

A Review of Soft Computing Approaches in Student Academic Performance Evaluation

A.L.Gudyalakar

Gogte Institute of Technology
Belgaum, India

S.R.Mangalwede

Gogte Institute Technology
Belgaum,India

V.B. Deshmukh

Gogte Institute of Technology
Belgaum,India

D. H. Rao

Cambridge Institute of Technology,
Bengaluru, India

Abstract-The valuable factor in the student academics is student performance and their success in the academics. Evaluation of that academic performance plays major role in the educational sector. Student must be benefited by that evaluation process. We are well known with classical averaging method. This paper reviews the study of different types of machine learning algorithm used for student performance evaluation and prediction and this paper also consist of various types of classifiers that can be used for performance evaluation of the student. But all these classifiers are unique from one another. This uniqueness can be found, depending upon their strength and weakness, accuracy, time taken for training etc.

Keywords: Academic performance; Evaluation; different classifier.

I.INTRODUCTION

The academic performance and grades in a student's carrier are very important. To achieve good grades or improve performance, timely guidance to student may be needed; this guidance can be only given when teacher is aware of strengths and weaknesses of the student. To identify these qualities we can use either different prediction methods or detailed diagnostic/formative assessments of every student. This can be done considering different factors. Academic execution relies upon 3 factors (1) 1) intellectuals 2) capability to study 3) individual student personality. By considering these factors if a model predicts the student performance at the initial stage and if the student is found weak then some guide lines can be given to that student by the teacher to improve the performance in future, this can be very important application in the admission process, or placement activities or employee performance evaluation etc. (2). Evaluating the student performance plays a important role where students must be justified to their efforts. Hence various types of soft computing algorithms are available for performance evaluation. Student performance evaluation can be done in to two types by using 1.Rule based system and 2.Non Rule based system.we have discussed here the most common soft computing techniques employed in performance evaluation process such Neural Network (NN), Fuzzy inference system (FIS), Neuro fuzzy system(NFS), Genetic Algorithm(GA),Decision Tree(DT), Linear Discriminator (LD), K Nearest neighbour (KNN), Support Vector

Machine (SVM),Naïve Bayes (NB),Ensemble. The performance of these all classifiers are different from one another and this differentiation can be obtained by using different parameters such as confusion matrix, Root Mean Squared Error value (RMSE), Training time, Accuracy, etc. In all these techniques various parameters are considered to evaluate the student performance such as previous performance of the student, economical background, family background, location, all semester cumulative Grade Point Average(CGPA) etc.

A. Neural Network

Artificial Neural Network is procedure learning the human brain component. As shown in the figure (1) it involves three layers input layer, hidden layer, output layer. In this system different types of algorithm can be obtained to train the data set like Broyden-Fletcher-Goldfarb-Shanno(BFGS), Levenberg Marquardt Algorithm (LM), Resilient Back Propagation Algorithm (RBP).NN involves two stages 1) Training phase- Is directed to enable the system to do learning in view of the input 2) Online phase- where ANN will have the capacity to achieve nonlinear forecast of yield with new specimens nourished in to the information layer.

B. Fuzzy Inference System

FIS maps input to output using fuzzy set theory. It works with 3 stages 1) Fuzzification- Determining the crisp input values to the fuzzy membership function 2) Inference- rules can be generated to produce specified output. 3) Diffuzification- involves the converting process from fuzzy output to physical values (11) that is shown in the figure (2).

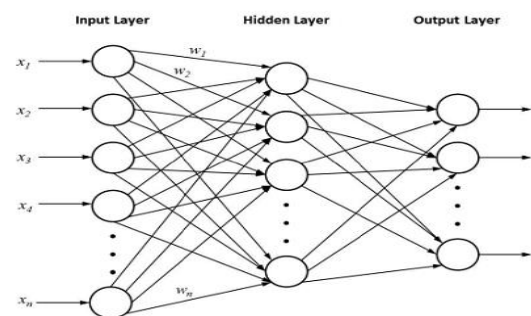


Fig. 1. Neural Network

There are two sorts of systems present here Mamdani and Sugeno, in Mamdani we can define the membership function to the output but in sugeno output is constant type. As per the need of application different kinds of membership function are used such as Gaussian, Triangular, and Trapezoidal, Sigmoidal etc.

