International Journal of Engineering Applied Sciences and Technology, 2016
Vol. 1, Issue 8, ISSN No. 2455-2143, Pages 19-22
Published Online June - July 2016 in IJEAST (http://www.ijeast.com)

ANFIS APPROACH WITH GENETIC FEATURE SELECTION FOR PREDICTION OF STUDENTS' ACADEMIC PERFORMANCE IN DISTANCE EDUCATION ENVIRONMENT

Majd Latah
Computer Engineering Department
Ege University, Izmir, Turkey

Abstract— **Recently the number of distance education platforms has been increased significantly. These platforms provide isolation between the student and teacher. Thus, there is a need for predicting the students who are possible to fail in a specific course and take the precautions like starting face-to-face on demand lectures for individual cases. Both of artificial intelligence and data mining techniques can be used perfectly for this task. In this paper a neuro-fuzzy inference approach has been used for prediction of academic performance of students in distance education system. The proposed system uses Takagi Sugeno Kang fuzzy inference system for the generation of the fuzzy rules. In addition the genetic algorithm used as feature selection method. The experimental results have shown that the proposed system in this paper can over perform both of neuro-fuzzy and conventional neural network approaches.**

Keywords— **ANFIS, Genetic algorithms, Neuro Fuzzy systems**

I. INTRODUCTION

The e-learning systems represent a promising education system for people need for a certain degree without the obligations of going to the education institutes frequently. In this environment the interaction between the teacher and students is less compared with conventional education systems. For that there is a need for the teacher to predict the student's performance before failing in the exams. It can be used for building more adaptive systems that help the students to fill the gap in certain subjects because unlike traditional education systems some of student came to e-learning system after many years of leaving the high school. This prediction process also helps in increasing the quality of the output of the educational process by taking these results to raise the prerequisites to the next course related to the current course or increasing admission requirements for the whole program that failed students included in. Because failed students or the potential dropout students are using the web resources for the university platform without any benefit for them or for the educational process. On the other hand, new emerging technologies like big data have introduced a new types of implicit unstructured data. Combination of these data and artificial intelligence approaches can be an effective tool for more accurate predictions in the new e-learning systems.

II. RELATED WORKS

In [1] a combination of three classifiers has been used with a voting online learning approach. These classifiers are: (i) Winnow, (ii) 1-Nearest Neighbour and (iii) Naive Bayes. Winnow , a linear online algorithm which is very similar to the perceptron algorithm. The main difference is that Perceptron additively updates the weight, and Winnow multiplicatively updates the weight. In addition, they use different norms. Normalized Euclidean distance has been used as distance metric in 1NN algorithm. The ensemble predicts the pattern classification based on majority vote of the predictions and continually modifies its hypothesis in online learning manner. The accuracy was better than individual using of the previous algorithms. In the same context authors in [2] used a genetic algorithms-based approach for building more accurate decision trees. Binary decision trees used to predict the fail/success patterns of the students. The GA in every new generation selects a certain number of decision trees according to the fitness value where the crossover operation can help in exchanging parts between decision trees and mutation operation can modify a specific attribute. The fitness value include factors such as correct classification and size factor. Increasing the size factor also increases the accuracy of the model but it opens the door for over fitting problem. The results show more accurate decision trees than the traditional ones. Whereas the authors [3] have proposed a gene-fuzzy model for the same problem. Two fuzzy logic models were implemented classical fuzzy and expert fuzzy. The researches have used the genetic algorithm approach to optimize the

International Journal of Engineering Applied Sciences and Technology, 2016
Vol. 1, Issue 8, ISSN No. 2455-2143, Pages 19-22
Published Online June - July 2016 in IJEAST (http://www.ijeast.com)

interval between membership function in these models. The error ratio had been used as fitness function. Moreover, a single point crossover type had been used and the accuracy of the classic fuzzy model had been increased from 72% to 82.5 %. In [4] also a feed forward neural network approach has been proposed for the same topic and it was more efficient than linear regression method.

