ISSN 2278 – 0211 (Online)

Determining Students Performance Using the Tool of Artificial Neural Network

Jateen Shet Shirodkar
Assistant Professor, Department of General Engineering (Mechanical),
Shree Rayeshwar Institute of Engineering and Information Technology, Goa, India
Viren Pereira
Faculty, Shree Rayeshwar Institute of Technology, Goa, India

Abstract:
Technical Education has grown rapidly over a few decades and is one of the key drivers in the knowledge driven economy. The systematic growth of quality in technical education plays a vital role in development. In any academic organization, early prediction of student's performance is very important to management so that strategic intervention can be pre-planned before they could appear for the semester examination. Artificial Neural Network (ANN) is a superior mathematical tool used to identify trends in data by developing relationships between various inputs and outputs. Feed Forward Back Propagation Network (BPN) is powerful technique of ANN as it can simulate any continuous or non-linear function. The proposed model allows prediction using a Feed Forward Neural Network (FFNN). The trained model helps to accurately predict at-risk students and reduces the student dropout rate. The output of this study showed that first semester percentage is strongly influenced by fundamental subjects. Comparison between predicted and actual output indicated that the ANN model holds promise for estimation of student's performance.

Keywords: *Artificial neural network, prediction, multilayer, back-propagation network, students performance, weights*

1. Introduction

The Indian economy has been growing at the rate of around 8 percent per year. A planned and integrated approach is required to put India on the development track. Orienting technical education towards the right path will determine the future of Indian economy. For the past 25 years there had been an exponential growth in the number of engineering institutions. Due to an increase in number of engineering institutions over last few years across India, there is an increase in intake capacity which has resulted in number of engineering seats remaining vacant every year. Many colleges are therefore facing problems for enrollment process. For survival, institutions need to adopt innovative strategies to attract the stakeholders. To maintain reputation, engineering institution must have continuous improvement in results and have better and sustained placement records. Over the years, most of the top ranked companies for campus placement insist on a first class throughout the engineering course. Hence marks scored in each semester have a great impact on selection criteria for campus placement during final year. Early detection of percentage of marks can be used to identify such students and later special treatment can be provided like counseling through mentorship, tutorials for below first class level students.

Every institute aims in evaluating and predicting student's academic performance to determine what percentage of students fall in intelligent, average or poor category. At the beginning of the course, the teachers need to know the levels of their students, which may provide them a guideline to decide how much effort they have to put in while teaching in the classroom. This may also guide them either to lower down or to exalt their standard of teaching in classroom for imparting adequate knowledge for satisfactory academic progress.

Typically, the first-year students are at greatest risk of failure from the study because of delay in admission process as well as due to increase in difficulty level as compared to higher secondary education. The students who join late to engineering institution fail to capture the basic concepts taught by the teachers during the start of the semester which results in students facing difficulty while application of those concepts. In the present study, past data of fundamental subjects is collected from admission forms of students of Goa University. The study only considered the fundamental subjects scored by the students at higher secondary level without considering family background.

2. Artificial Neural Network

The term artificial neural network gets inspiration from neurobiological studies of human brain which is a massively interconnected network of neurons as shown in figure 1. The artificial neurons have input connections (dendrites); these neurons interact with their neighboring neurons through transmission lines (axon) which are connected through synapses to dendrites of other neurons [1, 2]. Transmission of signal from one biological neuron to other is on account of chemical activity. Effect of this activity is to increase or lower potential. When potential is increased beyond certain threshold, we say that neuron is fired and signal is transmitted through synaptic junction. Generally, the chemical activity is confined in cell region of a biological neuron as shown below [3]. In an inactive state, the protoplasm is negatively charged against the surrounding liquid having Na^+ ions. Resting potential is -70 mV at which cell membrane is impervious to Na^+ ions. This causes deficiency of sodium ions in the cell. The signals received from the neighboring neurons can cause temporary depolarization of resting potential. When potential is increased to -60 mV, the membrane loses impermeability for Na^+ ions. Na^+ ions enter the cell and reduce the potential difference. This sudden change in potential causes the neuron to discharge and neuron is said to have fired. Neuron is fired if signal is positive, else remains inactive. The cell body acts as a summing device which receives signals from the various other neurons and when the potential difference has reached threshold (10 mV), neuron fires. The various axons carry electric signals having different strengths. The strength of the signal determines the extent of knowledge passed on to the connecting neuron. In general, ANNs are most useful in tasks such as model selection and classification, function estimate, determination of the optimum value and data categorization [4].

