Think India Journal
ISSN: 0971-1260 Vol-22, Special Issue-11
6th International Conference on Innovation and Research on Technology and Engineering
held on 19-20 December 2019
Venue- **Padmabhushan Vasantdada Patil Pratishthan's College of Engineering**.Vasantdada Patil Education Complex, Sion - Chunabhatti, Near Everard Nagar, Mumbai, Maharashtra 400022

A Multilayer Back-Propagation Model to Predict Students' Performance

Sachin R. Barahate
Associate Professor
Department of Computer Engineering
PVPP's College of Engineering, Mumbai, India
bara.sachin@gmail.com

Abstract

One of the challenges faced by the higher education system today is to improve the quality of education and decisions support system in the institute. The decision making process in the institute becomes more intricate as the complexity of data related with educational system increases. Educational institute seek out more efficient technology to support decision making procedures and to formulate better management plans. This can be achieved by utilizing valuable implicit knowledge, which is currently unknown.

Artificial Neural Networks is the one of the promising mining tool. It has better performance than that of many other traditional data mining techniques, so it can be used for narrowing knowledge gap that exists in higher learning institutes. Educational data mining through artificial neural networks will help to enhance traditional educational procedures.

This paper proposes a model for predicting students' academic performance of final year BE students. A Multilayer Back-Propagation Network Algorithm is applied to the students' data in order to predict the students' academic performance based on certain variables. The extracted knowledge from model will be used to predict the student's academic performance in final examination.

Keywords: Data mining, Educational data mining, Back propagation algorithm, Multilayer artificial neural network.

1.Introduction

The plenty of valuable information is hidden within the educational data. This information is often useful but often remains buried and untapped. It becomes very difficult for institutions or organizations with limited technological resources to explore and gain any insight from the data.

"Data Mining is a bundle of algorithms for the extraction or mining knowledge, pattern identification from large amounts of data". There are many synonymous terms with slightly different meaning to data mining, such as knowledge mining from databases, knowledge extraction, data pattern analysis, data archaeology, and data dredging. Furthermore, identification of unknown patterns and relationships can also result in prediction of future outcomes. The importance of data mining has been established in many fields like criminal investigations,

business applications, bio-medicine. Data mining can be applied on huge amount of data to get some insight in data which can be useful in strategic decision making.

Educational Data Mining (EDM) is an emerging multidisciplinary research area, in which methods and techniques for exploring data originating from various educational information systems have been developed. Educational data mining is a rich application area for data mining and learning science, due to the growing availability of educational data. Educational data mining contributes to the study of how students learn, and the settings in which they learn.

These days there is another field of computational science that coordinates the diverse techniques for critical thinking that can't be so effectively portrayed without an algorithmic traditionally core interest. These techniques, in somehow, have their cause in the copying, pretty much smart, of the conduct of the organic frameworks. It is another method for processing designated Artificial Intelligence, which through various techniques is equipped for dealing with the impressions and vulnerabilities that show up when attempting to take care of issues identified with this present reality, offering solid arrangement and of simple usage. One of those strategies is known as Artificial Neural Networks (ANN). Motivated, in their source, in the working of the human mind, and entitled with some insight. These are the blend of an incredible measure of components of procedure counterfeit neurons interconnected that working parallels get the chance to take care of issues identified with parts of arrangement.

The paper endeavours to develop a Back propagation ANN model for prediction of the students' academic performance.

2.Literature Survey

Bo Liu, & Jian-Hua Zhao in [1], has tried to identify how to use data mining techniques in mainly online computer and web based leanings. Author has mainly describing about applying non-linear correlation technique in educational data mining. The suggested technique is more efficient and faster for identifying interesting patterns. In combination with clustering technique or association rule mining technique NLCDM is more useful for identifying the class of student. This technique has not been used yet for actually predicting performances of student or classifying the students. This technique needs instructor to visualize the obtained pattern in order to detect problematic situation as explained technique is capable of providing patterns only and not the knowledge.