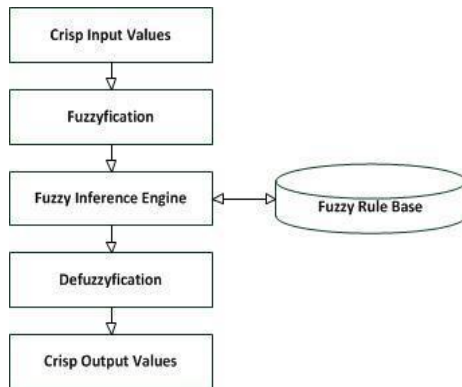


Fig. 2. Fuzzy Inference System

C. Neuro Fuzzy

NFZ is combination of both FIS and NN that is shown in the figure (3) and it uses the positive factors of the both the system that is explicit knowledge depiction of the FIS and learning power of the NN. It has following steps (1) Defining input and output variables by linguistic statements, (2) Deciding on the fuzzy partition of the input and output spaces, (3) Choosing the membership functions (MFs) for the input and output linguistic variables, (4) Deciding on the types fuzzy control rules, (5) Designing the inference mechanism and (6) Choosing a de-fuzzification procedure (4).

D. Genetic NN and Genetic Fuzzy algorithm

GA is natural genetics based search algorithms. It efficiently optimizes the parameters and then learns the rules of a rule based system. The parameter optimization has been the widely used application of genetic algorithms and this optimization is done with proper flow that is shown in the figure (4).

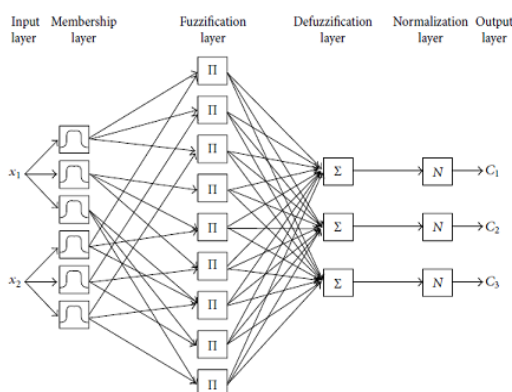


Fig.3 Basic block diagram of NFZ

The main challenge in the process of classifying the student performance is identifying and deciding the factors which will influence the student performance. The main

advantage of the Genetic algorithm (GA) is to locate the arrangement of elements that best predict academic achievement by keeping in mind the end goal to actualize a framework of educational information mining to precisely group the student's academic execution (1).

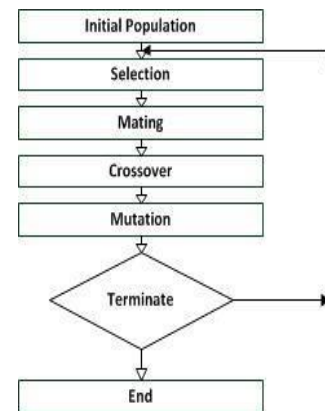


Fig.4. Basic flow of Genetic Algorithm

E. Decision Tree

A decision tree involves of a set of tree-structured decision tests operating in a divide-and-conquer technique. Each individual non-leaf node is related with a feature test also called a split; According to the feature test data will be divided in to the subset depending upon values. Each leaf node is associated with a label, which will be assigned to instances falling into this node. To classify the new data decision tree learning generates something like flowchart and at the each and every point values are verified that is shown in the figure (5). The information gain criterion is employed for split selection. Given a training set D, the entropy of D is defined as E

$$Ent(D) = -\sum_{y \in Y} P(y|D) \log P(y|D) \quad (1)$$

F. Linear discriminator

The linear discriminator algorithm flows with 2 steps initially all data instances are aligned on single dimensional surface and the instances are classified. linear classifier consists of a weight vector w and a bias b. Given an instance x, the predicted class label y is obtained by

$$y = \text{sign}(w^T x + b) \quad (2)$$

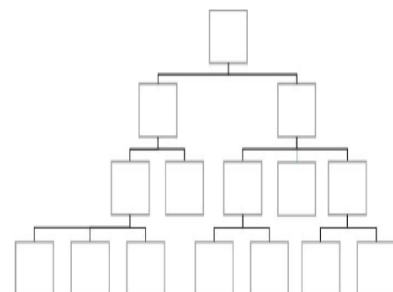


Fig. 5. Tree diagram

G. KNN

In KNN neighbors are chosen in group of objects in the manner of class and property of that object is known. Here