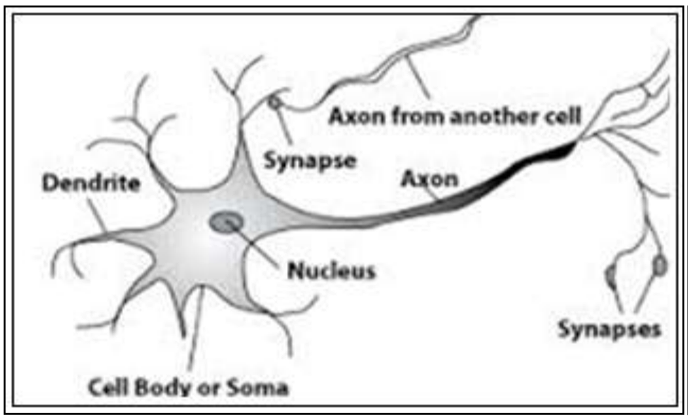

Figure 1: Simple model of Biological Neuron

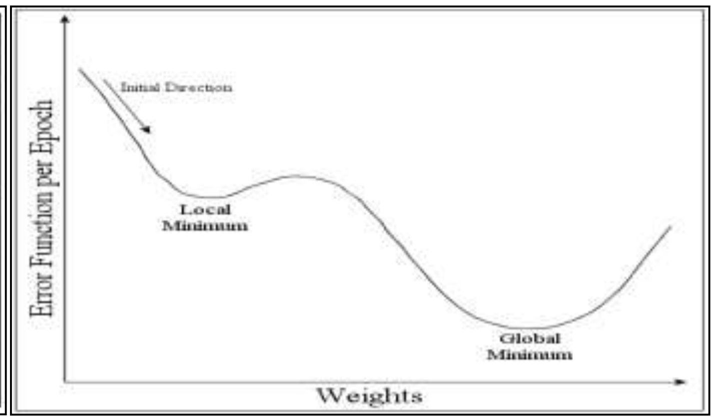

Figure 2: A plot showing Local and Global minima.

Quantitative methodologies like Regression analysis, moving average method, ARIMA (auto regressive integrated moving average), exponential smoothing method, Markov analysis which can predict using historical data but most of the time predication or forecast goes wrong as quantitative methods used to model fails to capture various parameters such as trend or pattern in data, seasonality etc. Even the use of differential equation to map the function relationships proves mathematically exhaustive as unfortunately, there are few differential equations which have exact solutions [5-7]. In last two decades, ANN has been extensively used in engineering and management as a data modeling tool. Multilayer Back-propagation Network trained with error back-propagation algorithm is found to be a successful prediction tool of ANN [8-12]. In practice the BPN has one or more layers of neurons between input and output layer called hidden layer. Activation function used is normally a function of the form $f(x) = \frac{1}{1+e^{-x}}$ [13].

The effectiveness and optimal convergence depends on the value of constant learning rate (η). By using momentum constant we can accelerate the conversion of error back-propagation learning and avoids the problem of getting stuck at local minima instead of Global minima [2, 12] as shown in figure 2. Selection of weights, bias weights, learning rate and momentum factor is done randomly because of which algorithm has to re-iterate several number of times. As a result the final output predicted by the network varies at every instance the training data is fed to the network. The algorithm then tries to minimize the error between actual and predicted values [14-17]. Selection of proper momentum factor and learning rate is very important as it saves time and reduces the error at faster rate [18]. Once a BPN is trained, the number of hidden neurons and the weights are fixed. In present research, it is proposed to predict the first semester percentage by using FFNN trained using Back Propagation Algorithm.

