of the student. Instead of using the conventional statistical methods, we learn the pattern using artificial neural network. Similarly for non-linear data we design a system which can predict the output current of a Semiconductor Diode, given an input voltage. The accuracy obtained for predicting GPA is around 97%, while for non-linear data the accuracy is 87.5%.**

Keywords **– Electronics, Computer, Neural Networks, Data Prediction, Non-Linear, Pattern Recognition.**

I. Introduction

Artificial Neural Networks has been applied and become objects of everyday use. Their superior performance in optical character recognition, speech recognition, signal filtering in computer modems etc. have established ANN as an accepted model & method. However, neural networks have not yet been established as a valid and reliable method in the business prediction domain, either on a strategic, tactical or operational level.

Data can be classified into two types - Linear Data and Non-Linear Data. In this work, we classify the task of predicting the GPA of the students as prediction of linear data while the task of predicting the output voltage is considered a non-linear data prediction task. Students have a good bonding with their institutions. The performance of the student is important both to the student as well as the institution. Therefore, performance control is an important issue. It is necessary for both the institute as well as the student to keep a check on the performance of the students. Identifying the current status of the student is not enough. We need systems which can predict the performance of the student so that one can take proper measures and improve it. In order to achieve this we build a system which can predict the GPA of the student. The system takes in input the past scores of the students and generates a neural network. This model is then employed to predict the GPA. The I-V characteristic of a Semiconductor Diode is non-linear in nature. One needs to carry out experiments in order to know the exact characteristics. This is an expensive task. We need systems which can predict data with such non-linear characteristics as well.

One can differentiate between two types of networks, networks with feedback and networks without feedback. In networks with feedback, the output values can be traced back to the input values. However there are networks wherein for every input vector laid on the network, an output vector is calculated and this can be read from the output neurons. There is no feedback. Hence only, a forward flow of information is present. Networks having this structure are called as feed forward networks. These are various nets that comes under the feed-forward type of nets. One of the main type of feed forward network is the Back Propagation network. We use Back Propagation network to train the mode.

In the next section we discuss in detail about the Back Propagation Network. Section III discusses the past work followed by our contribution. In Section V the algorithm is explained and Section VI highlights the results of the experiments performed with linear and non-linear data.

II. Background

A. Back Propagation Network

Back Propagation is a systematic method for training multi-layer artificial neural networks. It is a supervised method and therefore needs to know the output for each input value in order to calculate the loss/error. It is a generalization of the delta rule to multi-layered feed forward networks, made possible by using the chain rule to iteratively compute gradients for each layer. It is a multi-layered forward networking using external gradient-descent based delta-learning rule, commonly known as back propagation (of error) rule. Back Propagation provides a computationally efficient method for changing the weights in a feed-forward network, with differentiable activation function units, to learn a training set of input-output example. The aim of this network is to train the net to achieve a balance between the ability to provide good response to the input that are similar.

B. Delta Learning or Back Propagation Rule.

The total squared error of the output computed by a gradient descent method known as back propagation or generalized deltarule.

Let the activation function be f(x), its derivation is F(x)

$$y_{-ink} = \sum_{i} z_i w_{jk}$$

International Journal of Engineering Innovation & Research
Volume 3, Issue 6, ISSN: 2277 – 5668

$$z_{-inJ} = \sum_i v_{ij} x_i$$

$$Y_k = f(y_{-ink})$$

The error (E) to be minimized is

$$E = (0.5\sum_k [t_k - y_k]^2)$$

The updated output weight unit is (α is the learning rate and δ is the error at output unit k)

$$\Delta w_{jk} = -\alpha \frac{\partial E}{\partial W_{jk}}$$

$$= \alpha \delta_k z_j$$

The updated output weight for the hidden unit is

$$\Delta v_{ij} = -\alpha \frac{\partial E}{\partial v_{ij}}$$

$$= \alpha \delta_j x_i$$

where *y* is the output unit, *z* is the hidden unit, *w* is the bias on output unit, *v* is the bias on the hidden unit, *t* is the output target vector, x is the input training vector.

