

Optimising the parameters of a RBFN network for a teaching learning paradigm

Pamela Chaudhury, Hrudaya Kumar Tripathy

School of Computer Engineering, Kalinga Institute of Industrial Technology (KIIT) University, India

Article Info

Article history:

Received Sep 12, 2018

Revised Nov 17, 2018

Accepted Feb 28, 2019

Keywords:

Academic performance

Classification

DE

RBFN

TLBO

ABSTRACT

Academic performance of students has been a concern worldwide. Despite efforts made by educational institutions there has been a rise in poor academic performance. In our research study we have proposed a model to pre-determine the academic performance of students using a Radial Basis Function network (RBFN) using primary data. The proposed model has been developed by using algorithms like differential evolution (DE) and teaching learning based optimization (TLBO). This model can be used by academic institutions to identify the academically weaker students and take preventive steps to reduce the number of academic failures.

Copyright © 2019 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Pamela Chaudhury,
School of Computer Engineering,
KIIT University,
Bhubaneswar, Odisha, India.
Email: pamela.chaudhury@gmail.com

1. INTRODUCTION

Data mining techniques have been used in the field of education since several years. Educational data mining has been used to do several tasks like academic performance prediction, curriculum designing, student modeling based on behavior, analyzing the learning patterns of students and investigating the academic datasets to discover patterns in the data [1, 2]. Academic performance prediction is one of the most interesting applications of educational data mining. It involves analyzing students' academic, non-academic data and using classification, prediction models or regression models to pre-determine how a particular student might perform in the upcoming examination [3, 4]. Accurate determination of academic results would enable the stakeholders to take appropriate measure and reduce academic failures. Several researchers have been conducted to predict academic performance of students.

The students' data used by researchers consisted of different types of attributes. Academic performance prediction has been done using past CGPA of students [5]. Some researchers have used psychometric attributes of students for determining the student grades [6]. In-depth analysis has been done using academic and demographic attributes like social background, family support etc. [7, 8]. Attributes like internet access patterns have a profound impact on academic performance and can be used for the classification/predictive models [9, 10]. Researchers have used feature selection algorithms to remove irrelevant students' attributes from the dataset [11-13]. Algorithms like information gain, correlation based feature selection and relief based feature selection are used to identify the most important set of features of students' datasets. Different classification algorithms are used by researchers for generating the academic performance determination model. Decision tree algorithm has been used for the classification model [14].

Support vector machines, neural network and naïve Bayesian classifiers are the most popular classification algorithms in this domain and have been used by several researchers [15-17]. The problem with the existing academic performance determination models is as follows:

- a) Academic performance determination models were built using classification models which did not give very accurate results in pre-determining academic results of students.
- b) The datasets of students used in different research works often consisted of student attributes randomly without much analysis.
- c) Many of the attributes of students' dataset were often non numeric in nature. Most data mining algorithms cannot work on non numeric data and hence these attributes were not considered for student grade prediction.

In this research work we have proposed a model that aims to handle the above shortcomings. We have used optimization algorithms to enhance the classification model. The dataset used for this study was decided after an indepth analysis on the attributes of the student dataset. The non numeric attributes were converted into a form so that it could be used in data mining algorithms without losing valuable information contained in them. The paper is organised as follows: Section 2 describes the research method used to obtain the proposed model. Section 3 discusses the result obtained and its implication. Section 4 concludes the paper.

2. RESEARCH METHOD

2.1. Data Collection

The first step of the study consisted of collecting real students' data from engineering students of Biju Patnaik University and KIIT University. A lot of deliberation was done as to what should be the attributes of student dataset incorporated for the study. Detailed literature review was also done to find out student attributes used for academic performance prediction. Feature selection algorithms were used to find out the significant features in [18]. A detailed discussion was held with the parents, professors, placement officers and students. Finally the following attributes were included in the study as shown in Table 1.

Table 1. The Attributes of the Student Dataset

Student Attributes
Internal Grade
Attendance Percentage
Internet Usage Hours (weekly)
Study pattern: Daily or Before exam only approach
Previous backlogs (in any semester)
Participation in extra curricular activities
Division secured in Secondary
Division secured in higher secondary
Financial /family or health issues affecting studies
Past semester CGPA

The data was collected from 209 students of Biju Patnaik University and KIIT University. Out of the 209 students questionnaire only 9 were incomplete and 200 were included in the study.

