Predicting Student Academic Performance using Fuzzy ARTMAP Network

Nidhi Arora\textsuperscript{1}, Jatinderkumar R. Saini\textsuperscript{2}

Assoc. Prof., Institute of Technology & Management Universe, Vadodara, Gujarat, India\textsuperscript{1}
Assoc. Prof., Narmada College of Computer Application, Bharuch, Gujarat, India\textsuperscript{2}

Abstract: Predicting student academic performance plays an important role in academics. Classifying students using conventional techniques cannot give the desired level of accuracy, while doing it with the use of soft computing techniques may prove to be beneficial. A student can be classified into one of the available categories based on his behavioral and qualitative features. The paper presents a Neural Network model fused with Fuzzy Logic to model academic profile of students. The model mimics teacher’s ability to deal with imprecise information representing student’s characteristics in linguistic form. The suggested model is developed in MATLAB which takes into consideration various features of students under study. The input to the model consists of data of students studying in any faculty. A combination of Fuzzy Logic ARTMAP Neural Network results into a model useful for management of educational institutes for improving the quality of education. A good prediction of student’s success is one way to be in the competition in education system. The use of Soft Computing methodology is justified for its real-time applicability in education system.

Keywords: Education, Student Academic Performance, Fuzzy ARTMAP Network

I. INTRODUCTION

Educational universities are becoming more and more focused on behavioural assessment of students. The effect can easily be seen on teaching and learning which leads to a significant amount of time spent for preparing and taking standardized tests. Any move away from standardized and non-personalized tests holds promise for increasing deep learning \[7\]. The ability of an institute to predict the students’ performance accurately lends a significant hand in its success. There are many reasons to support this argument. The teachers can identify weak students in the class who are likely to have low achievement during the term or in the examination. The early recognition of the students’ weaknesses, the teachers can warn the students in advance and may provide them additional support thereby increasing the quality of their teaching in the classroom. Over a period of years, researchers have proposed various methods for predicting the academic performance of students. Early prediction of the students’ academic performance offer multi-fold advantages to the institute \[2\].

This paper examines various socio-demographic and study environment variables which influence the performance of the student. These variables are the determinants in identifying in advance the successful and unsuccessful students. A number of models have been developed in past to identify dropouts in e-learning environment that may impact attrition. Jun \[9\] classified them into five constructs, i.e. factors: individual background, motivation, academic integration, social integration and technological support. A good student performance prediction tool may help the instructor to motivate the student better by providing early feedbacks. The model presents a fuzzy ARTMAP network to predict performance of students and accordingly classify them in good, average or poor category.

The rest of the paper is structured as follows: Section II summarizes related work pertaining to significance of predicting academic performance of students for an educational institute and for predicting academic performance of student with the application of variety of techniques. Section III gives a brief on the architecture of Fuzzy ARTMAP network and describes applicability of neural Fuzzy ARTMAP in real world. Section IV discusses experiments conducted and results thereof performed on the available data. In section V the paper is concluded by presenting possible future work.

II. RELATED WORK

Literature presents a wide variety of methods for predicting student’s performance with different approaches. Both types of approaches viz. Conventional statistical approaches and soft computing approaches are applied by researchers in the past to predict academic performance of students. Olani \[15\] applied standard multiple regression analysis in order to explain the degree to which the GPA of first year university students were predicted from their prior academic achievements and psychological behaviours. The results showed 17\% of variance in students’ university GPA scores. Osmanbegovic \[16\] collected data from the surveys conducted during the summer semester at the University of Tuzla,
the Faculty of Economics, academic year 2010-2011, among first year students and compared data mining techniques for predicting students' success on the data. Desmarais & Baker [8] proposed a model for Predicting student performance in an Intelligent Tutoring System. Cen [5] argued that an improved model for predicting student performance could save millions of hours of students' time and effort in learning Algebra. This time can be utilized by the students in other fields of study or doing things they enjoy. An accurate and reliable model in predicting student performance may replace the current standardized tests reducing the pressure, time, as well as effort on teaching and learning for examinations [7].