III. NEURO FUZZY SYSTEMS

These systems combine between the fuzzy logic and neural network approaches. Fuzzy logic (FL) was invented by Lofti Zadeh in 1965. Fuzzy logic represents multivalued logic which can handle with imprecise and vague information. Unlike boolean logic the variable can be partial true of partial false based on the degree of membership. These linguistic variables are used in fuzzy rules. There are two types of Fuzzy inference systems [5]: The first one is Mamdani fuzzy inference system which is proposed in 1975 by Ebrahim Mamdani. Mamdani's model consists of four phases: fuzzification of the input variables, rule evaluation, aggregation of the rule outputs, and finally defuzzification. The other one is Sugeno Fuzzy inference system which is proposed by Michio Sugeno in 1985. The difference in this method the output is constant or linear function of the input variable. In spite of fuzzy logic systems are easy to implement but they need experts to determine the membership functions and appropriate rules. But if we use a learning technique like neural networks it could lead to optimize the membership functions with least output error. FL systems are easy to explain whereas neural Networks are used as a black box; that's why the hyWbird approach lead to take the advantages of these two important methods and can bypass their weakness points. There are several neuro-fuzzy architectures like: (i) Adaptive Network based Fuzzy Inference System (ANFIS), (ii) Dynamic/Evolving Fuzzy Neural Network (DENFIS) and (iii) Hybrid Neural Fuzzy Inference System (HyFIS). [6]

IV. ADAPTIVE NEURO FUZZY INFERENCE SYSTEMS

(ANFIS) approach implements a Takagi Sugeno fuzzy inference system which has linear output function and the system has six layers as shown in Figure 1. The first one is for crisp input, the second one is the fuzzification layer and the parameters in this layer called premise or antecedent parameters, the third layer is the rule layer, fourth layer is the normalization layer and the fifth layer is defuzzification layer which includes consequent parameters and the last layer is output layer which includes a single summation neuron [5].

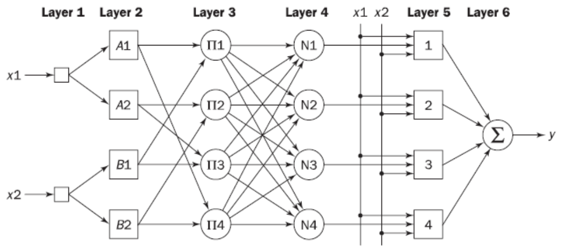

Fig 1. Adaptive Neuro-Fuzzy Inference System (ANFIS) [5]

The first layer is the input layer which only pass crisp signals to the next layer:

$$y_i^{(1)} = x_i^{(1)} \tag{1}$$

Whereas the fuzzification neurons in the second layer have a bell activation function as shown below:

$$y_i^{(2)} = \frac{1}{1+(\frac{x_i^{(2)}-a_i}{c_i})^{2bi}} \tag{2}$$

In the third layer the strenght of the Sugeno-type fuzzy rule is calculated by the operator product as the third output of neuron (i) in the third layer obtained by:

$$y_i^{(3)} = \prod_{j=1}^{k} x_{ji}^{(3)} \tag{3}$$

The 4th layer which is the responsible of calculating the strength of a normalized rule as follows:

$$y_i^{(4)} = \frac{x_{ii}^{(4)}}{\sum_{j=1}^{n} x_{ji}^{(4)}} = \frac{\mu_i}{\sum_{j=1}^{n}\mu_i} = \bar{\mu}_i \tag{4}$$

Where $x_{ji}^{(4)}$ is the input from neuron j located in Layer 3 to neuron i in Layer 4, and n is the total number of rule neurons.
The defuzzification operation in the fifth layer is calculated by:

$$\begin{aligned} y_i^{(5)} &= x_i^{(5)}[k_{i0} + k_{i1}x_1 + k_{i2}x_2] \\ &= \bar{\mu}_i^{(5)}[k_{i0} + k_{i1}x_1 + k_{i2}x_2] \end{aligned} \tag{5}$$

Where $x_i^{(5)}$ is the input and $y_i^{(5)}$ is the output of defuzzification neuron i in Layer 5, and $\boldsymbol{k_{i0}}$, $\boldsymbol{k_{i1}}$ and $\boldsymbol{k_{i2}}$ is a set of consequent parameters of rule i.

