PREDICTING STUDENT ATTENTION DURING LECTURES THROUGH ARTIFICIAL NEURAL NETWORK: A SURVEY

Kavitha Elango
Assistant professor, Department of computer science, Indian Academy Degree College Autonomous, Bangalore, Karnataka

Abstract: Drawing the attention of student for the complete lecture session has always been a central concern of educators in almost all institutions. This is due to the lack of concentration of students in between the lectures for several reasons. An early prediction of student attention helps the educators to identify the student level of learning and to improve their delivering mechanism in the teaching-learning process. The identified set of input feature variables are classified into two groups such as student factors considered as primary parameters and teacher factors considered as secondary parameters. This paper presents a brief survey on BPNN and PNN in order to find out the perfect, efficient solutions for classifying the student attentiveness into different categories.

Keywords: Student attentiveness, Teaching-learning process, Back Propagation Neural Network (BPNN), Probabilistic Neural Network (PNN).

1. INTRODUCTION

Nowadays, all educational institutions are focusing to improve the student learning capabilities so as to improve the quality of teaching as teachers are playing the crucial role in awakening the student knowledge. There are several reasons which influence the students to lose their attention in between the lectures including the student past history, family background and teacher's contributing factors. This paper presents the input variables which influences the student to lose attention and analyses the different ANN techniques to predict the student attentiveness in the class room. An early prediction of assessing the student attention would help the instructors to fine tune their delivering mechanism so as to improve the student performance in the final exam.

Neural network have emerged an important key component in many data mining applications due to their abiity to "learn" from the data, their nonparametric nature, and their ability to generalize.(Priyanka Gaur,2014).Neural network method in data mining is used for classification, clustering, feature mining, prediction and pattern recognition (Sanjesh Ghore,2014). So far researchers have used ANN for predicting student academic performance(Nidhiarora, Jitinderkumar R .Saini,2013), student enrollment decision(Malaya Dutta Borah, Rajni Jindal , Daya Gupta, Ganesh Chandra Deka,2011) HSC examination performance(A T M Shakil Ahamed1, Navid Tanzeem Mahmood, Rashedur M Rahman,2016) by considering different factors. This research focuses on predicting student attention in the class room and classifies them into attentive, partially attentive and not attentive. This paper is organized as follows: The section (2) reviews related work followed by the section (3) that presents the input feature variables identified and section (4) analysis the neural networks classifiers to find out better solution for predicting student attention during lectures. Finally, in section (5) the conclusion and the future work are presented.

2. RELATED WORK

Many researchers have developed ANN models in educational data mining. The table below shows the different ANN techniques and the results.

S. No	Author name	Technique	Result
1	(Md. Fahim Sikder,Md. Jamal Uddin and Sajal Halder,2016).	Back Propagation Neural Network	Predicted student yearly performance in the form of cumulative grade point with the accuracy of more than 99%.
2	(Nidhiarora, Jitinderkumar R .Saini, 2013).	Fuzzy probabilistic neural network to predict student academic performance	They have used input parameters such as merit, family background, attention in class etc.. The results revealed that FPNN takes less time to be trained and achieved the accuracy of 98.56%.
3	(KritiPurswani,PankajDalal, 2013).	Feed Forward Neural Network with fuzzy c means clustering.	Achieved approximately 81.8182 % accuracy to predict software faults.
4	(Qing Yang, JingranGuo, Dongxu Zhang, Chang Liu, 2011).	Probabilistic Neural Network with fuzzy c means clustering.	The results showed that PNN approach can make classification more stable and have higher precision and it reduces the iterative amount.
5	(Han-XiongLi , ZhiLiu, 2008).	PNN with fuzzy logic to handle complex stochastic uncertainties.	They showed that fuzzy logic is used along with PNN to handle vagueness, randomness, and time-varying uncertainty existing in the process
6	(Kai Goebel , Bill Wood, 1994).	fuzzy logic with neural networks and probabilistic neural network to monitor and diagnose tool wear in unattended milling machines	It is proved that fuzzy-neural system classifies with high reliability due to its redundant structure and efficiency of the preclustering.
7	(A T M Shakil Ahamed, Navid Tanzeem Mahmood, and Rashedur M Rahman, 2016).	Random forest, ANN and Naïve Bayes to predict academic performance of HSC examination using Socioeconomic, Psychological and Academic Factors.	They have concluded that ANN model achieved highest accuracy compared with other two models.
8	(Stamos T. Karamouzis, Andreas Vrettos,2008)	ANN model for predicting community college students graduation rate with the application of sensitivity analysis	The results were shown that the disability services, the need for support services, and the student's age at the