2.2. Conversion of Non Numeric Attributes Into Dummy Variables

Since most attributes were categorical (non numeric) in nature there was a need to convert the data attributes into a format so that it could be used as an input for classification algorithms. While some researchers ignore the non numeric variables others substitute the categorical values with numbers. For example the attribute 'Participation in extra curricular activities' can have two values: 'frequently' or 'rarely'. So the value frequently could be substituted with 1 and the value rarely could be substituted with value 2. However this approach could bias the classification algorithm towards values 'frequently' since it has a higher value substituted numerically than the value 'rarely'. This approach often makes the classification models biased towards particular values and reduces the accuracy of these models.

We used the concept of dummy variables [19] which converts categorical attributes into a format suitable for classification algorithm without introducing bias. Table 2 demonstrates how a categorical attribute can be handled by introducing dummy variables for each attribute.

Table 2. Showing Introduction of Dummy Variables

Original Dataset Attribute: Participation in extra curricular activities	Converted categorical attribute: Participation in extra curricular activities frequently/rarely	
Attribute: Study Pattern	Attribute: Frequently	Attribute: Rarely
Frequently	1	0
Rarely	0	1
Rarely	0	1
Frequently	0	1

The Table 2 shows the introduction of dummy attributes: ‘Participation in extra curricular activities Frequently’ and ‘Participation in extra curricular activities: Rarely’. For each attribute value a 0 or 1 is assigned. Hence for each data instance 0 or 1 value will be added for each dummy attribute added. The number of dummy attributes will depend upon the number of values a categorical attribute would have. For example the attribute study pattern has two values: ‘daily’ and ‘before exam’. Hence two dummy attributes are introduced replacing the attribute study pattern.

2.3. RBF as a Classifier

Radial Basis Function are a class of neural network introduced by Broomhead and Lowe [20] and are now increasingly used by different researchers as an improved alternative to multilayer neural networks. The structure of RBFN consists of an input layer, a hidden layer and an output layer. The input layer consists of source nodes to feed the n dimensional input vector. The hidden layers are responsible for applying a linear transformation on input data using radial function and the output layer which implements linear transformation on the hidden layer output. The architecture of RBFN is displayed in Figure 1.

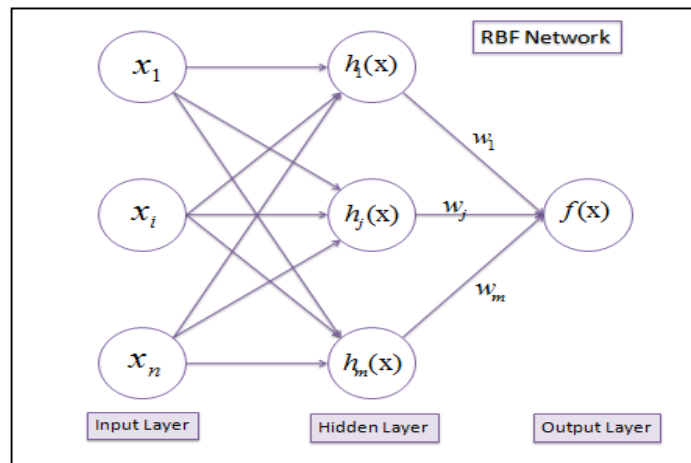


Figure 1. The architecture of RBFN

Each of the neurons of the hidden layer is used to keep the centers for the RBFN and applies the radial basis function on the center and the input data. We have used the Gaussian function as the radial basis function. The width of the bell curve of the Gaussian function is determined by the parameter called spread. Hence the output of the i th hidden neuron with center ϕ_i and spread μ_i .

$$\phi_i(x) = \exp\left(\frac{\|x - \mu_i\|^2}{2\sigma^2}\right) \tag{1}$$

The output layer consists of the same number of units as the number of classes. We have divided the output class into three categories of students’ academic performance: poor, average and outstanding. Therefore there are three units in the output layer of the RBFN. These output units are called activations and are multiplied with the weight of the links from the hidden layer to the output layer as shown in (2).

$$y_{net} = w_0 + \sum_{i=1}^N w_i e^{\left(\frac{-\|x - \mu_i\|^2}{2\sigma^2}\right)} \tag{2}$$

The classification performance of the RBFN highly depends upon the selection of center and spread of each neuron in the hidden layer. The best performance would be obtained if every instance in the training could be used as a center and the spread is calculated based on the average Euclidean distance from center to examples in training set. As the data set is huge it is not possible to use all instances as centers, hence a limited amount of selected instances are used for hidden neurons.