A number of theoretical models have been developed focussing on the ways to lock in the students in a course. Jun [9] did extensive study on dropouts in e-learning environment and identified the variables which have impact on attrition. He classified the students based on factors like individual background, motivation, academic integration, social integration and technological support. Woodman [21] found using the binary logistic regression that the most significant factors to whether students passed, failed or dropped out depends on the marks for the first assignment, the number of maths courses passed in the previous two years, the course level, the points the course is worth and the occupation group of the student especially for courses in the mathematics and computing faculty at the Open University in UK. Kotsiantis et al. [11] considered demography and assignment marks as the key factors in supervised machine learning algorithms including decision trees, artificial neural networks, naive Bayes classifier, instance-based learning, logistic regression and support vector machines to predict student’s performance at the Hellenic Open University. It was also observed that the prediction accuracy varied from 58.84% to 64.47% when using neural networks instead of using support vector machines only with the demographic variables. On the other hand, when remaining factors are considered besides demography, the naive Bayes classifier was found to be the most accurate algorithm for predicting students' performance.

Minaei-Bidgoli et al. [13] classified students using genetic algorithms to predict their final grade. Using regression method, Kotsiantis & Pintelas [12] predicted whether a student belongs to category of passing or failing students. Vandamme & Meskens [20] predicted a student’s academic success into one of the three classes i.e. low, medium or high risk using decision trees and neural networks techniques. Al-Radaideh et al. [1] applied a decision tree model to predict the final grade of students who studied in C++ course in Yarmouk University. Romero et al. [17] compared different methods of data mining in order to predict final assessment based on the data obtained from the system of e-learning. Thai-Nghe et al. [19] presented a model for improving student’s academic performance prediction by dealing with class imbalance.

Bekele & Menzel [3], Kotsiantis & Pintelas [12] and Minaei-Bidgoli et al. [13] proposed using Bayesian networks, decision trees, and other classification techniques to predict the student results. Romero et al. [17] compared different techniques like neural networks and decision trees to classify students based on their Moodle usage data and the internal marks obtained in their respective courses. Cortez & Silva [6] collected information on past performance of students as well as their socio-economic attributes and worked on it to predict their grade. It was found that the tree based algorithms outperforms the methods like Neural Networks and SVM. Nghe, Janecek & Haddawy [14] compared the two methods viz. decision tree and Bayesian networks in predicting the academic performance of undergraduate and postgraduate students at two different academic institutes. Schmitt et al. [18] conducted a study to determine the validity of non-cognitive and cognitive predictors of the performance of college students at the end of their 4th year in college. They concluded that both the biodata and situational judgment measures could be useful supplements to cognitive indexes of student potential in college admissions. Karamouzis et al. [10] considered students from higher education and developed a three-layered perceptron for predicting student graduation outcomes using the back propagation.

III. FUZZY ARTMAP NETWORK

Fuzzy ARTMAP neural network, first proposed by Carpenter et al. [4], is a supervised clustering algorithm that operates on vectors with analog or binary valued elements. When fuzzy ARTMAP is used on a learning problem, it is trained to the point that it correctly classifies all training data. This feature may often result into fuzzy ARTMAP to over-fit some data sets, especially those in which the underlying pattern has to overlap. One solution to avoid the problem of over-fitting is to allow for some error during training.

The network, as shown in figure 1, is composed of two fuzzy ART modules [4], ARTa and ARTb, where ARTa (active input categories) processes the input vector and ARTb (active output categories) handles the desired output vector. The associative memory module, Fab, also called inter-ART or map Field, connects ARTa and ARTb. A match tracking process that increases the ART vigilance parameter achieves this by the minimum amount needed to correct a predictive error. Learning of fuzzy ARTMAP takes place step by step: normalization of the input/output vectors, code complement execution, recognition, comparison, search and learning.
During training, input signal consisting of independent variables are fed as input to ART\textsubscript{a} and output signal comprising dependent variables is provided as input to ART\textsubscript{b}. In recall phase, inputs are supplied only to ART\textsubscript{a} and the template chosen at ART\textsubscript{b} will serve as the predicted output. The purpose of MAP field is to ensure maximum code compression at ART\textsubscript{a} templates for generating minimum predictive error at ART\textsubscript{b} templates. The match tracking mechanism allows the network to increase vigilance parameter \( \rho \) of ART\textsubscript{a} module correcting the error in ART\textsubscript{b} module when the network executes a wrong prognostic. This ensures maximizing the generalization and minimizing the error. The ART\textsubscript{b} module begins the search until a correct prognostic or the creation of a new category for the current input is found. The inputs to ART\textsubscript{a} and ART\textsubscript{b} are in complement code form; for ART\textsubscript{a} module when the network ....