Layer 6 is a single summation neuron as shown below:

$$y_i^{(6)} = \sum_{i=1}^{n} x_i^{(6)} = \sum_{i=1}^{n} \bar{\mu}_i[k_{i0} + k_{i1}x_1 + k_{i2}x_2] \tag{6}$$

ANFIS uses back propagation learning algorithm to determine the antecedent parameters which are the input membership functions parameters and the least mean square method to determine the consequents parameters. The learning algorithm includes two steps in each iteration: (i) Refining the consequent parameters by treating all antecedent parameters as fixed parameters and applying the minimum squared method, (ii) Refining the antecedent parameters by treating all consequent parameters as fixed parameters and applying back propagation learning algorithm [6].

International Journal of Engineering Applied Sciences and Technology, 2016
Vol. 1, Issue 8, ISSN No. 2455-2143, Pages 19-22
Published Online June - July 2016 in IJEAST (http://www.ijeast.com)

V. GENETIC ALGORITHMS

Genetic Algorithms (GAs) are adaptive heuristic search algorithm based on the evolutionary ideas of natural selection and genetics. Genetic algorithms encode a potential solution to a specific problem on a chromosome data structure and try to improve this chromosome by the selection, crossover and mutation operations. The GA starts from a group of initial solutions called the initial population. Furthermore, a fitness function is used to evaluate the performance of the solutions. Each time two solutions are chosen from the population and a crossover operation is applied to produce two new solutions of the next generation. The new solutions will take the place of the old ones if they have a better fitness value. [7]

The steps of the algorithm can be summarized as follows [8]:

Initialize Population
Repeat
  Evaluation
  Reproduction
  Crossover
  Mutation
Until requirements are met

Moreover, genetic algorithms can be used as a feature selection method in order to reduce the dimensional of the used dataset.

VI. PROPOSED ANFIS SYSTEM

The architecture of the proposed system is shown in Figure 2. To reduce the dataset dimensions and achieve higher performance of this system, different methods can be used. In this paper genetic algorithm approach with 10-fold cross-validation has been used as wrapper feature selection method. As a wrapper method the features will be divided to subsets and each subset will be evaluated separately and take the subset that achieve the best prediction performance on that dataset. On the other hand, for the ANFIS system we used triangle membership functions and Centre Of Gravity (COG) as a defuzzification method. The type of used conjunction operator is MIN and disjunction operator is MAX.

The architecture of the proposed system is shown in Figure 2. To reduce the dataset dimensions and achieve higher performance of this system, different methods can be used. In this paper genetic algorithm approach with 10-fold cross-validation has been used as wrapper feature selection method. As a wrapper method the features will be divided to subsets and each subset will be evaluated separately and take the subset that achieve the best prediction performance on that dataset. On the other hand, for the ANFIS system we used triangle membership functions and Centre Of Gravity (COG) as a defuzzification method. The type of used conjunction operator is MIN and disjunction operator is MAX.

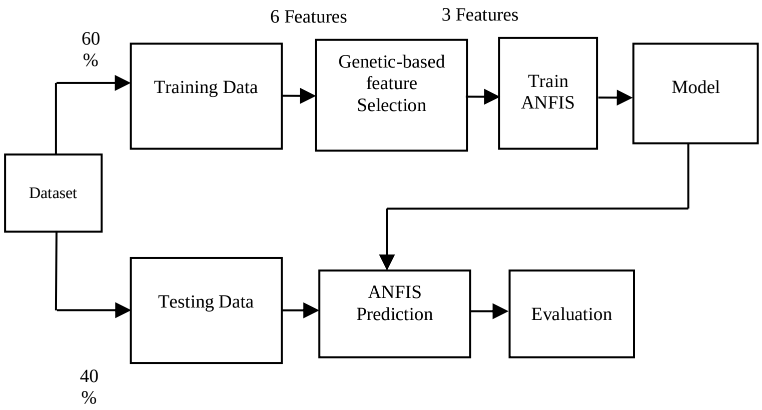

Fig 2. The architecture of the proposed system.

We used a unity–based normalization for the dataset by the formula:

$$X' = \frac{(X - X_{min})}{X_{max} - X_{min}} \qquad (7)$$

The structure of the dataset we have used includes 100 record with 6 features as shown in Table 2.

