

ABSTRACT

As we know that in maximum university the evaluation of the student's performance is done manually by the faculties. This System of student performance evaluation is non-transparent and often leads to dissatisfaction among students. This project aims to solve the problem by designing a user interface which would work on supervised learning using Neural Network. Data mining techniques are widely used in educational field to find new hidden patterns from student's data. The hidden patterns that are discovered can be used to understand the problem arise in the educational field. Data Mining (DM), or Knowledge Discovery in Databases (KDD), is an approach to discover useful information from large amount of data. DM techniques apply various methods in order to discover and extract patterns from stored data. The pattern found will be used to solve a number of problems occurred in many fields such as education, economic, business, statistics, medicine, and sport. The large volume of data stored in those areas demands for DM approach because the resulting analysis is much more precise and accurate.

KEYWORDS: education, KDD, Neural Network, performance, evaluation.

INTRODUCTION

Normally in Education System, the student performance evaluation is done by faculty manually. This System of student performance evaluation is non-transparent and often leads to dissatisfaction among students. This project aims to solve the problem by designing a user interface which would work on supervised learning using Neural Network. Data Mining techniques are widely used in educational field to find new hidden patterns from student's data. The hidden patterns that are discovered can be used to understand the problem arise in the educational field. Data Mining (DM), or Knowledge Discovery in Databases (KDD), is an approach to discover useful information from large amount of data. DM techniques apply various methods in order to discover and extract patterns from stored data. The pattern found will be used to solve a number of problems occurred in many fields such as education, economic, business, statistics, medicine, and sport. The large volume of data stored in those areas demands for DM approach because the resulting analysis is much more precise and accurate.

This project aims at mining student's data using MATLAB and to categorise students according to input data as Best, Average and Worst. Educational Data Mining (EDM) is concerned with developing methods and analysing educational content to enable better understanding of student's performance. It is also important to enhance teaching and learning process.

VARIOUS DATA MINING TECHNIQUES

Association: Association is one of the best-known data mining techniques. In association, a pattern is discovered based on a relationship between items in the same transaction. That is the reason why association technique is also known as relation technique. The association technique is used in market analysis to identify a set of products that customers frequently purchase together.

Classification: Classification is a classic data mining technique based on machine learning. Basically, classification is used to classify each item in a set of data into one of a predefined set of classes or groups. Classification method makes use of mathematical techniques such as decision trees, linear programming, neural network and statistics. In classification, we develop the software that can learn how to classify the data items into groups. For example, we can apply classification in the application that "given all records of employees who left

the company, predict who will probably leave the company in a future period.” In this case, we divide the records of employees into two groups that named “leave” and “stay”. And then we can ask our data mining software to classify the employees into separate groups.

Clustering: Clustering is a data mining technique that makes a meaningful or useful cluster of objects which have similar characteristics using the automatic technique. The clustering technique defines the classes and puts objects in each class, while in the classification techniques, objects are assigned into predefined classes.

Prediction: The prediction, as its name implied, is one of a data mining techniques that discovers the relationship between independent variables and relationship between dependent and independent variables. For instance, the prediction analysis technique can be used in the sale to predict profit for the future if we consider the sale is an independent variable, profit could be a dependent variable.

Sequential Patterns: Sequential patterns analysis is one of data mining technique that seeks to discover or identify similar patterns, regular events or trends in transaction data over a business period.

Decision trees: The A decision tree is one of the most common used data mining techniques because its model is easy to understand for users. In decision tree technique, the root of the decision tree is a simple question or condition that has multiple answers. Each answer then leads to a set of questions or conditions that help us determine the data so that we can make the final decision based on it. Machine learning through Neural Network by using classification technique provide the most accurate results in data mining. So supervised learning is done through Neural Network in this project.

LITERATURE SURVEY

A thorough literature review has been conducted in preparation for the following project methods and to inform findings and recommendations. Data Mining is recently widely used in the field of education for student performance evaluation, performance prediction, future predictions etc. In our project titled “**Student**

Performance Evaluation through Supervised learning using neural network” we have used various concepts of data mining such as concept of Neural Network and back propagation algorithm which would evaluate students performance and would classify them as best, average or worst.