		on the ANN parameters in order to identify the factors that contribute more to a successful/unsuceessful graduation outcome.	time of admission to the college were the three factors identified as a key factors for their prediction.
9	(Hoang Tieu Binh ,Bui The Duy, 2017).	Artificial Neural Network with 3 layer perceptron to predict student academic performance based on learning style.	They achieved 80.63 % of accuracy rate using multilayer perceptron.
10	(Pauziah Mohd Arsad, Norlida Buniyamin, and Jamalul-lail Ab Manan, 2013).	Neural network model for predicting academic performance of engineering students.	The outcomes showed that fundamental subjects at semester one and three have strong influence in the final CGPA upon graduation.
11	(Malaya Dutta Borah, Rajni Jindal ,Daya Gupta and Ganesh Chandra Deka.,2011).	Application of knowledge based decision technique to Predict student enrollment decision	ANN algorithm places the highest accuracy rate in terms of correctly classified tuples for both Training & Testing phase
12	(Shaymaa E. Sorour, Tsunenori Mine and Kazumasa Goda and Sachio Hirokawa, 2014).	ANN and Latent semantic Analysis for predicting student's grade based on their free style comments.	The results were shown that the accuracy of student's predicted grade in ANN and similarity measuring method were 82.6% and 78.5%, respectively.
13	(Julian Webber, Norisato Suga, Abolfazl Mehbodni ya, Kazuto Yano, Tomoaki Kumagai,2018)	Study on Fading Prediction for Automated Guided Vehicle Using Probabilistic Neural Network.	A probabilistic neural network (PNN) estimates the most likely signal by performing pattern matching between a stored and current fading signal window.
14	(Adeli H. , Panakkat A, 2009).	A probabilistic neural network for earthquake magnitude prediction.	The PNN model constructed in this research satisfies good prediction accuracies for earthquakes of magnitude between 4.0 and 5.5 .
15	(Purwono Prasetyawan, Imam Ahmad and Rohmat Indra Borman,2018).	Classification of the Period Undergraduate Study Using Back-propagation Neural Network	The results shown that BPNN algorithm is suitable for the classification of undergraduate study periods with accuracy rates above 85%.

3. ATTRIBUTES IDENTIFIED FOR PREDICTION

The input parameters identified are divided into two parts such as factors from students and teachers perspective. The factors are described below in table I. Student details such as Family background, last education details, interest and hobbies are considered as primary parameters. Teacher's contributing factors are considered as secondary parameters as teachers are playing crucial role to make student attentive in the class room.

Table I Input feature variables identified for predicting student attention

Student Factors	Teacher's Contributing Factors
• Details of last education 1. Grade in last exam 2. Medium of study 3. Attendance 4. Board of study 5. Brach studied 6. Time before admission • Family Background 1. Family income 2. Parent's education 3. Parent's status • Interest 1. Interest in learning 2. Interest in course 3. Interested subjects • Hobbies 1. Playing games 2. Surfing web 3. Reading books	• Problem solving skills • Teaching techniques • Self-directed learning skills • Collaborative skills • Communication • Information handling • Usage of ICT