The most naïve way of using a RBFN classifier is to randomly select some centers and calculate the spread, however this method gives varying levels of accuracy of the classification model. Therefore we optimized the centers and spreads using the evolutionary algorithms.

2.4. Optimization Algorithms

Optimization techniques are used for selection of the best element based on some criterion from a set of variety of elements available. Optimisation algorithms are able to produce the best solution. Many of these evolutionary optimization techniques are inspired from biological processes like reproduction, mutation etc [21]. We have used the Teaching Learning based Optimisation techniques (TLBO) and the differential evolution methods for optimizing the center and spread of our proposed model based on the RBFN classification.

TLBO: The TLBO algorithm is proposed by Rao and Vakharia [22]. It is based on how a student improves his/ her performance in the class. The student not only learns from the teacher but also from his classmates and this enables to improve the overall performance of the students in the class. TLBO simulates this behavior of teachers and learners inside a classroom. The algorithm consists of two phases: the teaching phase and the learner phase. The group of students inside the classroom are called the population, result is the objective function, the teacher is the best solution, different subjects are the design variable. The design variables are the parameters of the objective function which has to be optimized. The best solution is the best value of the objective function. In TLBO the objective function is taken as the error value between 0 and 1 and hence the effort is to reduce or minimize the objective function.

$$T_F = \text{round}(1 + \text{rand}(0, 1)) \quad (3)$$

$$Y_{i,t+1} = Y_{i,t} + \text{rand}(0, 1) * (Y_{\text{teacher}} - T_F * Y_{\text{mean}}) \quad (4)$$

Where, TF is teaching function with a value either 1 or 2 and is randomly selected by the algorithm and Y_{mean} is the mean value attribute wise. The following is the algorithm for teaching learning based optimization.

Algorithm: Teaching Learning Based Optimization

1. Generate an initial population randomly. ($Y_i, I = 1, 2, \dots, N$)
 2. Evaluate the objective function for each member
 3. Untill the termination criterion is not met
 - 3.1. Find the best teacher
 - 3.2. $Y_{\text{new},i} = \text{Difference_Mean}$
 - 3.3. Calculate the objective function for $Y_{\text{new},i}$
 - 3.4. if $F(Y_{\text{new},i}) > F(Y_i)$ then
 - 3.5. $Y_i = Y_{\text{new},i}$
 - 3.6. For $i = 1$ to N do
 - 3.7 Select two students Y_j and Y_k randomly
 - 3.8 if $F(Y_j) > F(Y_k)$ then
 - Obtain new solution $Y_{\text{new},i}$ using Eq.5
 - else
 - Obtain new solution $Y_{\text{new},i}$ using Eq.6
 - 3.9. Calculate the objective function for $Y_{\text{new},i}$
 - 3.10 if $F(Y_{\text{new},i}) > F(Y_i)$ then
 - 3.3.5.1. $Y_i = Y_{\text{new},i}$
 4. find the best value
-

For minimizations:

If $f(Y_{i,t})$ is less than $f(Y_{j,t})$

$$Y_{i,t+1} = Y_{i,t} + \text{rand}(0, 1) * (Y_{i,t} - Y_{j,t}) \quad (5)$$

If $f(Y_{j,t})$ is less than $f(Y_{i,t})$

$$X_{i,t+1} = Y_{i,t} + rand(0, 1) * (Y_{j,r} - X_{i,t}) \tag{6}$$

DE: We have also used the differential evolution which is one of the most famous differential evolutionary algorithms proposed by Storn and Price [23]. The algorithm consists of mutation, combination and selection steps which continue till the termination criteria are met. The DE algorithm is as follows:

Algorithm: Differential Evolution Optimization

1. Set the generation number $G = 0$
 Select a random value from the dataset
2. while(end criterion is not met) do
 - 2.1. **Mutation**
 - 2.1.1. select three random indices $s1, s2, s3$
 - 2.1.2. calculate $v_{i,G+1} = x_{s1,G} + F(x_{s2,G} - x_{s3,G})$
 - 2.2. **Recombination**
 - 2.2.1. for every attribute in dataset
 - 2.2.2. if CR is better than or equal to random probability or index == random integer with range D
 - 2.2.3. then assign that attribute of v to u
 - 2.2.4. if CR better than random probability and index different to random integer with range D
 - 2.2.5. then assign that attribute of x to u
 - 2.3. **Selection**
 - 2.3.1. create a copy of input matrix, mat1, and replace the x value with v
 - 2.3.2. if accuracy of mat1 is greater than accuracy of original matrix
 - 2.3.3. return mat1
 - 2.3.4. else
 - 2.3.5. return original matrix