K defines the number of nearest neighbors that we want to consider for each object. It calculates distance between object and other objects present in the training data set and selects closest neighbor finally it classifies the object into particular class the complete flow of algorithm is shown in the Figure (6). It works with Euclidian distance algorithm.

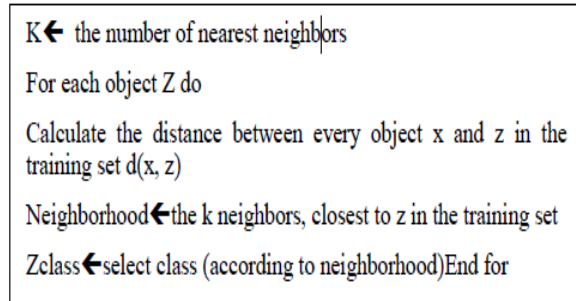


Fig. 6.Flow diagram of KNN Algorithm

H. Naïve Bayes

The main advantage of the Naïve Bayes classifier is most possible outputs are calculated based on the input(7). Consider that training samples are present (x_i, y_i) , where $x_i = (x_{i1}, x_{i2}, \dots)$ is an n-dimensional vector and y_i is the corresponding class. For a new sample x_{tst} , we wish to predict its class y_{tst} using Bayes' theorem:

$$y_{tst} = \arg \max_y P(y | x_{tst}) = \arg \max_y \frac{P(x_{tst} | y) P(y)}{P(x_{tst})} \quad (3)$$

By the following equation we come to know that, A Naive Bayes classifier makes a robust independence assumption on this probability distribution:

$$P(x | y) = \prod_{j=1}^n P(x_j | y) \quad (4)$$

It means each distinct components of x are conditionally independent given its label y . The assignment of classification now proceeds by estimating n single-dimensional distributions $(x_j | y)$.

I. Support vector machine

Support vector machine create the hyper plane to classify the objects. Individual subjects performance can be classified using support vector machine shown in figure (7). Consider a training dataset of feature-label pairs (x_i, y_i) with $i = 1, \dots, n$. The optimum separating hyper-plane is represented as

$$g(x) = \text{sign}(\sum_{i=1}^n y_i a_i K(x_i, x_j) + b) \quad (5)$$

Where $K(x_i, x_j)$ is the kernel function; is a Lagrange multiplier; and b is the offset of the hyper-plane from the origin. This is subject to constraints $0 \leq a_i \leq C$ and $\sum a_i y_i = 0$, where a_i is a Lagrange multiplier for every training point and C is the penalty. Some of the a_i is nonzero which are near to the support vector machine. The optimum hyper-plane can be represented as

$$y_i (w \cdot x_i + b) \geq 1 - \epsilon_i, \epsilon_i \geq 0 \quad (6)$$

Where w is the weight vector that determines the orientation of the hyper-plane in the feature space and ϵ_i is

the i th positive slack variable that measures the amount of violation from the constraints.