In proposed model of [2] the authors are identifying the parameters of undergraduate students after their first year degree examinations that have impact and affect their performance. Authors also have compared the combination of decision tree classification techniques,clustering algorithms and artificial neural network for finding better technique for data mining in education domain. The ANN and Decision tree are compared in order to select best of the data mining techniques. This paper has able to prove that ANN is better techniques as it is less complex and provides accurate result. In this paper students are classified only in two groups i.e. successful and unsuccessful. In order to get accurate result of class of student, clustering is used in

combination with decision tree classification so that increasing complexity of model as compared to ANN.

Vo Thi Ngoc Chau, in [3] suggested a system which uses a knowledge-driven educational decision support approach by taking advantage of EDM. They have pointed out the difficulties in knowledge discovery from education system are, First, Educational programs most of the times changes over the particular period of time so that the models designed for knowledge discovery may become obsolete. Second difficulty is that, the training data set used is most of the time related to part of educational program, so the model designed may not be applicable for each and every student. Third difficulty is that most of the knowledge discovery models talks about the knowledge of courses students successfully obtained but it is also important to consider the knowledge student will obtained. The fourth difficulty is that data set used in the knowledge discovery model may b small as compared to the overall data generated by all the educational organizations.

In the proposed work of [4] multilayer back propagation algorithm is used to predict average rainfall over Udupi district of Karnataka. The paper suggests that along with ANN we can use other statistical and mathematical model also. Such models are very accurate in calculation but failed in accurate prediction. To overcome such type of real-life situations it is observed to use the artificial neuron which can learn from past precepts i.e. by adjustment of weights the system improves its performance in every iteration.

The author of [5] discusses the EDM technique used to identify the different dimensions and their relevance to predict the academic performance of students. They used the different dimensions like students demographics, past academic performance etc. to build the prediction model. The data is divided into multiple splits that fragments the dimension's range into multiple sub range. As a result algorithm needs the huge amount of data to get the acceptable results.

J. Jacob and team in their research work [6] explained the different EDM techniquesand how these techniques are beneficial for all the stakeholders like students, teachers, management etc. This paper given a very brief comparison of major EDM techniques, which is very helpful in understanding the different approaches for the application these data mining techniques on data belong to education system.

Artificial Neural Network (ANN) is one such field which can be connected for data mining capacities [7]. This work concentrates on data characterization capacity conventional data mining procedures and also simulated neural system methods. Author utilized some datasets under characterization, apply traditionally and additionally ANN systems for arrangement lastly analyse the aftereffects of two orders. This will give us more exactness of grouping and will likewise help in selecting better calculation for arrangement.

3.Proposed System

One important aspect of higher education system is related with quality of education. Higher learning institutes face various difficulties to achieve their quality objective. Several of these problems arise due to un-adequate knowledge. Knowledge gap is the lack of information or

significant knowledge at the educational main processes such as counselling, registration and evaluation etc.

In this proposed system the associations, hidden patterns and anomalies that are discovered by EDM algorithms can help bridge this knowledge gap in higher educational institutions. The pattern and knowledge discovered by EDM algorithms would enable the educational institutions in making better academic decisions, predicting individual behaviours with higher accuracy, having more advanced planning in directing students, and enabling the institution to allocate resources and staff more effectively. It results in improving the effectiveness and efficiency of the processes thus, maintaining quality of education.

This system is the application of ANN which aims at predicting student's academic performance at the end of the semester depending on the data captured during that semester. The work flow of the proposed system is as shown in figure 3.1.