4. NEURAL NETWORK ARCHITECTURE

Artificial Neural networks has been successfully implemented in the fields such as medical sciences, engineering, business and education etc.. Neural networks learn by examples. They can therefore be trained with known examples of a problem to acquire knowledge about it. Then the network can be used to solve 'unknown' or 'untrained' instances of the problem. The three fundamental classes of NN architectures are single layer feed forward architecture, multilayer feed forward architecture, and recurrent networks architecture. The learning mechanisms of NNs are broadly classified as supervised, unsupervised, and reinforced learning methods. Neural network, Fuzzy system and Evolutionary algorithms have been individually and integratedly applied to solve various problems in the real world (S. Rajasekaran and G.A. Vijayalakshmi pai, 2017). Transfer function in neural networks is classified into three categories: Linear, threshold and sigmoid. Linear activation function generates the network into a single input-output linear product, regardless of the structure of the network. For threshold units, the output is set depending on the threshold value. For sigmoid units, the output varies continuously but not linearly as the input changes(S. Rajasekaran and G.A. Vijayalakshmi pai,2017). The following section analyses the BPNN and PNN classifiers to find out the better solution for the present research.

4.1 Back Propagation Neural Network (BPNN)

Academic researchers have developed BPNN in different areas of education such as student yearly performance (Md. Fahim Sikder , Md. Jamal Uddin and Sajal Halder,2016), period undergraduate study (Purwono Prasetyawan, Imam Ahmad, Rohmat Indra Borman, Ardiansyah, Yogi Aziz Pahlevi, Dwi Ely Kurniawan,2018).etc.. Back Propagation algorithm is a supervised learning method for training multilayer feed forward neural network. It is built on high mathematical foundation and has very good application potential. The Back Propagation algorithm uses a technique called the delta rule or gradient descent to looks for the minimum value of the error function in weight space. The weights that minimize the error function is then considered to be a solution to the learning problem (S. Rajasekaran and G.A. Vijayalakshmi pai, 2017).

4.1.1 Pseudo code for Back-propagation: (Jason Brownlee, 2015)

Input: ProblemSize, InputPatterns, ,
Output: Network
Network < - ConstructNetworkLayers()
Network $_{weights}$ <- InitializeWeights(Network, ProblemSize)
For (i=1 to iterations $_{max}$)
Pattern $_i$ <- SelectInputPattern(InputPatterns)
output $_i$ <- ForwardPropagate(pattern, Network)
BackwardPropagateError(pattern $_i$, output $_i$, Network)
UpdateWeights(pattern $_i$, output $_i$, Network, learn $_{rate}$)
End
Return (Network)

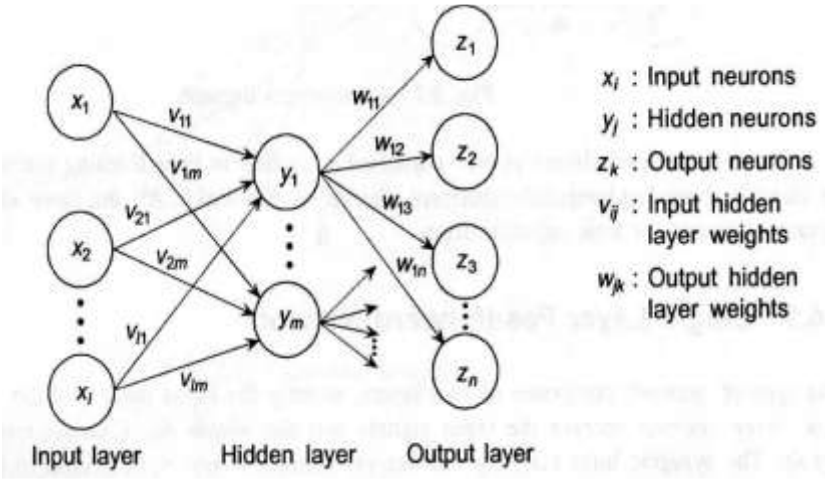

Fig. 1: A multilayer feedforward neural network (S.Rajasekaran, G.A. Vijayalakshmi pai, 2017)