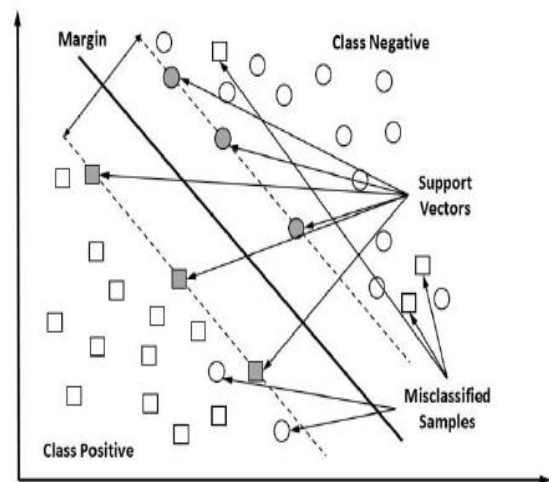


Fig.7.Support Vector Machine

J. Ensemble

Ensemble classifier is used to classify a new data and predict the data and its supervised learning where the output is labeled in the training data set. Normally, an ensemble is built in 2 steps, i.e., producing the base initiates, and then merging them as shown in Figure (8). To get a good ensemble, it is generally understood that the base learners must be as correct as possible, and as different as possible. It can be classified in to three categories, they are Random Forests (RF), Logit Boost (LB) and Cost-Sensitive Classifier (CSC).

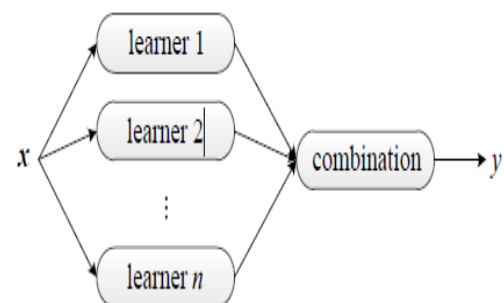


Fig. 8.Basic Architecture of Ensemble

II RELATED WORK

Student performance evaluation can be done using different soft computing techniques. Here we have studied few papers depicting applications of NN, FIS, GA, NFS, SVM, LD, DT, NB, KNN, Ensemble. for student performance prediction and evaluation.

In (1) authors predict academic performance of the student using artificial neural networks. Many factors influence the student performance like mental pressure, environment where they are living or studying, activities on social media, some youth problems, financial conditions, family support, involvement in other activities etc. Under such circumstances optimization technique of factors is employed. Few of these factors cannot be ignored. The

authors mention about 39 factors and it optimized to 8 using GA. Feed forward neural network with back propagation algorithm can be used as learning algorithm and also for evaluation system. Each factor can be sub classified into three categories such as Very Bad, Bad and Average.

In (2) authors predict student performance using Artificial Neural Networks (ANN). Here they have taken set of student's performances and evaluated the performances using BFGS and LM Algorithm. The later gives accurate results. This is used for performance prediction of B.Tech students. Some of the parameters considered for prediction are 10th results, 10+2 results, previous semester results, residence, gender, age and other competitive exam performances etc. RMSE values are calculated to compare the results of these algorithms.

In (3) authors classify the student academic performance using NFZ model. Different types of classifiers are analysed such as SVM, NB, NN, and DT etc. Data set consisting of certain parameters such as University entrance examination scores Subject 1, Subject 2, Subject 3, High school scores, student gender, Type of high school, location of high school, time between high school and graduation etc. Classification of students is done into three categories namely Very good Good and Poor.

Classifiers performances are compared using confusion matrix. True positive values are the main factor to compare the result.

In (4) authors predict the student academic performance using Sugeno-type NFS architecture. The membership functions generated by NFS using the training datapredict the student performance. Quiz (Q), Major (M), Midterm (MD), Final (F), Performance Appraisals (P), and Survey (S) scores , these are the input variables.

In (5) authors predict student performance in the distance education using genetic-fuzzy model. The use of fuzzy logic here is to build the genetic component and achieve classification using the membership functions. Optimization of the membership functions and fuzzy rules are done by the genetic algorithm. For each individual input variables membership function may be different and accuracy obtained by using that membership function is different. These membership functions intervals are selected by the Genetic algorithm.

In (6) authors classify the student characteristics using different types of algorithms such as KNN, SVM, NB, K means algorithm and their percentage of performance is compared. The data set consisting of student's sex, age, school, address, mother's education, mother's job, parents status, no of failures, guardian, family size , quality of relationship, study time, reason for choosing that school, travel time, school support, family support these all factors are considered in the student characterization.

In (7) author presents model which predict the Students Performance using KNN and NB algorithm. The data mining used to obtain the hidden knowledge about the student performance. There are three major steps to obtain this knowledge such as the Prior processing, Grouping and Post preparing. The data set consisting of the individual student ID, student name, male/female, birth place, branch,

academic year, location, telephone number, marriage status, fathers job, student status.