This is the generalized delta/back propagation rule used in the network for training.

C. Architecture

A multilayer feed forward back propagation network with one layer of z-hidden units (see Fig 1). The Y output unit has W_{ok} bias and Z hidden unit has V_{ok} as bias. It is found that both the output units and the hidden units have bias. The bias acts like weights on connection from units whose output is always one. This network has one input layer, one hidden layer and one output layer. There can be any no of hidden layers. The input layer is connected to the hidden layer and the hidden layer is connected to the output layer by means of interconnection of weights. Increase in the number of hidden layers results in computational complexity of the network. As a result, the time taken for convergence and to minimize the error may be very high.

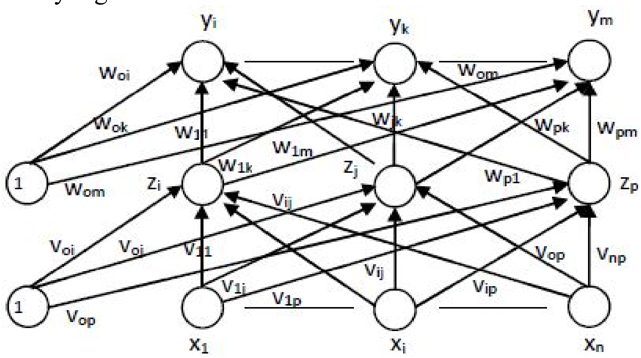

Fig.1. Architecture of Back Propagation Network

III. Literature Survey

A lot of work has been done in the area of application of Back Propagation Algorithm in the past. Examples like Patton and Timothy [1] worked on differences in the prediction of college GPA and the ACT-C scores. Chang and Shih [2] used BPN and multivariate linear regression analysis (MLR) to distinguish the classification in the mixture of organic molecules. Zimmermann et al. [3] with his group predicted the performance of post graduate level by taking into considerations of undergraduate achievements. T. Chang and Chao [4] predicted the flow of debris in the country Taiwan using BPN considering seven different factors. Gandhi and Srivastava [5] used the BPN in the field of Intrusion Detection to increase the security in Computer Networks. Wang et al. [6] worked for paleoclimate modelling where annual mean temperature and precipitation resulting in fairly accurate paleoclimate estimation. Yang et al. [7] incorporated root mean square error (RMSE), the mean absolute error (MAE) and coefficient of efficiency (R^2) for modelling process and accuracy in forecasting the ground water resources. Xueli et al. [8] used the algorithm for the wind turbine application. Their work was to narrow down the vibrations signals which were highly nonlinear and non-stationary, a fault diagnosis method for direct-drive wind turbine is been used. Li et al. [9] used improved back propagation algorithm in which they have introduced self-adaptive learning rate for fault diagnosis in air cooling condenser. Jiang et al. [10] worked on image processing where they improve the high resolution remote sensing image using back propagation algorithm. Lu [11] worked on geographical terrain slope. Their investigation was to find the mean slope of an area where they get the input from geographical information system (GIS). Kaensar [12] used the algorithm for recognition of handwritten digit. They worked on different structures like Simple Back Propagation, Back Propagation with momentum terms and Back Propagation using conjugate gradient descent working on different parameters and finding out the suitable structure with the relevant parameters in order to find better recognition. Sengto and Leauhatong [13] used the algorithm in reducing the fall fatalities of elderly people. They worked with tri-axial accelerometers which help in evaluation of fall occurrence. Du et al. [14] used the BPN in predicting the rainfall over the period in a particular area. Their results shows higher accuracy and with better stability.

IV. Our Work

In this work we **propose** a**novel** system which can predict the data irrespective of its characteristics. More specifically, we predict the GPA of the students which is **linear** in nature. We also predict the current vs. voltage (I-V) characteristics of an electronic device, namely, Semiconductor Diode when an input voltage is given which is **non-linear** in nature. Our model is general in nature, in the sense that it can handle both linear and non-linear systems.