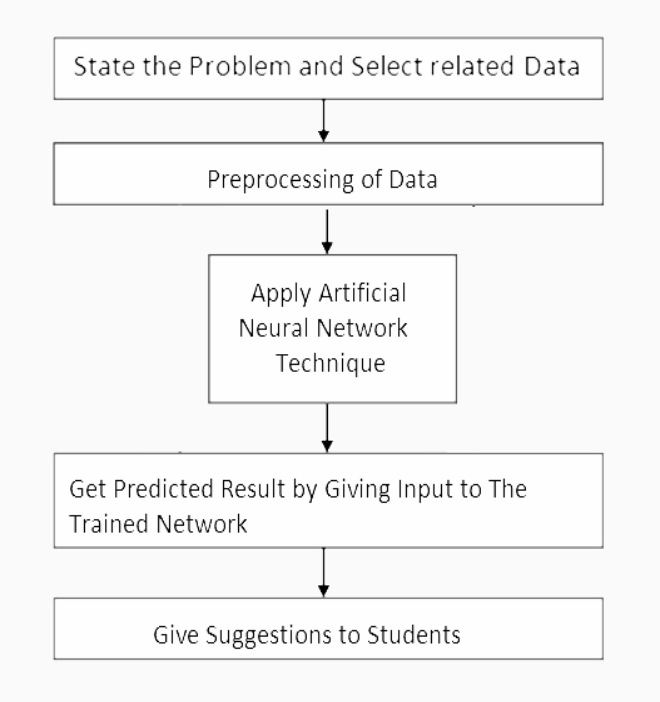

Fig 3.1 Proposed model

Here we will see what are inputs, algorithm and output.

3.1 Inputs

In this proposed system we have selected seven input variablesto predict the student's academic performance, which are shown in table 1.

Table 1: Parameters for prediction of students' performance

Parameters	Attribute Type/ % Range
Marks in previous Sem	(Number, Range 0 - 100)
Marks in Mid Term Test	(Number, Range 0 - 100)
Marks in End Term Test	(Number, Range 0 - 100)
Marks in oral / practical	(Number, Range 0 - 100)
Percentage attendance	(Number, Range 0 - 100)
Extra-curricular activities	(Binary, (yes, no))
Gender	(Binary, (Male, Female))

These inputs cannot be directly applied to the ANN. So, initially different data normalization technique must be applied to data to get the quality data. All numeric columns are scaled during data normalization. All columns are scaled using the below formulae:

$$SF = \frac{(USR_{max} - LSR_{min})}{(A_{max} - A_{min})}$$

$$Yp = (USR_{max} + (Y - A_{min}) * SF$$

Where,
Y - Actual value of a numeric column
LSR_{min} - Lower scaling range limit

A_{min} - Minimum actual value of the column
A_{max} - Maximum actual value of the column
USR_{max} - Upper scaling range limit
SF - Scaling factor
Y_p - Pre-processed valueFor input columns scaling range is [-1, 1].

3.2 Multilayer Perception's Back-propagation Algorithm

Multilayer Perception's Back-propagation or propagation of mistake is a neural system knowledge calculation. It is a technique for educating manufactured neural systems how to perform a given undertaking. The back-propagation calculation is utilized as a part of layered feed forward ANNs. This mean the artificial neurons are sorted out in layers, and send their signals forward and afterward then errors are proliferated in reverse. The back-propagation algorithm utilizes supervised learning, which mean that we give the algorithm cases of the inputs and output we need the system to figure, and afterward the mistake (contrast in the middle of real and expected results) is ascertained. The thought of the back-propagation algorithm is to overcome these errors, until the ANN takes in the preparation data.

A Back-propagation neural network is an arrangement of associated input/output units where every association has a weight connected with it. Amid the learning stage, the system learns by modifying the weights in order to have the capacity to foresee the right class label of the inputs.