3. Literature Survey

A study of literature shows that various researchers have applied artificial neural networks in last two decades for various applications in prediction and classification.

Research conducted by Kumar [10] on the possibility of predicting average rainfall over Udupi, district of Karnataka has been analyzed through various ANN models namely Back Propagation Algorithm (BPA), Layer Recurrent Network (LRN) and Cascaded Back-Propagation (CBP). Input parameters were the average humidity and the average wind speed whereas the output parameter is average rainfall. It was concluded that Larger the amount of input data, lower is the MSE after training. It was also concluded that BPA is the best algorithm out of the three tested for convergence.

Researchers [11] recommended some practical steps and interventions that the management can take to improve the Electrical degree students' academic performance before graduation. They also suggested that those students with low CGPA at early semester may not fully absorb the core subjects. They concluded that there is a direct correlation between students' strong academic ability on fundamental subjects at early semester and their overall academic performance upon graduation. This leads to a conclusion that fundamental subjects must be fully understood and grasped because without them understanding of subjects in the subsequent higher semesters will be very difficult.

In 1995 Michaelides et al compared the performance of ANN with multiple regressions in estimating missing rainfall and found that ANN produced better results than multiple regressions [19]. Study using ANN technique conducted by Oladokun, Adebanjo and Owaba [20] observed that a very poor quality of graduates of Nigerian Universities has been partly traced to inadequacies of the National University Admission Examination System. This study has shown the potential of the Artificial Neural Network for enhancing the effectiveness of a university admission system. Experiments carried by Toth, Brath and Montanari [21] compared short-time rainfall prediction for studying flood forecasting which uses auto-regressive moving average (ARMA) model, ANN and nearest-neighbors approaches were applied for forecasting. Koizumi in 1999 applied ANN model to results obtained by the Japan Meteorological Agency (JMA) Asian Spectral Model and training was carried out with one year data. He found that the ANN was better than forecast by the linear regression prediction, and the numerical model. It was also suggested that performance of ANN can be improved if data sets are large [22].

Research carried out by Braunstein [23] used correlation and regression techniques to determine the variables most closely related to academic success of recent graduates of an MBA program. They found undergraduate GPA and GMAT score to have the strongest positive correlation with the graduate GPA. However these statistical models usually assume normality and variances. When these assumptions are violated in real world data structures, the predictive ability of regression model is diminished. Walczak and Sincich [24] compared results of NN model to that of logistic regression analysis for modeling student enrollment decision-making to show improvements gained by using NN. The authors have concluded that the level of performance of NN is slightly higher than statistical models.