Prediction of marks helps the students as well as the institute to analyze the performance of the student. By knowing the GPA in advance, the student can take necessary steps in order to improve their performance. We use Back Propagation Algorithm in order to model the past marks of the students. The predicted value enables the students to provide ample time to work on their GPA and score more than the predicted marks. A weighted linear combination of the marks of a particular semester(s) of a student is replace by a single value. This helps in analyzing a student's performance quickly and effectively. In our system we are evaluating the student performance and compress many subjects' marks into a single value. The system used the algorithm for predicting the GPA for ten students. We used the actual data marks of 5^{th} and 6^{th} semester having 17 different parameters (total no of subjects of 5^{th} and 6^{th} semester is 17) as training input and used the combined YGPA of 7^{th} and 8^{th} semester as target training output. Due to data compression, the space complexity reduces thus enabling to analyze a student better as in this case.

For prediction of I-V characteristics we perform an experiment on a lab on different Diode (having same configurations). There are many cases where we don't have proper resources for evaluation of Diode characteristics. The data which we got is used for training the network. One can predict the output current by giving voltage as an input.

A. Algorithm

Step 1: Initialize weights to small random value.]

Step 2: For each input vector follow steps 3-5.

Step 3: For i=1,…., n; set activation of input unit, x_i ;

Step 4: Feed Forward

For j=1,…., p;

$$z_{-inj} = v_{oj} + \sum_{i=1}^{n} x_i y_j$$

applying activation function $Z_{-inj} = f(z_{inj})$

Step 4: For k=1,…., m;

$$y_{-ink} = w_{ok} + \sum_{j=1}^{p} z_j w_{jk}$$

applying activation function

$$Y_{-ink} = f(y_{-ink})$$

Step 5: Finding out Back Propagation Error

Step 6: Updation of Weights and Biases

V. Experiments And Results

A. Back Propagation Network

Dataset: In order to experiment we use different data sets for the task of predicting linear and non-linear data. In order to predict linear data, we use the marks of the students for a semester. For non-linear data prediction task we generate the I-V characteristic of Diode.

B. Experimentation using Back Propagation Algorithm

During experimentation we vary the percentage of training and testing data. In Table I we report the results for linear data and Table II reports for Non-Linear data and obtained error and efficiency for different percentage of training and testing data. It is found that in **both the cases** accuracy of the system is maximum on taking 80% of the data as training and 20% as testing.

Table I: Comparison of different Input Training and Testing datawith Efficiency for Linear Data

Input Data		Error	Efficiency
Training Data	Testing Data		
20%	80%	.2176	78.237%
40%	60%	.1463	85.363%
60%	40%	.1061	89.382%
80%	20%	.0763	92.364%

Table II: Comparison of different Input Training and Testing data with Efficiency for Non-Linear Data

Input Data		Error	Efficiency
Training Data	Testing Data		
20%	80%	4.3057	50.0%
40%	60%	3.442	50.0%
60%	40%	2.3625	75.0%
80%	20%	2.1303	87.5%

C. Linear Data

Next we predicted the SGPA of the students. Table III reports the actual and the predicted SGPA of ten students for a particular semester. Since the accuracy of the system was not very high, we tuned the parameters of the system. Fig 2 shows the graphical representation of the same. After some more experimentation the best accuracy if obtained by decreasing the learning rate by 0.01 i.e. 0.08, fixing the momentum as 0.4. Table IV reports the accuracy for these parameters. As evident in Fig 3, the gap between actual and predicted GPA values is reduced considerably.