4.2 Probabilistic Neural Network (PNN)

Probabilistic Neural Network is one of the feed forward Neural Network introduced by D.F. Specht in the early 1990's used in classification problems (Specht, D. F, 1990). It is organized as a multilayer feed-forward network with four layers. The first layer is a input layer with the set of measurements. The second layer is a Gaussian function made by the given set of data points. The third layer is a output from the second layer formed by performing average operation for each class. The fourth layer is formed by selecting the largest value which belongs to one of the classification. (Shahsavarani, S, 2012).The main advantage of PNN is that it has fast training process compared with Back Propagation networks. It is best suitable for real time applications as the training samples can be added or removed without extensive retraining. The PNN is a type of Radial Basis Function network which uses the radial basis function as a activation function in the hidden layer to make a decision (Tran T.P., Cao L., Tran D. and Nguyen C.D, 2009). The PNN is based on the Bayes classification rule with the implementation of Parzen nonparametric probability density function (PDF) estimation. (Parzen E, 1962).

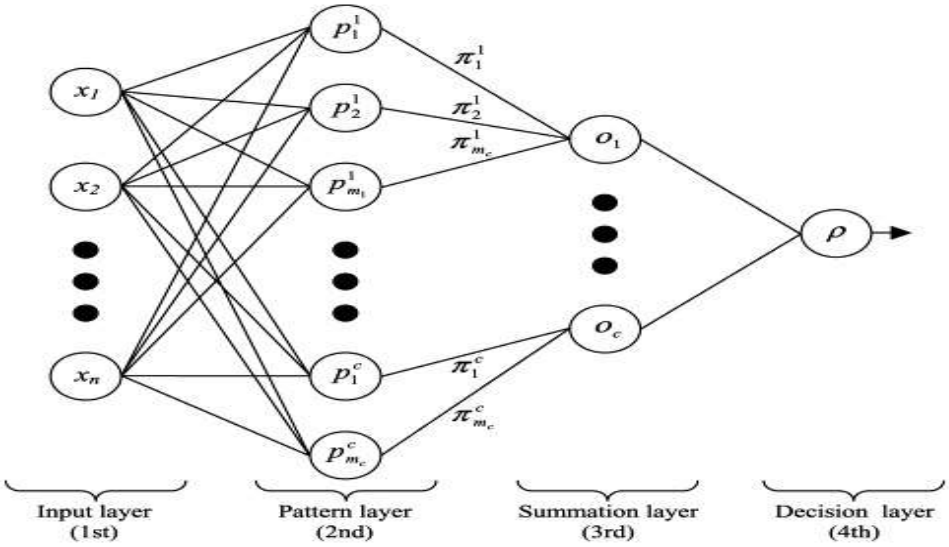

Fig. 2: Probabilistic Neural Network (Specht, D. F, 1990)

4.2.1 Working principle of PNN

The PNN creates a set of multivariate probability densities from the training vectors presented to the network. After receiving the input the each node in the pattern layer computes the output as:

$$\pi_i^c = \frac{1}{(2\pi)^{n/2}\sigma^n}\exp\left[-\frac{(x-x_{ij})^T(x-x_{ij})}{2\sigma^2}\right]$$

where d is the number of features of the input instance x, \sigma is the smoothing parameter, and x_{ij} is a training instance corresponding to category c. The summation layer neurons compute the maximum likelihood of pattern x being classified into c by summarizing and averaging the output of all neurons that belong to the same class.

$$p_i(x) = \frac{1}{(2\pi)^{n/2}\sigma^n}\frac{1}{N_i}\sum_{j=1}^{N_i}\exp\left[-\frac{(x-x_{ij})^T(x-x_{ij})}{2\sigma^2}\right]$$

where N_i denotes the total number of samples in class c. If the priori probabilities for each class are the same, and the losses associated with the incorrect decision for each class are the same, the decision layer unit classifies the pattern x in accordance with the Bayes's decision rule based on the output of all the summation layer neurons

$$C(x) = \text{argmax } \{p_i(x)\},\ i = 1,2,...,c$$

where C(x) denotes the estimated class of the pattern x and m is the total number of classes in the training samples. If the priori probabilities for each class are not the same, and the losses associated with the incorrect decision for each class are different, the output of all the summation layer neurons will be

$$C(x) = \text{argmax } \{p_i(x)\ cost_i(x)\ apro_i(x)\},\ i = 1,2,...,c$$

where $cost_i(x)$ is the cost associated with misclassifying the input vector and $apro_i(x)$ is the prior probability of occurrence of patterns in class c(Shahsavarani, S, 2012).