\( x^{a} = \{ x_{1}^{a}, ..., x_{2M_{a}}^{a} \} \)
\( y^{a} = \{ y_{1}^{a}, ..., y_{N_{a}}^{a} \} \)
\( w_{j}^{a} = \{ w_{j1}^{a}, ..., w_{j2M_{a}}^{a} \} \)
\( x^{b} = \{ x_{1}^{b}, ..., x_{2M_{b}}^{b} \} \)
\( y^{b} = \{ y_{1}^{b}, ..., y_{N_{b}}^{b} \} \)
\( w_{j}^{b} = \{ w_{j1}^{b}, ..., w_{j2M_{b}}^{b} \} \)
\( \rho \)

The map field \( F_{ab} \) is activated whenever one of the categories ART\textsubscript{a} or ART\textsubscript{b} is active. If node J of \( F_{1}^{a} \) is chosen, then its weight \( w_{j}^{ab} \) activates \( F_{ab} \). If node K of \( F_{1}^{b} \) is chosen, then the node K in \( F_{ab} \) is activated by 1-to-1 pathways between \( F_{1}^{b} \) and \( F_{ab} \). If both ART\textsubscript{a} and ART\textsubscript{b} are activated, then \( F_{ab} \) becomes active only if ART\textsubscript{a} predicts the same category as ART\textsubscript{b} via the weight \( w_{j}^{ab} \). At the beginning of each input presentation to the ART\textsubscript{a}, vigilance parameter \( \rho_{a} \) equals a baseline vigilance \( \rho_{a0} \). The map field vigilance parameter is \( \rho_{ab} \). If
\[
|x^{ab}| < \rho_{ab} |y^{b}|
\]
Then \( \rho_{a} \) is increased until it is slightly larger than \( |A \wedge W_{j}^{b}| |A|^{1} \), where A is the input to \( F_{1}^{a} \) in complement coding form and
\[
|x^{a}| = |A \wedge W_{j}^{b}| < \rho_{a} |A|
\]
Where, J is the index of the active \( F_{j}^{a} \) node. When this occurs, ART\textsubscript{a} search leads either to activation of another \( F_{j}^{a} \) node J with:
\[
|x^{a}| = |A \wedge W_{j}^{b}| \geq \rho_{a} |A|
\]
and
\[
|x^{b}| = |y^{b} \wedge W_{j}^{ab}| \geq \rho_{b} |y^{b}|
\]

Table 1. Fuzzy ARTMAP Architecture

![Fuzzy ARTMAP Architecture Diagram](image-url)
Predicting Student Academic Performance using Fuzzy ARTMAP Network

or, if no such node exists, to the shutdown of $F^n$ for the remainder of the input presentation. Learning rules determine how the map field weights $w^{ab}_{jk}$ change through time. This can be done as follows: Weights $w^{ab}_{jk}$ in $F^n \rightarrow F^n$ paths initially satisfy:

$$w^{ab}_{jk}(0) = 1$$  \hspace{1cm} (5)

During resonance with the ART$_a$ category $J$ active, $w^{ab}_{jk}$ approaches the map field vector $x^{ab}$. With fast learning, once $J$ learns to predict the ART$_b$ category $K$, that association is permanent, i.e., $w^{ab}_{jk} = 1$ for all time.

IV. EXPERIMENTS AND DISCUSSION

This study focuses on a technique that considers student data gathered from two different streams. The reason of taking data from different streams is to check the generalized applicability of the model in any stream. Table 1 shows the factors used in this work for predicting academic performance of the students.