Table III: Results of single semester marks [learning rate=0.09; momentum=0.5; Efficiency = 86.2602%; Compression Ratio=1.67]

Name	Actual SGPA	Predicted SGPA	Error (%)
A	8.2900	8.2096	0.9698
B	8.4800	8.1102	4.3608
C	8.5200	8.4162	1.2183
D	8.5200	8.6154	-1.1197
E	7.4300	7.9979	-7.6433
F	8.4800	8.1084	4.3821
G	8.6700	8.2926	4.3529
H	8.3300	8.1675	1.9507
I	8.6200	8.4591	1.8665
J	9.3800	8.0192	14.5074

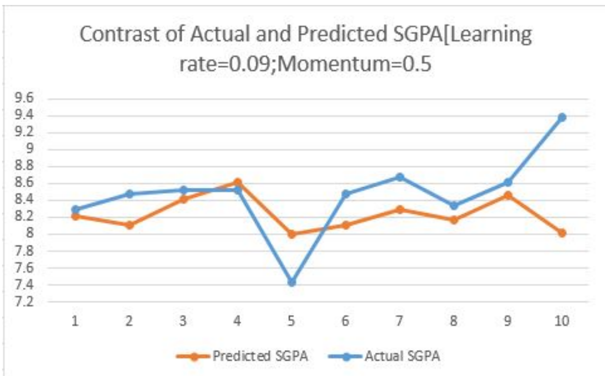

Fig.2. Contrast of predicted SGPA and actual SGPA [Table III data]

Table IV: Prediction of SGPA
[Learning rate=0.08; momentum=0.4; Efficiency=97.2701%; Compression Ratio=3]

Name	Actual SGPA	Predicted SGPA	Error (%)
A	8.2900	8.4584	-2.0313
B	8.4800	8.3913	1.0459
C	8.5200	8.5553	-0.4143
D	8.5200	8.4139	1.2453
E	7.4300	8.3144	4.7604
F	8.4800	8.4692	0.1273
G	8.6700	8.5215	1.7128
H	8.3300	8.3704	-0.4826
I	8.6200	8.4071	2.4698
J	9.3800	8.5044	9.334

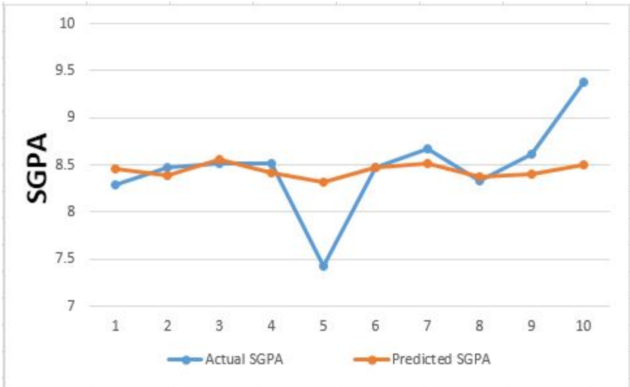

Fig.3. Contrast of predicted SGPA and actual SGPA [Table IV data]

Table V: Prediction of YGPA [learning rate=0.03; momentum=0.4; Efficiency=95.0%; Compression Ratio=5.66]

Name	Actual YGPA	Predicted YGPA	Error (%)
A	7.7700	8.4906	-9.2741
B	8.7300	8.4479	3.2313
C	8.6500	8.5943	0.6439
D	8.1500	8.4814	-4.0662
E	8.5600	8.3497	2.4567
F	7.5400	8.4994	-12.7241
G	8.0400	8.5103	-0.4703
H	8.3800	8.5082	-5.8495
I	8.1900	8.4699	-3.4175
J	7.8300	8.4729	-8.2107
K	8.8500	8.3431	5.7276
L	8.6700	8.4965	2.0011
M	9.0000	8.5884	4.5733
N	8.4600	8.4019	0.6867
O	9.3300	8.4022	9.9442
P	8.8100	8.5526	0.2574
Q	8.0000	8.5635	-2.9216
R	9.1900	8.4942	7.5712
S	7.9200	8.4880	7.1717
T	8.5200	8.5740	0.6338

Next we worked on the YGPA of twenty students. In this experiment we take the number of parameters as total subjects covered in two semesters which is seventeen. We take the training data as two semester marks. Table V reports the actual and predicted YGPA of twenty students, keeping the optimum Learning Rate and Momentum which we had reported in the next section.