The output of mutation goes as a input to the process of recombination. In the process of recombination, a value u is generated with the help of the main input (x) and the mutated output(v)[24]. This process uses random comparisons to generate the value of u, which is either a part of x or vi i.e. the attribute of the output is either that of x or of v. The output is calculated by the following process:

$$u_i(t+1) = \begin{cases} v_{ip}(t+1) & \text{if } (rand \leq c_r) \text{ or } (i = rand(ind)) \\ x_{ip}(t+1) & \text{if } (rand > c_r) \text{ and } (i \neq rand(ind)) \end{cases} \tag{7}$$

c_r is the crossover constant [25]. If the child instance created or the trial vector has higher fitness function than the parent will be replaced.

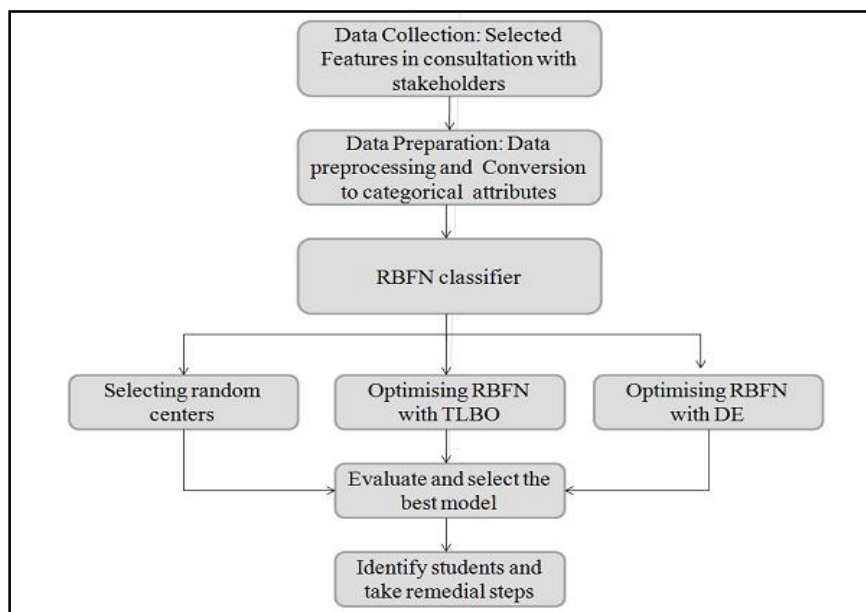


Figure 2. Our approach for developing the proposed model

The proposed model was created by evaluating and selecting the best model amongst the three models as shown in Figure 2. The RBFN classification model optimized with TLBO had the best classification accuracy.

Since RBF network performs best when optimized with TLBO, hence it is used for identifying the students who might fail in the upcoming examination. Accurate identification would enable educational institutions and students to take measures and remedial steps like bridge classes, increase in study hours etc. This might help to reduce the number of academic failures and improves the educational system.

3. RESULTS AND ANALYSIS

This section presents the experimental results of the proposed work and compares it with the traditional approach used so far. The data consisted of 10 attributes of students each and after the converting the attributes into dummies there were a total of 26 attributes. We have used classification accuracy to evaluate the classification model. The results of experiments have been demonstrated in Figure 3.

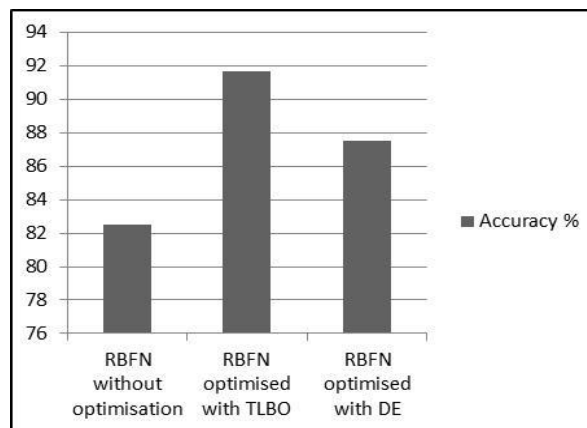


Figure 3. Classification accuracy of the optimised model for academic performance determination vs the traditional approach

The results clearly indicate that RBFN classification model built through our proposed model performs better the traditional approach. The existing models of determining academic performance of students do not use optimization techniques to fine tune the classification model. In our approach optimization techniques were used with RBFN classification model. Our proposed model uses two optimization techniques to find the best set of centers and spread for the RBFN. The accuracy of the classification model is highest when it is optimized with teaching learning based optimization model. We also evaluated the model with the Root mean square errors for each model and are shown in Table 3.