Table 2. The structure of the used dataset

Attribute	Description
Last long	Stood for period last a student logged on to the section of the system related to the artificial intelligence course.
Frequency	Stood for frequency at which a student logged on to the system.
Time Spent	Stood for the amount of time a student spent on the section of the system related to the artificial intelligence course.
Written exam results	Stood for the sum of midterm exam and quiz results for the course.
Resource downloading interval	Stood for the difference between the day of uploading the resources related to the course and the day of downloading the resources by the student.
Response to the lecturer questions	Stood for the student's response to the questions asked by the teacher during the online lecture.

VII. EXPERIMENTAL RESULTS

We used both of the (frbs) and (caret) libraries from the R programming language to implement the proposed system shown in Figure 2. The used methods from (frbs) library are described in Table 1.

Table 1. FRBS method's used in the proposed system

Method name	FRBS Model
ANFIS	Takagi Sugeno Kang
DENFIS	Clustering
HYFIS	Mamdani

The proposed system has achieved a less prediction error compared with other neuro-fuzzy approaches such as non-genetic ANFIS, DENFIS and HYFIS. Also the Genetic ANFIS system outperformed both of neural networks and Schwarz-Bayesian information criterion (SBC) approaches. According to Coefficient of determination (R^2) we see that the model we build explains 99.34 % of the variability of the response data around its mean.

Table 3. Performance evaluation criteria

Criteria	Calculation
Root mean squared error (RMSE)	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - A_i)^2}$
Coefficient of determination (R^2)	$1-(\frac{\sum_{i=1}^{n}(P_i - A_i)^2}{\sum_{i=1}^{n}A_i^2})$

Table 4. Comparing the performance of Genetic ANFIS and other systems

Method / Performance measure	RMSE	R^2 %
Genetic ANFIS	**5,19**	**99,34**
ANFIS	6,41	99,14
DENFIS	7,09	98,95
HYFIS	9,85	97,96
Neural Network	6,08	99,22
Schwarz-Bayesian information criterion	5,85	99,28

The proposed system has achieved least root mean squared error compared with other neuro fuzzy approaches such as non-genetic ANFIS, DENFIS and HYFIS. Also the Genetic ANFIS system outperformed the classical neural networks approach.

VIII. CONCLUSION

In this paper we proposed a neuro-fuzzy approach for prediction the students' academic performance in distance education systems. Genetic algorithms were used to reduce the input features from 6 to 3. The proposed system has shown better performance compared with non-genetic feature selection step. The previously mentioned also outperformed other neuro fuzzy systems, neural network and SBC approaches.

IX. REFERENCE

[1] S. Kotsiantis, K. Patriarcheas, M. Xenos, "A combinational incremental ensemble of classifiers as a technique for predicting students' performance in distance education," *Knowledge-Based Systems*, vol. 23, pp. 529-535, 2010.

[2] D. Kalles and C. Pierrakeas , "Using Genetic Algorithms and Decision Trees for a posteriori Analysis and Evaluation of Tutoring Practices based on Student Failure Models," *in 3rd IFIP Conference on Artificial Intelligence Applications and Innovations (AIAI)*, Athens, pp. 9-18, 2006.

[3] O. Yildiz, A. Bal, S. Gulsecen. "Improved fuzzy modelling to predict the academic performance of distance education students," *The international review of research in open and distributed learning*, vol. 14, no. 5, pp.144-165, 2013.

[4] I. Lykourentzou, I. Giannoukos, G. Mpardis, V. Nikolopoulos, & V. Loumos. "Early and dynamic student achievement prediction in e-learning courses using neural networks," *Journal of the american society for information science and technology (JASIST)*, vol. 60, no.2, pp.372–380, 2009.

[5] M. Negnevistsky. *Artificial Intelligence - A guide to intelligent systems*, 2 nd ed. Pearson. Uk, 2005.

[6] J. Vieira, F. Morgado Dias, "A. Mota. Neuro-Fuzzy Systems: A Survey," *WSEAS Transactions on Systems*, Issue 2, vol. 3, pp. 414-419, 2004.

[7] K. Sastry, D. Goldberg, G. Kendall, *Search Methodologies - Introductory Tutorials in Optimization and Decision Support Techniques*, Springer, 2005.

[8] D. Karaboğa, B. Akay, "A comparative study of Artificial Bee Colony algorithm," *Applied Mathematics and Computation*, 214 (2009), pp 108–132.