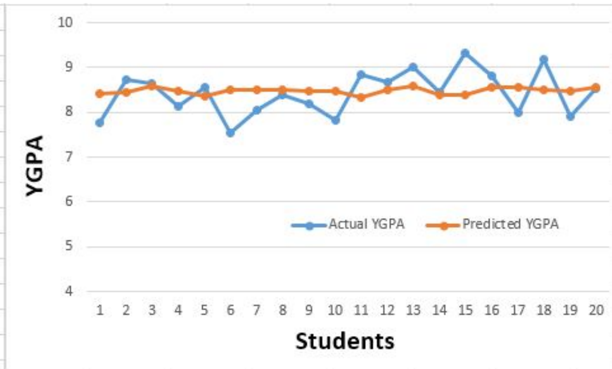

Fig.4. Contrast of predicted YGPA and actual YGPA [Table V data]

D. Optimum Learning Rate and Momentum

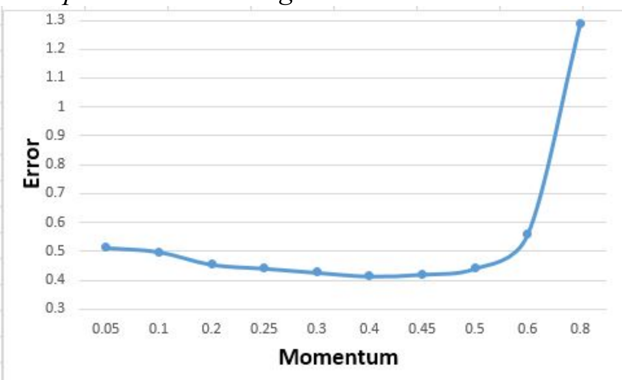

Fig.5. Linear-Momentum vs. Error [Minimum error at 0.4]

While executing the code for predicting YGPA we compared the values of Learning rate and the momentum, such that the error is minimum. It was found that the Momentum for the minimum error is at 0.4 (see Fig 5). Learning rate for which the error is minimum is found to be 0.03 (see Fig 6). Hence we use the best value of the Momentum and Learning Rate for our application for predicting the YGPA. It was found in the Table V the efficiency of our system is to be 95.0%.

Fig.6. Linear-- Learning Rate vs. Error [Minimum error at 0.03]

E. Non-Linear Data

Many real world applications consists of non-linear data. In order to predict nonlinear data we first generated our own dataset by using a Diode and obtained the I-V characteristics. After we obtain this dataset, this data is used for training and the intermediate input voltage value is used for testing.

Table VI: Prediction of output Current based on given input Voltage [learning rate=0.8; momentum=0.4; Efficiency=87.5%; Compression Ratio=3]

Input Voltage(mV)	Actual Output Current(mA)	Predicted Output Current(mA)	Error (%)
0.57	5.5	8.6071	-56.49
0.59	8.0	9.4303	-17.87
0.61	9.5	9.626	-1.32
0.63	13.0	12.7871	1.63
0.65	20.0	16.9237	15.38
0.67	32.0	27.2291	14.91
0.69	42.0	40.4672	3.64
0.70	55.5	53.2821	3.99

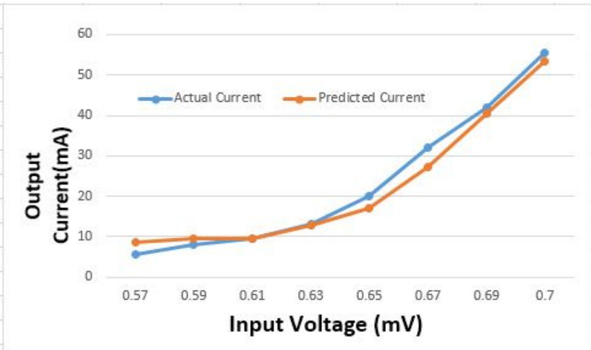

Fig.7. Contrast of predicted Current and actual Current [Table IV data]

The efficiency calculation process for the non-linear data is different than that of linear data. For non-linear data we apply a method that if the error is less than 20% we ignore the error otherwise the error is taken into consideration. Hence, in this case we have 8 data from which there is only one data whose error is greater than 20% (see Table VI). This is the only data which is considered as an error in this calculation. So, the efficiency is 87.5%.