Table 1. Factors for Predicting Academic Performance of Student

<table>
<thead>
<tr>
<th></th>
<th>Merit in the last exam</th>
<th>Average merit throughout study</th>
<th>Interest in the course</th>
<th>Interest in the curriculum</th>
<th>Interest in learning</th>
<th>Belief in hard work</th>
<th>Interest in learning from mistakes</th>
<th>Study with reference books</th>
<th>Study with revision</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Merit in the last exam</td>
<td>Average merit throughout study</td>
<td>Interest in the course</td>
<td>Interest in the curriculum</td>
<td>Interest in learning</td>
<td>Belief in hard work</td>
<td>Interest in learning from mistakes</td>
<td>Study with reference books</td>
<td>Study with revision</td>
</tr>
</tbody>
</table>

A fuzzy ARTMAP Network has been proposed to model the academic performance of students on a training sample of size 600 on an 18-dimensional input. A test set and a validation set each of size 100 datasets is used to test the network and to avoid the chances of over-fitting respectively. Root mean square error (RMSE) value in testing the network’s performance is found to be decreasing with increase in number of training epochs for first few runs but starts increasing later on. This implies that the training epochs with least error must be selected. The results of optimized network with 60 epochs resulted in best predictive power of fuzzy ARTMAP on the unseen test set. The model is found to achieve correct classification of 88.26% with RMSE of 0.1174. Table 2 represents the experimental results of network’s performance on the training data. The smoothing parameter is considered to be 0.1 as it was found to minimize the classification error.

Table 2. Performance Accuracy of Fuzzy ARTMAP on Test Set

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Epochs</th>
<th>Smoothing Parameter</th>
<th>Root mean Square Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>0.01</td>
<td>0.6019</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>0.01</td>
<td>0.5429</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>0.01</td>
<td>0.5264</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>0.01</td>
<td>0.4293</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>0.01</td>
<td>0.3705</td>
</tr>
<tr>
<td>6</td>
<td>60</td>
<td>0.01</td>
<td>0.3449</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>0.1</td>
<td>0.5654</td>
</tr>
<tr>
<td>8</td>
<td>20</td>
<td>0.1</td>
<td>0.4012</td>
</tr>
<tr>
<td>9</td>
<td>30</td>
<td>0.1</td>
<td>0.3999</td>
</tr>
<tr>
<td>10</td>
<td>40</td>
<td>0.1</td>
<td>0.3132</td>
</tr>
<tr>
<td>11</td>
<td>50</td>
<td>0.1</td>
<td>0.2879</td>
</tr>
<tr>
<td>12</td>
<td>60</td>
<td>0.1</td>
<td>0.1174</td>
</tr>
<tr>
<td>13</td>
<td>70</td>
<td>0.1</td>
<td>0.2099</td>
</tr>
<tr>
<td>14</td>
<td>80</td>
<td>0.1</td>
<td>0.2293</td>
</tr>
<tr>
<td>15</td>
<td>90</td>
<td>0.1</td>
<td>0.3614</td>
</tr>
<tr>
<td>16</td>
<td>100</td>
<td>0.1</td>
<td>0.5016</td>
</tr>
</tbody>
</table>

The training error value 1.3257 suggests that fuzzy ARTMAP network is capable of clustering in fuzzy environment. The results of experiments done on the trained network show that with 18 input neurons the Fuzzy ARTMAP network can predict the students in near correct categories.

V. CONCLUSION

A model has been created for predicting academic performance of students using Fuzzy ARTMAP Network. The results show that the model can be useful in educational institutes in predicting the categories of students in the class.
before they actually appear in the examination. The two fold advantages of prediction include providing an approach to the teachers to prepare their lectures based on the level of students in the class and motivating students by giving them timely advice to focus on the studies for timely improvement in their result. Conventional model for predicting academic performance of student are limited in their performance accuracy and generalization, but the suggested methodology can deal with uncertainty in its determinants. The model presented proves to have prospective application in educational system. Future work in the same area may include exploring additional determinants which affect academic performance of a student in an institute. Such determinant factors can be analysed for their significance in contribution in successful prediction.

REFERENCES
7] Feng, M., Heffernan, N. & Koedinger, K., (2009), Addressing the assessment challenge with an online system that tutors as it assesses, User Modelling and User-Adapted Interaction, 19, 243-266.
Predicting Student Academic Performance using Fuzzy ARTMAP Network