4. Back-Propagation Network Algorithm

4.1. Basic Description of Artificial Neural Network

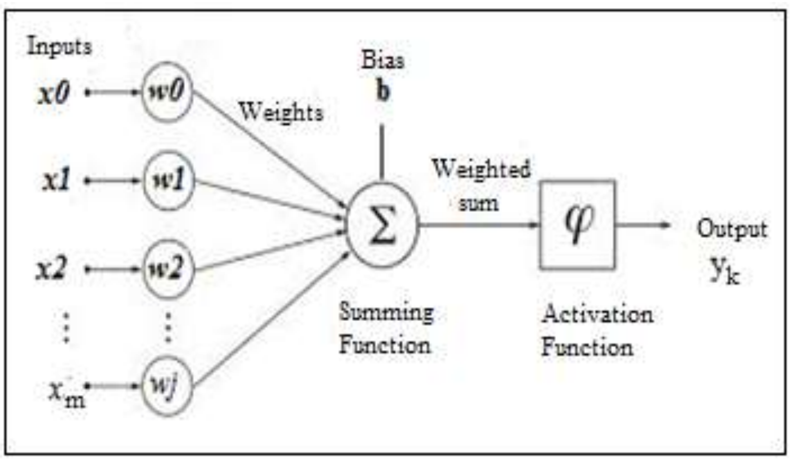

Figure 3: Simple model of Artificial Neuron

ANN simulates the complicated functioning of biological neuron into electrical signals by using mathematical operations. The ANN is having three or more number of layers, each composed of certain number of hidden neurons. Hidden layers, which can be one or more in number, are sandwiched between the input and output layers. Bias is also introduced at hidden and output layer for threshold adjustment. Figure 3 represents a simple computational model of ANN. Let x_0, $x_{1,\ldots,}x_m$ are m input parameters. W_j is the synaptic weight connecting the hidden node. Synaptic weights represent the synapse strength of network neurons. Each node receives a signal from the nodes of the previous layer and these signals are multiplied by separate synaptic weights. The weighted inputs are summed up to get the total input. The weighted sum is then passed through an activation function, which produces the output Y_k. The output activation function is then transmitted to next layer and similar calculations are carried out. The final output is produced at output layer. This is called feed forward pathway and flow of information takes place in this direction. The prediction accuracy of ANN depends on training. During training optimization is required on weights. Among various training algorithms available to train ANN, the back-propagation algorithm is known to be most popular [2, 13].

4.2. Modeling of Network

ANN model is constructed to predict the first semester percentage of first year engineering students. In the present work, a FFNN was employed which is trained with the error-back-propagation algorithm and an activation function as sigmoidal. The FFNN model consists of an input layer, an output layer and one or two hidden layers. The NN model was coded using MATLAB programming language. To design the neural network, datasets for training and testing were normalized between 0.1 and 0.9. During training, the influence of the number of hidden layers and the number of neurons in the hidden layer on the convergence criterion was deeply studied. Training was stopped when the mean squared error (MSE) was reduced to desire or the maximum number of iterations reached, whichever occurred first. The maximum number of iterations executed was limited to prevent over fitting of the model. Initially, the learning rate (η) and the momentum constant (α) were taken as 0.90 and 0.10, respectively. The number of hidden layers and the number of neurons in the hidden layers were fixed based on the MSE. The MSE was given by

$$MSE=\frac{1}{n}\sum_{i}^{n}(Yp-Yi)^{2}$$

There is no fixed formula for selecting the optimum of hidden neuron. However, some thumb rules are available for calculating number of hidden neurons and a few of them are listed below:

- "A rule of thumb is for the size of this hidden layer to be somewhere between the input layer size and the output layer size.'' by Blum [25].
- Swingler [26] proposed that "You will never require more than twice the number of hidden units as you have inputs'' in an MLP with one hidden layer."
- "How large should the hidden layer be? One rule is that it should never be more than twice as large as the input layer.'' by Berry [27].
- "Typically, we specify as many hidden nodes as dimensions needed to capture 70% to 90% of the variance of the input data set." suggested by Boger [28].

5. Comparison of Actual Data and Predicted Data

Training was performed by varying the learning rate and momentum constant from 0.1 to 0.9 with a step size of 0.05. Based on the MSE, the learning rate and momentum constant was selected. The number of neurons in the hidden layer and the learning rate and momentum were then fixed while training the network. The optimum NN model architecture achieved is used for predicting the first semester percentage. The model used in present study consists of data collected which includes marks at HSSC examination as well as higher secondary marks in subjects such as Physics, Chemistry and Mathematics (PCM), GCET Score. Selection of learning rate, momentum constant, number of hidden neurons and hidden layers is a trial and error process. The NN model showed excellent convergence with one hidden layer and six neurons in hidden layer. The network converged with η = 0.75 and α = 0.25. After the testing is done, the results are saved and a graph is plotted between the actual output and the predicted output so that a comparison can be made. The graph is an efficient way of comparing the results.