Before beginning with the implementation of the project we have researched various papers on Educational data mining, each of which explains different algorithm for educational data mining. A explanation of researched papers is provided below:

- Brijesh Kumar Baradwaj, Research Scholar, Singhaniya University, Rajasthan, India
Mining Education Data to Analyse student’s Performance, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 6, 2011[1] Present paper is designed to justify the capabilities of data mining techniques in context of higher education by offering a data mining model for higher education system in the university. In this research, the classification task is used to evaluate student’s performance and as there are many approaches that are used for data classification, the decision tree method is used here. By this task knowledge is extracted that describes students’ performance in end semester examination. It helps earlier in identifying the dropouts and students who need special attention and allow the teacher to provide appropriate advising/counseling.
- Firstman Noah, BaahBarida, Taylor Onate Egerton, Evaluation of Student Performance Using Data Mining Over a Given Data Space, International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-2, Issue-4, September 2013 [2] In this paper, the performance evaluation of students, were presented using data mining technique and cluster checking. The system examined students who gained admission into the University of Port-Harcourt through the University Matriculation Examination (UME) and through Basic studies programme with the aim of finding out variations in their performance when they graduate from the university. The evaluation was done using data mining technique to find out the ratio that falls into grouping of the grading in the various classes using the cumulative grade point average (CGPA) and the students who failed out.
- Ashwani Kharola, SwarnimaKunwar, Gopa B Choudhury, Institute of Technology Management, Defence Research & Development Organisation, Ministry of Defence, Government of India, Students Performance Evaluation: A fuzzy logic reasoning approach

PM World Journal Vol. IV, Issue IX – September 2015 [3] This paper presents a new fuzzy logic reasoning based approach for performance evaluation of students in school or college. The attributes considered for evaluation cover academic as well as personality traits of the students. A Stage-wise fuzzy reasoning approach has been used to eliminate the issues of rule explosion. The comparison between fuzzy and traditional average technique shows the advantage of weightage allocation in fuzzy approach.

- Vaneet Kumar, Dr. Vinod Sharma, Student's Examination Result Mining: A Predictive Approach, International Journal of Scientific & Engineering Research, Issue11, November 2012

[4] This paper takes into consideration the various factors and their influences on student performance in education and predicts their final examination result whether pass or reappear. Various factors such as previous year results, attendance, financial status of family, parental education qualification, internet use for study material, parents visit to school, extra class etc. play an important role in education. This paper discusses the most common measurable factors among students. The result of this paper present an idea about these factors and their influence on student performance. MATLAB is used to predict outcome which helps to predict final examination result.

- Azwa Abdul Aziz, NurHafieza Ismail, Fadhilah Ahmad, Mining Students' Academic Performance, Journal of Theoretical and Applied Information Technology, 31st July 2013. Vol. 53 No.3

[5] This paper explains that Data Mining techniques are widely used in educational field to find new hidden patterns from student's data. The hidden patterns that are discovered can be used to understand the problem arise in the educational field. This paper surveys the three elements needed to make prediction on Students' Academic Performances which are parameters, methods and tools. This paper also proposes a framework for predicting the performance of first year bachelor students in computer science course. Naïve Bayes Classifier is used to extract patterns using the Data Mining Weka tool. The framework can be used as a basis for the system implementation and prediction of Students' Academic Performance in Higher Learning Institutions.

- C. Anuradha and T. Velmurugan, A Comparative Analysis on the Evaluation of Classification Algorithms in the Prediction of Students Performance, Indian Journal of Science and Technology, Vol 8(15), IPL057, July 2015 [6]The primary objective of this research work is to apply the classification techniques to the prediction of the performance of students in end semester university examinations. Particularly, the decision tree algorithm, Bayesian classifiers, Nearest Neighbor algorithm and two rule learner's algorithms namely OneR and JRip are used for classifying the performance of students as well as to develop a model of student performance predictors. The result of this study reveals that overall accuracy of the tested classifiers is above 60%. In addition classification accuracy for the different classes reveals that the predictions are worst for distinction class and fairly good for the first class. The JRip produces highest classification accuracy for the Distinction. Classification of the students based on the attributes reveals that prediction rates are not uniform among the classification algorithms.