4.3 Comparison of BPNN and PNN for predicting attention of student in the class room

When we compare the accuracy of PNN with BPNN the performance difference increases with smaller datasets (Specht D.F., Shapiro P.D, 1991) and the training process of PNN is faster than BPNN. When the size of the training data set increases, PNN guarantees to converge to an optimal classifier than BPNN (Nidhiarora, Jitinderkumar R .Saini, 2013). Hence, PNN is well suitable for real time applications because as soon as one pattern representing each category is recognized, the new pattern can be generalized by the network. The main drawback of BPNN is that, it requires extensive retraining when we add or remove training samples (Specht D.F., Shapiro P.D, 1991). Hence, BPNN is not suitable for real time applications as the new samples will be added frequently.

5. CONCLUSION

It is found that PNN is preferred as it requires lesser time to train the network and the new data samples can be added without extensive retraining compared with BPNN. So far researchers have used BPNN and PNN techniques in education field to predict student academic performances (Nidhiarora, Jitinderkumar R .Saini, 2013), (Hoang Tieu Binh and Bui The Duy,2017), (Pauziah Mohd Arsad, Norlida Buniyamin, and Jamalul-lail Ab Manan,2013) by considering different input parameters. From this survey the Probabilistic Neural Network is been identified to classify student attentiveness into different categories as it requires less time to train the network and it has proven to be more time efficient than Back Propagation network. The various factors has been identified from students and teachers perspective and presented in this paper. In the future work an Artificial Neural Network based model will be delivered that could be used for an early prediction of assessing the student's attention in the class room. This research would serve as a strategy to understand the student's level of attention and how it varies across the lecture hours.

References:

[1] Nidhiarora, Jitinderkumar R .Saini, 2013. A Fuzzy Probabilistic Neural Network for Student's Academic Performance Prediction, International Journal of Innovative Research in Science, Engineering and Technology, Vol. 2, Issue 9.

[2] Priyanka Gaur, 2014.Neural Networks in Data Mining.International Journal of Electronics and Computer Science Engineering, ISSN- 2277-1956.

[3] Sanjesh Ghore. 2014. Data Mining used of Neural Networks Approach. IJISET - International Journal of Innovative Science, Engineering & Technology,Vol. 1 Issue 6.

[4] A T M Shakil Ahamed, Navid Tanzeem Mahmood, Rashedur M Rahman3, 2016. Prediction of HSC Examination Performance using Socioeconomic, Psychological and Academic Factors, 978-1-5090-2963-1/16, IEEE.

[5] Stamos T. Karamouzis, Andreas Vrettos.2008. Sensitivity Analysis of Neural Network Parameters for identifying the Factors for College Student Success, 978-0-7695-3507-4/08 IEEE.

[6] Hoang Tieu Binh and Bui The Duy, 2017. Predicting Students' performance based onLearning Style by using Artificial NeuralNetworks, 9th International Conference on Knowledge and Systems Engineering(KSE) -5386-3576-6/17/2017 IEEE.

[7] Pauziah Mohd Arsad, Norlida Buniyamin, and Jamalul-lail Ab Manan. 2013. "A Neural Network Students' Performance Prediction Model (NNSPPM),IEEE International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA).

[8] Shaymaa E. Sorour, Tsunenori Mine, Kazumasa Goda, Sachio Hirokawa.2014. Predicting Students' grades Based on free style Comments Data by Artificial Neural Network, IEEE Frontiers in Education Conference (FIE) Proceedings. ISBN: 978-1-4799-3922-0.

[9] S. Rajasekaran and G.A. Vijayalakshmi pai, Second edition: 2017. Neural Networks,Fuzzy systems, and Evolutionary algorithms, Sythesis and applications, ISBN-978-81-203-5334-3.

[10] Specht D.F., Shapiro P.D. 1991.Generalization accuracy of probabilistic neural networks compared with back-propagation networks, Int Joint Conference on Neural Networks, pp. 887-892.