F. Optimum Learning Rate and Momentum

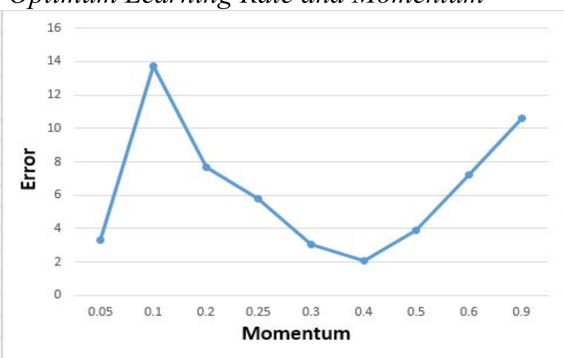

Fig.8. Non-Linear --Momentum vs. Error [Minimum error at 0.4]

While execution of non-linear systems of I-V characteristics of Diode it was found that the system gives the best performance when the Momentum is 0.4 (see Fig 8) and Learning Rate parameter is 0.8 (see Fig 9). The efficiency of the system was found to be 87.5%.

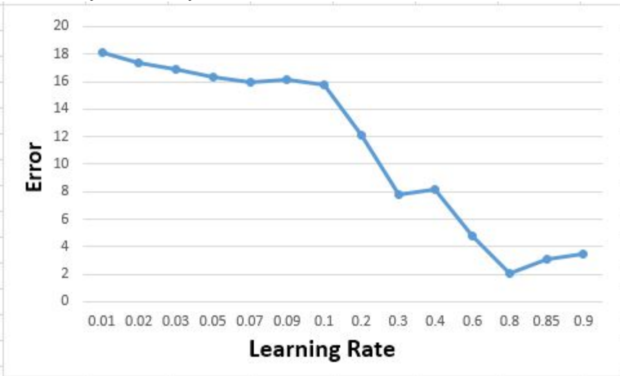

Fig.9. Non-Linear--Learning Rate vs. Error [Minimum error at 0.8]

VI. Conclusion

An approach of data compression and prediction is presented. It is based on artificial neural network using a model named as Back Propagation Algorithm. The prediction of marks can be used in any institute. Improved performance of the student leads to satisfaction of the student. Also, it adds to the repute of the institute as well. In case of non-linear data as in this case we would easily predict the output of any components without giving any bias to it. Thus for supervised input pattern, the output is obtained with good level of accuracy.

In spite of the fact that the data availability for this research was not optimal, it can be concluded that the presented approach has a potential in solving the problems, allowing for efficient corrections of predictions.

References

[1] Patton, Timothy K. (1998). "Differential Prediction of College Performance between Gender," in Education Resources Information Center (ERIC). Available: http://files.eric.ed.gov/fulltext/ED427030.pdf.

[2] Ping Chang and Jeng-Shong Shih, "The Application of Back Propagation Neural Network of Multi-channel Piezoelectric Quartz Crystal Sensor for Mixed Organic Vapours", *Journal of Science and Engineering*, vol. 5, No. 4, 2002, pp. 209-217.

[3] J.Zimmermann, K.H. Brodersen, J.P. Pellet, and J.M Bhumann, "Predicting graduate-level performance from undergraduate achievements". ETH Zurich.

[4] Tung-Chueng Chang and Ru-Jen Chao. "Application of back-propagation networks in debris flow prediction". *Engineering Geology*, 2006, pp. 270–280.

[5] Meera Gandhi and S.K Srivatsa, "Application of Back Propagation Algorithm in Intrusion Detection in Computer Networks". *International Journal of Soft Computing*, 2008, pp. 277-281.

[6] Hongli Wang, Xueyuan Kuang, and Jian Liu, "Application of back propagation neural network in paleoclimate". *International Conference on Information Science and Technology (ICIST)*, 2011.