OUR WORK

Grouping students as best, average and worst in performance. Admin can upload and save students data for machine learning. Faculty can upload and maintain student's record. Faculty can modify (insert, update and delete) student's record. Supervised learning of the machine is done using Neural Network. The admin can view uploaded student data. Excel sheet can be directly imported to upload students record. Project provides a graphical user interface. The final result can be directly exported to excel file. The project uses neural network (back propagation) which provide the most accurate results. The faculty as well as the admin can view final mining result.

The Back Propagation Algorithm

Backpropagation, or propagation of error, is a common method of teaching artificial neural networks how to perform a given task. The back propagation algorithm is used in layered feedforward ANNs. This means that the artificial neurons are organized in layers, and send their signals "forward", and then the errors are propagated backwards. The back propagation algorithm uses supervised learning, which means that we provide the algorithm with examples of the inputs and outputs we want the network to compute, and then the error (difference between actual and expected results) is calculated. The idea of the back propagation algorithm is to reduce this error, until the ANN learns the training data.

Summary of the technique:

1. Present a training sample to the neural network.
2. Compare the network's output to the desired output from that sample. Calculate the error in each output neuron.

3. For each neuron, calculate what the output should have been, and a scaling factor, how much lower or higher the output must be adjusted to match the desired output. This is the local error.
4. Adjust the weights of each neuron to lower the local error.

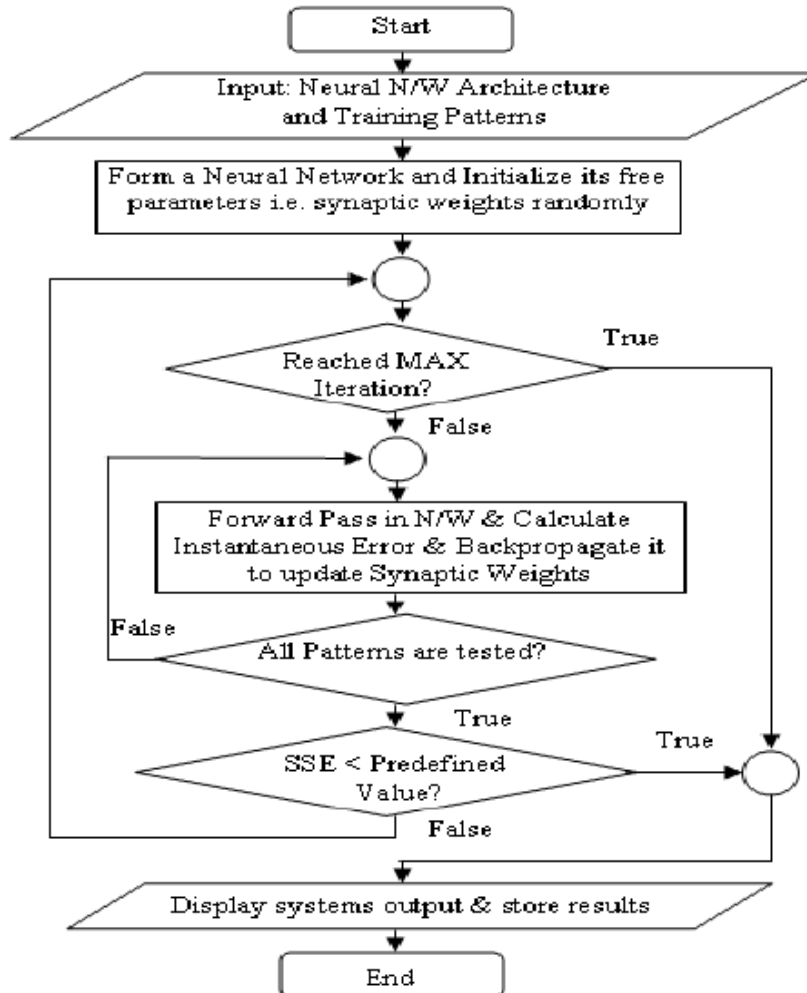


Figure: Flowchart of back propagation neural network algorithm.