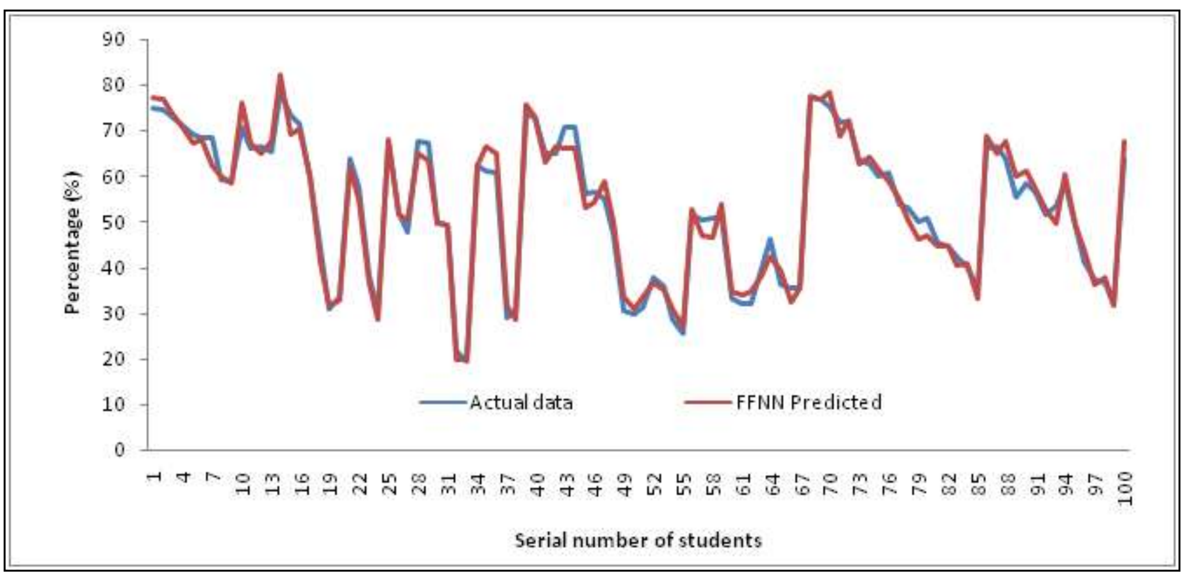

Figure 4: A graph showing comparison between Actual and FFNN Predicted values

6. Conclusion

The study on the data collected for predicting academic performance of students shows similarity between actual and predicted output. This indicates that the NN model can be used to predict student academic performance. It is observed that as the number of neurons increases in the hidden layer, the MSE decreases up to a threshold limit. It is also found that a larger bank of input data improves the

prediction capability of the network to a great extend. In the process of training the network, it was found that using two hidden layers rarely improves the model and it also introduces a greater risk of converging towards local minima. Network with large number of nodes and connections are capable of memorizing the training set. So, it is better to use smaller size of networks. Learning constant helps in minimizing the error whereas momentum factor speed up the process of conversion. Gross use of momentum factor may lead to increase the error so proper selection of momentum factor is also required. The problem of what architecture to select is still under study with no strong answer yet. It was found to be a trial and error method and left up to the network builder.
However such model is not limited to first year degree students and can be extended to all four years of degree course with suitable input variables. With this knowledge strategic intervention can be done to help below average students to improve the overall academic performance of students. Furthermore, model can be use to carry out sensitivity analysis of every input vector so that impact of each input on the end semester percentage can be identified.