Table 3. Comparison of the Three Classification Models

Classification model used:	RMSE
RBFN without optimisation	0.2821
RBFN optimized with TLBO	0.2319
RBFN optimized with DE	0.2382

The error of the classification model using RBFN using TLBO was least when optimized with TLBO. Different set of experiments were conducted with 20, 40, 60, 80 and 100 iterations for the optimization techniques. Both TLBO and DE perform the best with around 40 iterations as shown in Figure 4. The RBFN model optimized with TLBO performs best with 40 iterations compared to all other models. The experimental results show a significant improvement in academic performance determination using our approach of optimizing the RBFN classification model compared to the previous approaches used by other researchers. Hence, this proposed model can be used by academic institutions to make accurate predictions about the students that might underperform in examinations and also gives them time to take appropriate remedial steps.

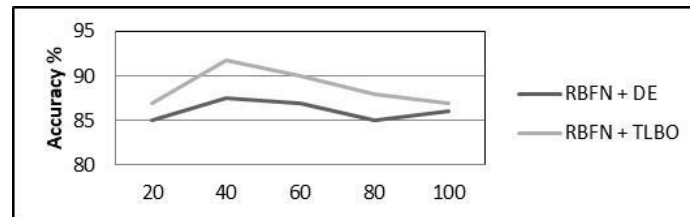


Figure 4. Depicts Accuracy % vs No of iterations of the classification models

4. CONCLUSIONS

This paper focuses on building an accurate classification model for determining the students' performance in a teaching learning environment like a classroom. The study was conducted on primary data. Selected student attributes were collected and used with a RBFN model. Optimisation algorithms were used for selecting the center and spread of RBFN model. Our proposed model consists of the RBFN optimized with TLBO and outperforms other classification models. This model when used by academic institutions can improve the teaching learning process. It helps academic institutions to take remedial steps in advance for students, who are indicated by the classification model as potential candidates for academic failures.

REFERENCES

- [1] C. Romero and S. Ventura, "Educational data mining: a review of the state of art", IEEE transactions on systems, man and Cybernetics – Part C : application and Reviews, vol. 40 (6), pp. 601-618, 2010.
- [2] Yu Zhike, "A New BP Neural Network Algorithm and Its Application in English Education Evaluation", TELEKOMNIKA, Vol.11, No.11, pp. 6441~6446, November 2013.
- [3] A.P. Ayala, "Educational data Mining: a survey and a data mining based analysis of recent works", Expert Systems with Applications, vol. 41, no. 4, pp. 1432-146, 2014.
- [4] C. Evandro et. al., "Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses", Computers in Human Behavior, vol. 73, pp. 247-256, 2017.
- [5] Natek, M. Zwilling, "Student data mining solution–knowledge management system related to higher education institutions", Expert systems with applications vol. 41 (14),pp. 6400–6407, 2014.
- [6] G. Gray, et. al., "An application of classification models to predict learner progression in tertiary education", Advance Computing Conference (IACC), IEEE International, pp. 549-554, 2014.
- [7] P. Cortez, and A. Silva, "Using Data Mining to Predict Secondary School Student Performance", in Proceedings of 5th FUTURE BUSINESS TECHNOLOGY CONFERENCE (FUBUTEC 2008), pp. 5-12, 2008.
- [8] B.K. Bharadwaj and S. Pal, 'Data Mining : A prediction for performance improvement using classification', in International Journal of Computer Science and Information Security, vol. 9, (4), pp. 136-140, 2011.
- [9] Paul A. Kirschner, Aryn C. Karpinski, "Facebook® and academic performance", Computers in Human Behavior, Vol. 26 (6), pp. 1237-1245, 2010.
- [10] P. Chaudhury & H. K. Tripathy, "A Study on impact of smartphone addiction on academic performance", International Journal of Engineering & Technology, Vol 7, June 2018, pp. 50-53.
- [11] M. Zaffar, et.al., "Performance analysis of feature selection algorithm for educational data mining," IEEE Conference on Big Data and Analytics (ICBDA), Kuching, pp. 7-12, 2017.
- [12] A. Acharya and D. Sinha, "Application of Feature Selection Methods in Educational Data Mining", International Journal of Computer Applications, (0975 – 8887) Vol. 103 (2), October 2014.
- [13] Yuan-Dong Lan, "A Hybrid Feature Selection based on Mutual Information and Genetic Algorithm", Indonesian Journal of Electrical Engineering and Computer Science, Vol. 7, No. 1, July 2017, pp. 214 – 225.
- [14] A. Raheela et. al., "Predicting Student Academic Performance at Degree Level: A Case Study", International Journal of Intelligent Systems and Applications. Vol. 7. pp. 49-61, 2014.
- [15] P. Chaudhury, et al., "Enhancing the capabilities of student result prediction system", in Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies (ICTCS '16), Udaipur, India, ACM, New York, NY, USA, Article 91, 04–05 March, 6pp.
- [16] P. Guleria and M. Sood, "Classifying Educational Data Using Support Vector Machines: A Supervised Data Mining Technique", Indian Journal of Science and Technology, vol. 9(34), pp. 1-5, 2016.
- [17] M. Mayilvaganan, and D. Kalpanadevi, "Comparison of classification techniques for predicting the performance of students academic environment", International Conference on Communication and Network Technologies (ICCNT), IEEE, pp.113–118, 2014.
- [18] P. Chaudhury & H. K., "An Empirical Study on Attribute Selection of Student Performance Prediction Model' in International Journal of learning Technology, Inderscience, vol. 12(3), pp. 241-252, 2017.
- [19] David S. Salkever, "The use of dummy variables to compute predictions, prediction errors, and confidence intervals", Journal of Econometrics, Vol. 4 (4), pp. 393-397, 1976.