Actual Algorithm

1. Initialize the weights in the network (often randomly)
2. repeat * for each example e in the training set do
 1. O= neural-net-output(network, e) ; forward pass
 2. T = teacher output for e
 3. Calculate error (T - O) at the output units
 4. Compute delta_wi for all weights from hidden layer to output layer ; backward pass
 5. Compute delta_wi for all weights from input layer to hidden layer ; backward pass continued
 6. Update the weights in the network * end
3. until all examples classified correctly or stopping criterion satisfied
4. return(network)

Advantages of Neural Network

1. High Accuracy: Neural networks are able to approximate complex non-linear mappings
2. Noise Tolerance: Neural networks are very flexible with respect to incomplete, missing and noisy data.
3. Independence from prior assumptions: Neural networks do not make a priori assumptions about the distribution of the data, or the form of interactions between factors.

4. Ease of maintenance: Neural networks can be updated with fresh data, making them useful for dynamic environments.
5. Neural networks can be implemented in parallel hardware
6. When an element of the neural network fails, it can continue without any problem by their parallel nature.

Design Problems

1. There are no general methods to determine the optimal number of neurones necessary for solving any problem.
2. It is difficult to select a training data set which fully describes the problem to be solved.

Solution To Improve Neural Network Performance

1. Designing Neural Networks using Genetic Algorithms
2. Neuro-Fuzzy Systems

CONCLUSION & FUTURE ENHANCEMENT

The conclusion reached is that we have created a system which would evaluate students' performance and thus classify students as best average and worst. In our project we have focused on various algorithms which could be used in various predictions in Educational Data Mining. When implemented on huge level it will be highly beneficial for organizations (especially colleges).

You can add more different features to this project like:

- Modification in proposed algorithm to increase efficiency of results.
- Integration of decision tree algorithm along with back propagation algorithm so as to get good efficiency as well as easy understanding.
- This system could be extended for evaluation of employees in various organizations.

REFERENCES

- [1] Brijesh Kumar Baradwaj, Research Scholar, Singhaniya University, Rajasthan, India, Mining Education Data to Analyse student's Performance, (IJACSA) International Journal of Advanced Computer Science and Applications Vol. 2, No. 6, 2011 <http://www.ijrte.org/attachments/File/v2i4/D0796092413.pdf>
- [2] Firstman Noah, BaahBarida, Taylor Onate Egerton, Evaluation of Student Performance Using Data Mining Over a Given Data Space, International Journal of Recent Technology and Engineering (IJRTE)ISSN: 2277-3878, Volume-2, Issue-4, September 2013 <http://arxiv.org/ftp/arxiv/papers/1201/1201.3417.pdf>
- [3] AshwaniKharola, SwarnimaKunwar, Gopa B Choudhury, Institute of Technology Management, Defence Research & Development Organisation, Ministry of Defence, Government of India, Students Performance Evaluation: A fuzzy logic reasoning approach, PM world gernal volume IV issue IX-September 2015, www.pmworldjournal.net
- [4] Vaneet Kumar, Dr. Vinod Sharma, Student's Examination Result Mining: A Predictive Approach, International Journal of Scientific & Engineering Research, Volume 3, Issue 11, November 2012.
- [5] Azwa Abdul Aziz, NurHafieza Ismail, Fadhilah Ahmad, Mining Students' Academic Performance, Journal of Theoretical and Applied Information Technology, 31st July 2013. Vol. 53 No.3
- [6] Han and Camber, Data Mining: Concepts and Techniques, 2nd edition, Morgan Kaufmann Publishers.
- [7] Deepa and Shivnandam, Introduction to Neural Networks Using Matlab 6.0, Tata Mcgrew Hills.
- [8] C. Anuradha and T. Velmurugan, Bharathiar University, Coimbatore, India, A Comparative Analysis on the Evaluation of Classification Algorithms in the Prediction of Students Performance Indian Journal of Science and Technology, Vol 8(15), IPL057, July 2015.