In (8) authors describe Classification using KNN. Classification using KNN can be implemented using certain distance algorithms such as Cosine, Correlation, Euclidean, City blocks. Integer numbers are used as classification data. The data is divided in to the training data and testing data. Classes obtained by the integer numbers are class1, class2, class3 using different types of distance algorithms and results are compared.

In (9) authors approaches prediction model which predicts instructor performance using different types of data mining algorithm such Discrimination Analysis (DA), CART, Decision tree, SVM and NN. By considering parameters such as how much student and instructor both are regular to the course, time utilization by the instructor, prior knowledge about the subject, preparation before coming to class, balance between application and theory, whether instructors encourage the student to ask the question and interactions, availability of course material, overall performance of the instructor, instructor grading was fair, given assignment/exams increases the ability to think more creatively such kind of instructor related 23 factors are considered to predict the instructor performance.

In (10) authors presented a model which classifies recorded heart sound using Ensemble algorithm. In the application of diagnosis of various types of heart disease recording of heart sound is needed. For the classification of the heart sound nested ensemble method is used and this method works with 3 major parts such as Pre-processing, Grouping and Evaluation.

In (11) author describes model which predicts the winning and losing chances of a participant in the election process using fuzzy inference system. Election results prediction has been done by using nine candidate related parameters such as Number of Years in Active Politics, Currently in Government, Performance In development; Inclination of Voters, Vote Bank, Internal War, Party / Independent are major issues. And using these parameters 91 rules are formed and finally winning chance and losing chances are obtained.

In (12) authors classify the student performance based on the student qualitative observation by using Neuro fuzzy student model. For e- learning application Neuro fuzzy can be used.. Strategies of student performance evaluation are unique from one educator to another educator. For e-learning evaluation the time taken to complete the test and scored marks in the test are used as the input variables and performance is classified as bad, good, very good, excellent.

In (13) author designed model which predicts performance of the electrical engineering students using neural network LM algorithm. Two models are designed here. In the first model circuit theory, fundamental electronics, signals & systems(S & S), communication theory, mathematics and all semester CGPA is considered. But in the second model only important parameters are considered such as the circuit theory, fundamental electronics, S&S, communication theory and results are obtained and compared with the first model. Model

comparison is done by considering some factors such as Network configuration, Learning rate, Momentum rate, Training Technique, Training goal, Training pattern, Correlation coefficient, Testing pattern and No. of variables.

In (14) authors presents model to predict the yearly performance of the student using back propagation algorithm. By considering student performance in academics and learning atmosphere of the student, future results of the student can be predicted. Some of the factors considered in the model are marks obtained in class test, class attendance, assignment, lab performance, number of study hours, family education, drug addiction, extra curricula activity, previous results etc.

In (15) authors predict the student academic performance using fuzzy clustering embedded regression. This prediction is done by using the input variable such as study-time, failure-count, support, study-aim, activities, etc. In the grouping process, the Euclidean distance between a given new instance and the centroid of individual cluster is calculated to determine which cluster the instance belongs to.

In (16) authors predict severity of diabetes using fuzzy cognitive maps. This is done by using patient's health condition parameters and level of those parameters. The different parameters are Pancreatic Dysfunction Positive, Complete Insulin Deficiency or its insufficient production, Insulin Resistance Obesity, Physical Inactivity Hereditary, Extreme Birth Weight, Drugs, Alcohol, Smoking Pregnancy, Stress Pollution, Poverty, Age, High Blood Pressure, and Metabolic Syndrome etc.

In (17) authors evaluate student academic performance using Fuzzy C-Means Clustering. It is consisting of fuzzy expert system which converts crisp output into fuzzy output set by using fuzzy c clustering system. Rule based evaluation uses first semester results and second semester results as input variable. Clustering means grouping the student data score into 5 classes as very high, high, average, low, very low.

In (18) authors predict short term traffic flow using neuro fuzzy. This system designed of 234 rules. The recorded traffic data has different attributes such as velocity, speed, vehicle classification and axle details etc. By considering these factors traffic flows are predicted.