[7] Zhongping Yang, Wenxi Lu, Yuqiao Long, and Ping Li, "Application of Back-Propagation Artificial Neural Network Models for Prediction of Groundwater Levels: Case study in

Western Jilin Province, China". *International Conference on Bioinformatics and Biomedical Engineering (ICBBE),* 2008, pp. 3203 – 3206.

[8] Xueli An, Dongxiang Jiang, and Shaohua Li, "Application of back propagation neural network to fault diagnosis of direct-drive wind turbine". *World Non-Grid-Connected Wind Power and Energy Conference (WNWEC)*, 2010, pp. 1-5.

[9] Yong Li,Yang Fu, Si-Wen Zhang, and Hui Li,"Improved algorithm of the Back Propagation neural network and its application in fault diagnosis of air-cooling condenser". *International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR)*, 2009, pp. 180-184.

[10] Jiefeng Jiang,Jing Zhang, Gege Yang, and Dapeng Zhang, "Application of back propagation neural network in the classification of high resolution remote sensing image: Take remote sensing image of Beijing for instance". *International Conference on Geoinformatics, 2010*, pp 1 - 6.

[11] Ming Lu. "Improved neural network modeling approach for engineering applications". *Proceedings of the 9thInternational Conference on Neural Information Processing (ICONIP)*, 2002, pp. 18-22.

[12] C. Kaensar, "Analysis on the Parameter of Back Propagation Algorithm with Three Weight Adjustment Structure for Hand Written Digit Recognition". *International Conference on Service Systems and Service Management (ICSSSM),* 2013, pp. 17-19.

[13] A. Sengto, and T. Leauhatong, "Human falling detection algorithm using back propagation neural network". *Biomedical Engineering International Conference(BMEiCON),*2012,pp.1- 5.

[14] Jiaxing Du, Yunnan Teng Chong Bur., Tengchong, China; Bin Zhao, and Shaohui Miao, "An Application on the Immune Evolutionary Algorithm Based on Back Propagation in the Rainfall Prediction". *International Conference on Computer Science and Electronics Engineering (ICCSEE),* 2012, pp. 313 – 317.

AUTHOR'S PROFILE

Aman Dalmia
is pursuing his Undergraduate in Electronics and Communication Engineering at Haldia Institute of Technology, Haldia, India. His field of interests are Robotics, Machine Learning, and Artificial Intelligence. He is working on Machine Learning as his final year project.

Nitish Kumar
is pursuing his Undergraduate in Electronics and Communication Engineering at Haldia Institute of Technology, Haldia, India. His field of interests are Microprocessors and Microcontroller, Machine Learning and Artificial Intelligence. He is working on Machine Learning as his final year project.

Malay K. Pandit
received his B.E and M. E degrees in Electronics and Telecom Engg, from Jadavpur University, India in 1989 and 1991, respectively. He received his PhD from UK's renowned Cambridge University in 1996. He did his post-doc from the Optoelectronics Research Centre, City University of Hong Kong till 2002 where he pioneered the use of polymers for optical waveguide applications. He then took a corporate career where he worked in a fibre optic company "FONS (I) Ltd" in the domain of optical networking. He has 45 international publications in this area. Dr. Pandit was awarded the National Scholarship of the Government of India in 1983 and the Nehru Scholarship of the Government of India during his Ph.D. studies. Currently he is a Professor in the Dept. of Electronics and Comm. Engg. at Haldia Institute of Technology, Haldia, India where he focuses on embedded systems, including their usage in WDM optical networks.

Asim K. Jana
received his B.E. in Electronics and Telecommunication Engg and M.E in Production Engg, from Jadavpur University, India in 1988 and 1995, respectively. He has 10 years of industrial experience in the domain of embedded systems. Currently he is an Associate Professor in the Dept. of Electronics and Comm. Engg. at Haldia Institute of Technology, Haldia, India where he focuses on Computer Systems. He has 10 international research papers.