[11]Shahsavarani.S. 2012. Probabilistic Neural Network, http://cse-wiki.unl.edu/wiki/index.php/ Probabilistic_Neural_Network#Architecture.

[12] Specht, D. F.,1990. Probabilistic neural networks, Neural networks, 3(1),109-118.

[13] Yang, Q., Guo, J., Zhang, D., & Liu, C. 2011. Fault Diagnosis Based on Fuzzy C-means algorithm of the Optimal Number of Clusters and Probabilistic Neural Network, International Journal of Intelligent Engineering & Systems, Volume 4, No.2, 51-59.

[14] Malaya Dutta Borah, Rajni Jindal, Daya Gupta, Ganesh Chandra Deka. 2011. Application of knowledge based decision technique to Predict student enrollment decision, International Conference on Recent Trends in Information Systems, 978-1-4577-0792-6/11,IEEE.

[15]Tran T.P., Cao L., Tran D. and Nguyen C.D. 2009. Novel intrusion detection using probabilistic neural network and adaptive boosting, International J Comput. Sci. Inform. Secur., Vol. 6, pp. 83-91.

[16] Parzen, E.1962. On estimation of a probability density function and mode,Ann. Math.Statist., Vol. 33, No. 3, pp. 1065-1076.

[17] Deniz D. Z., and Ersan I. 2002. An academic decision-support system based on academic performance evaluation for student and program assessment, International Journal of Engineering Education, Vol. 18, No. 2, pp. 236–244.

[18]Julian Webber, Norisato Suga, Abolfazl Mehbodni ya, Kazuto Yano, Tomoaki Kumagai. 2018. Study on Fading Prediction for Automated Guided Vehicle Using Probabilistic Neural Network, Asia-Pacific Microwave Conference (APMC).

[19] Adeli H. and Panakkat A. 2009. A probabilistic neural network for earthquake magnitude prediction, Neural Networks, No. 22, pp. 1018-1024.

[20] Md. Fahim Sikder , Md. Jamal Uddin and Sajal Halder,2016. Predicting Students Yearly Performance using Neural Network: A Case Study of BSMRSTU, 5th International Conference on Informatics, Electronics and Vision (ICIEV), 978-1-5090-1269-5/16, IEEE.

[21] Han-XiongLi & ZhiLiu, 2008. A Probabilistic Neural-Fuzzy Learning System for Stochastic Modeling, IEEE transactions on fuzzy systems, vol. 16, no. 4, august 2008 1063-6706, IEEE.

[22] KritiPurswani, PankajDalal.2013. Software Fault Prediction Using Fuzzy C-Means Clustering and Feed Forward Neural Network, International Journal of Digital Application & Contemporary research, volume 2, Issue 1, ISSN: 2319-4863.

[23] Kai Goebel, Bill Wood.1994. Comparing a Neural-Fuzzy Scheme with a Probabilistic Neural Network for Applications to Monitoring and Diagnostics in Manufacturing Systems, Compilation copyright © 1994, AAAI Technical Report SS-94-04.

[24] Qing Yang1,2, Jingran Guo, Dongxu Zhang, Chang Liu.2011.Fault Diagnosis Based on Fuzzy C-means Algorithm of the Optimal Number ofClusters and Probabilistic Neural Network, International Journal of Intelligent Engineering and Systems, Vol.4, No.2.

[25] Purwono Prasetyawan, Imam Ahmad, Rohmat Indra Borman, Ardiansyah, Yogi Aziz Pahlevi, Dwi Ely Kurniawan. 2018. Classification of the Period Undergraduate Study Using Back-propagation Neural Network, 978-1-5386-8066-7/18, IEEE.

[26]Purwono Prasetyawan, Imam Ahmad and Rohmat Indra Borman. 2018. Classification of the Period Undergraduate Study Using Back-propagation Neural Network, 978-1-5386-8066-7/18, IEEE.

[27] Jason Brownlee, 2015.Clever Algorithms: Nature-Inspired Programming Recipes, © Copyright 2015.