In (19) authors presents model to predict student academic performance using decision tree and genetic algorithm. Performance prediction and evaluation depends upon the parameters that are considered for the prediction analysis. For better accuracy selected attributes must be relevant and noise free. By using certain parameters such as internal score, sessional score and admission score the student performance in the particular subject can be evaluated.

In (20) authors classify data using different kinds of data mining tools for diagnosis diabetes such as Multilayer Perceptron, Bayes Net, Naïve Byes, J48graft, Fuzzy Lattice Reasoning (FLR), JRip (JRipper), Fuzzy Inference System (FIS), Adaptive Neuro-Fuzzy Inference System (ANFIS). By considering health condition of patient's classification can be done to check out whether patient is having

diabetes. The data set of patient consisting of blood pressure, age, number of time pregnant, glucose tolerance in the body etc. Classifier performance will vary depending upon the software's that are used to implement it. Here they have used three software's 1) MATLAB 2) TANGRA 3) WEKA. Using all these software diagnosis for diabetes model is implemented and classifier performance is compared.

In (21) authors presents model to search an alternative design in an energy simulation tool. This application of energy usage in the prior stage of constructing the building can be classified using decision tree, KNN, Naïve Bayes. The factors considered for this process are Wall U-value, Wall Height, Roof U-value, Floor U-value, Floor Area, Number of Floors, Window U-value, South Window Area, North Window Area, East Window Area, West Window Area, Door U-value, and Door Area. Using above parameters energy usage prior to construction of the building can be obtained.

In (22) authors predict student performance achievement. This achievement is predicted by considering certain variables such as income, age, gender, state, status, etc. They have used three systems logistic regression model, NN and NFZ model and RMSE values of these classifiers are obtained and compared.

III CONCLUSION

Student performance evaluation is very important because using selected parameters student results can be analysed. The purpose of this work is to study 11 selected classification algorithms. Some of the systems are based on rule. These all classifiers can be used evaluate the student performance. Every single classifier has its own advantages and disadvantages and is unique.

During our survey we observed that different accuracy levels were obtained for different data sets while classifying the data using classifiers like SVM, KNN, DT etc. When the data set is large the DT was preferred as the processing speed and memory requirements are minimum even though the accuracy level is lesser than NB or KNN classifier in spite of their good accuracy levels. Among ANN and NF when RMSE values are compared, the latter was preferred. Among the rule based algorithms FL gives the best results with highest accuracy levels of classification. The only disadvantage here is if data set is large the time taken for rule base formation and processing is more. We observed that genetic neural and genetic fuzzy are used to optimization of the membership function and feature selection.

REFERENCES

- [1] Echegaray-Calderon, Omar Augusto, and Dennis Barrios-Aranibar. "Optimal selection of factors using Genetic Algorithms and Neural Networks for the prediction of students' academic performance." In Computational Intelligence (LA-CCI), 2015 Latin America Congress on, IEEE, 2015, pp. 1-6.
- [2] Pathak, Pooja, Neha Bansal, and Shivani Singh. "Mulyankan: A prediction for student's performance using Neural Network." In Computing for Sustainable Global Development (INDIACom), 2015 2nd International Conference on, IEEE, 2015, pp. 46-49.
- [3] Do, Quang Hung, and Jeng-Fung Chen. "A neuro-fuzzy approach in the classification of students' academic