7. References

i. S. Haykin, Neural Networks, A Comprehensive Foundation, Macmillan, NY, USA, 1994.
ii. J. M. Zurada, Introduction to Artificial neural systems, ISBN 81-7224-650-1.
iii. R. E. Uhrig, Introduction to artificial neural networks Proceedings of the 1995 IEEE IECON 21st International Conference on, vol.1, pp.33-37 vol. 1, 1995.
iv. M. Veronez, S. De Souza, M. Matsuoka et al, Regional mapping of the Geoid using GNSS (GPS) measurements and an artificial neural network. Remote Sens 3(4):668–683. doi:10.3390/rs3040668.
v. S. Makridakis, S.C. Wheelwright, R. J. Hyndman, Forecasting Methods and Applications, John Wiley & sons, Singapore (2003).
vi. W. Steavenson, Operation management, ISBN-13: 978-0073525259.
vii. S.S. Rao, Engineering Optimization, theory and practice, ISBN 978-81-224-2723-3.
viii. N. W. Clarence, Tan & Gerhard E, A study of parameters of a backpropagation Stock market price prediction Model, IEEE reference: 0-8186-4260-2/1993.
ix. D. A. Komo, C. Chang, H. ko, Neural network technology for stock Market Index Prediction, IEEE reference: 0-7803-1865-x/1994.
x. A. Kumar, A. Kumar, R. Ranjan, S. Kumar, A Rainfall Prediction Model using Artificial Neural Network in 2012 IEEE Control and System Graduate Research Colloquium (ICSGRC 2012).
xi. P. M. Arsad, N. Buniyamin, J. Manan, Neural Network Students' Performance Prediction Model (NNSPPM) in Proc. of the IEEE International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA), 26-27 November 2013, Kuala Lumpur, Malaysia.
xii. S. Mohammed, E. Abbasy, A. Senouci, T. Zayed, F. Mirahadi, L. Parvizsedghya, Artificial neural network models for predicting conditions offshore oil and gas pipeline, www.elsevier .com/ locate/autcon.
xiii. D. E. Rumelhart, G. Hinton, R. J. Williams, learning in ternal representations by error propagation, In parallel distrbuted Processing, Vol 1, MIT press cambridge. 1986 pp.318-362.
xiv. E.J. Dayhoff, Neural Network Architecture an Introduction, VNR Press, New York, 1990.
xv. R. P. Lippman, An introduction to computing with neural nets IEEE ASSP, Magazine, (1987) 4.
xvi. K. Hornik, H. White, Neural Netw. 2 (1989) 359.
xvii. D. E. Rumelhart, E.H. Geoffrey, J.W. Ronald, Nature 323 (1986)533.
xviii. N. S. Reddy, Study of some complex metallurgical systems by computational intelligence techniques.
xix. S. C. Michaelides, C.C. Neocleous & C. N. Schizas Artificial neural networks and multiple linear regression in estimating missing rainfall data. In: Proceedings of the DSP95 International Conference on Digital Signal Processing, Limassol, Cyprus.
xx. V. O. Oladokun, A.T. Adebanjo, O.E. Owaba, Predicting Students Academic Performance Using Artificial neural Network: A case study of an Engineering Course, 2008, The pacific journal of Science and Technology, volume 9, number1 (Spring).
xxi. E. Toth, A. Brath, A. Montanari, Comparison of short-term rainfall prediction models for real-time flood forecasting, Journal of Hydrology.
xxii. K. Koizumi, An objective method to modify numerical model forecasts with newly given weather data using an artificial neural network.
xxiii. A. W. Braunstein, Factors determining success in a graduate business program, College Student Journal, (2002). Vol. 36, pp. 471–477.
xxiv. S. Walczak and T. Sincich, A comparative analysis of regression and neural networks for university admissions, (1999) Information Sciences, Vol. 119, pp. 1–20.
xxv. A Blum, Neural networks in C++: An object oriented framework for building connectionist systems (1992), Cambridge Information Group, John Wiley & Sons. Inc.
xxvi. K. Swingler, Applying Neural Networks, A Practical Guide, (1996), Academic Press, Larcourt Brace & company, Publishers, London.
xxvii. Berry and G.S. Linoff, Data Mining Techniques for Marketing, Sales, and Customer support, (1997), John Wiley & sons, Inc.
xxviii. Z. Boger and H. Guterman, Knowledge extraction from Artificial Nesural Network models, (1997), IEEE Systems, Man and Cybernetics Conference, Orlando, FL.