- [20] D. S. Broomhead and D. Lowe, "Radial basis functions, multi-variable functional interpolation and adaptive networks.", Royal Signals & Radar Establishment Memorandum 4148, 1988.
- [21] Jianfeng Qiu, Jiwen Wang, Dan Yang, Juan Xie "A Comparison of Improved Artificial Bee Colony Algorithms Based on Differential Evolution", Indonesian Journal of Electrical Engineering and Computer Science, Vol 11, No 10: October 2013.
- [22] R.V. Rao, *et.al.*, "Teaching-learning-based optimization: A novel method for constrained mechanical design optimization problems", Computer-Aided Design, Vol. 43(3), pp. 303-315, 2011.
- [23] R. Storn and K. Price, "Differential Evolution—A Simple and Efficient Adaptive Scheme for Global Optimization over Continuous Spaces", Technical Report TR-05-012: International Computer Science Institute, Berkely, 22,1995.
- [24] Liu, J., Lampinen, J., "A Differential Evolution Based Incremental Training Method for RBF Networks", In Proceedings of GECCO 2005, pp. 881-888.
- [25] Das, S., Suganthan, P. N., "Differential evolution: A survey of the state-of-the-art," IEEE Transactions on Evolutionary Computation, 2011, vol. 15, No. 1, pp. 4-31.

BIOGRAPHIES OF AUTHORS



Pamela Chaudhury is an Assistant Professor in the department of Computer Science and Engineering at Silicon Institute of Technology, Bhubaneswar, India. She is also a PhD scholar in school of Computer Science and Engineering, KIIT University, Bhubaneswar, India. Her research interests include machine learning, artificial intelligence based applications and educational data mining. She is currently working in the domain of educational data mining. She has many publications in the field of machine learning and voice based applications. She is member of several professional bodies including IEEE



Dr. Hrudaya Kumar Tripathy presently working as Associate Professor at School of Computer Engineering, KIIT (Deemed to be University), Bhubaneswar, India. He had been a visiting faculty in Asia Pacific University, Kuala Lumpur, Malaysia and University Utara Malaysia, Sintok, Malaysia. He is having 18 years of teaching experience with post-doctorate research experience in the field of Soft Computing, Machine Learning, Speech Processing, AI, Mobile Robotics and Big Data Analysis. Received many certificates of merits and highly applauded in a presentation of research papers at International conferences. He has published number of research papers in reputed international & national refereed journals & conferences. He is a senior member of IEEE, life member of CSI & having membership in other different professional bodies such as IET, IACSIT, IAENG.