- performance." Computational intelligence and neuroscience 2013 2013
- [4] Taylan, Osman, and Bahattin Karagözoğlu. "An adaptive neuro-fuzzy model for prediction of student's academic performance." *Computers & Industrial Engineering* 57, no. 3 2009, pp: 732-741.
 - [5] Yıldız, Osman, Abdullah Bal, Sevinç Gölseçen, and Fulya Damla Kentli. "A Genetic-Fuzzy Based Mathematical Model to Evaluate The Distance Education Students' Academic Performance." *Procedia-Social and Behavioral Sciences* 55, 2012, pp: 409-418.
 - [6] Troussas, Christos, Maria Virvou, and Spyridon Mesaretzidis. "Comparative analysis of algorithms for student characteristics classification using a methodological framework." In *Information, Intelligence, Systems and Applications (IISA)*, 2015 6th International Conference on, IEEE, 2015, pp. 1-5..
 - [7] Amra, Ihsan A. Abu, and Ashraf YA Maghari. "Students performance prediction using KNN and Naïve Bayesian." In *Information Technology (ICIT)*, 2017 8th International Conference on, IEEE, 2017, pp. 909-913.
 - [8] Kataria, Aman, and M. D. Singh. "A review of data classification using k-nearest neighbour algorithm." *International Journal of Emerging Technology and Advanced Engineering* 3, no. 6, 2013, pp. 354-360
 - [9] Agaoglu, Mustafa. "Predicting instructor performance using data mining techniques in higher education." *IEEE Access* 4, 2016, pp: 2379-2387.
 - [10] Homsı, Masun Nabhan, Natasha Medina, Miguel Hernandez, Natacha Quintero, Gilberto Perpiñan, Andrea Quintana, and Philip Warrick. "Automatic heart sound recording classification using a nested set of ensemble algorithms." In *Computing in Cardiology Conference (CinC)*, 2016, pp. 817-820.
 - [11] Singh, Harmanjit, Gurdev Singh, and Nitin Bhatia. "Fuzzy cognitive maps based election results prediction system." *International Journal of Computers & Technology* 7, no. 1, 2013, pp: 483-492.
 - [12] Sevarac, Zoran. "Neuro fuzzy reasoner for student modeling." In *Advanced Learning Technologies*, 2006. Sixth International Conference on, IEEE, 2006, pp. 740-744.
 - [13] Arsad, Pauziah Mohd, and Norlida Buniyamin. "Neural network model to predict electrical students' academic performance." In *Engineering Education (ICEED)*, 2012 4th International Congress on, IEEE, 2012, pp. 1-5.
 - [14] Sikder, Md Fahim, Md Jamal Uddin, and Sajal Halder. "Predicting students yearly performance using neural network: A case study of BSMRSTU." In *Informatics, Electronics and Vision (ICIEV)*, 2016 5th International Conference on, IEEE, 2016, pp. 524-529.
 - [15] Li, Zhenpeng, Changjing Shang, and Qiang Shen. "Fuzzy-clustering embedded regression for predicting student academic performance." In *Fuzzy Systems (FUZZ-IEEE)*, 2016 IEEE International Conference on, IEEE, 2016, pp. 344-351..
 - [16] Bhatia, Nitin, and Sangeet Kumar. "Prediction of Severity of Diabetes Mellitus using Fuzzy Cognitive Maps." *Advances in Life Science and Technology*, Vol. 29, 2015, pp:71-78.
 - [17] Yadav, Ramjeet Singh, and Vijendra Pratap Singh. "Modeling academic performance evaluation using fuzzy c-means clustering techniques." *International Journal of Computer Applications* 60, no. 8, 2012.
 - [18] Deshpande, Minal, and Preeti R. Bajaj. "Short term traffic flow prediction based on neuro-fuzzy hybrid sytem." In *ICT in Business Industry & Government (ICTBIG)*, International Conference on, IEEE, 2016, pp. 1-3.
 - [19] Hamsa, Hashmia, Simi Indiradevi, and Jubilant J. Kizhakkethottam. "Student academic performance prediction model using decision tree and fuzzy genetic algorithm." *Procedia Technology* Vol 25, 2016, pp: 326-332.
 - [20] Rahman, Rashedur M., and Farhana Afroz. "Comparison of various classification techniques using different data mining tools for diabetes diagnosis." *Journal of Software Engineering and Applications* 6, no. 03 2013.
 - [21] Ashari, Ahmad, Iman Paryudi, and A. Min Tjoa. "Performance comparison between Naïve Bayes, decision tree and k-nearest neighbor in searching alternative design in an energy simulation tool." *International Journal of Advanced Computer Science and Applications (IJACSA)* 4, no. 11, 2013.
 - [22] Rusli, Nordaliela Mohd, Zaidah Ibrahim, and Roziah Mohd Janor. "Predicting students' academic achievement: Comparison between logistic regression, artificial neural network, and Neuro-fuzzy." In *Information Technology*, 2008. ITSIm 2008. International Symposium on, vol. 1, IEEE, 2008, pp. 1-6.