Back-propagation Algorithm:

1. Initialize the weights in the ANN
2. Repeat the following statements
   * For each samplet in the training set do
   (a) Out = neural_net_ output (network; t);
   Forward pass
   (b) U = teacher output for t
   (c) Calculate error (U - Out) at the output units
   (d) Compute delta w_i for all weights from unseen layer to output layer;
   Backward pass
   (e) Compute delta w_i for all weights from input layer to unseen layer;
   Backward pass continued
   (f) Update the weights in the network
   * end
3. until all samples classified properly or stopping criterion mate
4. Return ;

The various parameters and their initial value considered as follows:

- Learning rate of the network (default) = 0.2
- Momentum = 0.7
- Transfer function : SIGMOID function
- Number of input neuron = 7
- Number of unseen neurons = 11
- Number of output neurons = 7

3.3 Output

We considered seven output neurons in proposed ANN. As we provided the sample input to the model one of the six output neuron is activated. This output of activated neuron will be the prediction of students' academic performance based on the data presented in table 1. The neuron will take the one of the following possible values.

("\Excellent", "\Very Good", "\Good", "\Average", "\Poor", "\Fail")

4.Result and Analysis

In this proposed system,we applied Naïve Bayesian Classifier, C4.5 Decision tree classifier algorithms and back-propagation neural network algorithm on the pre-processed data and assess the correctness of each algorithm.

All data mining algorithms can be assessed as far as their prescient accuracy, speed of calculations, power, adaptability, interpretability and so on. But, most basic parameter for calculation assessment and correlation is the prescient exactness. Prescient accuracy is only the capacity of a calculation to anticipate the obscure future variables accurately. For decision and execution of prevalent calculation following analysis has been conducted. All the execution correlations, assessments and investigation are performed with the assistance of WEKA instrument.

Regularly, preparing data is utilized to infer a classifier and afterward to gauge the exactness of the classifier. This can bring about misdirecting over-idealistic evaluations. Cross acceptance arbitrarily parcels the given data in k-folds. The utilization of cross approval procedures to gauge classifiers exactness builds the general calculation time, yet is valuable for selecting among a few classifiers. 10-fold cross validation is calculated as follows. To start with, the dataset is partitioned in two areas haphazardly. First segment is comprises of 90% dataset and called as training dataset. Another set area is incorporated 10% of dataset and called as testing datasets. Second, the training dataset is utilized for gaining rules, and the testing dataset is utilized for approving the standards remotely. Third, the procedure is rehashed 10 times. The exactness evaluation is the general number of right orders from the 10 emphases, isolated by the aggregate number of tests in the beginning data.

Taking after observations are made which includes rate exactness, rate mistake, root mean square mistake, accuracy, review, time to construct order model. It can be observed that prediction accuracy of ANN based calculation is most elevated inside of all calculations for the data sets under study.

Table 2: Data Classification using Naïve Bayesian Classifier

Cross Validation	Accuracy in percentage	Error in Percentage	RMS Error	Precision	Recall	Build Time (Sec.)
Use Directly	76.302	23.697	0.413	0.75	0.76	0.02
5-fold	76.432	23.567	0.416	0.76	0.76	0.00
10-Fold	76.302	23.697	0.413	0.75	0.76	0.00

Table 3: Data classification using C4.5 Decision Tree

Cross Validation	Accuracy in percentage	Error in Percentage	RMS Error	Precision	Recall	Build Time (Sec.)
Use Directly	84.114	15.885	0.345	0.84	0.84	0.09
5-fold	71.224	28.776	0.464	0.71	0.71	0.05
10-Fold	73.828	26.171	0.446	0.73	0.73	0.05

Table 4: Data classification using Multilayer Back-propagation Algorithm

Cross Validation	Accuracy in percentage	Error in Percentage	RMS Error	Precision	Recall	Build Time (Sec.)
Use Directly	80.599	19.40	0.381	0.81	0.80	3.16

5-fold	75.130	24.869	0.413	0.74	0.75	2.67
10-Fold	75.390	24.909	0.421	0.75	0.75	2.44

Despite the fact that Multilayer Perceptron back-propagation calculation devours generally more time for learning than different calculations, the result (exactness) is apparent. The interpretability of C4.5 decision tree algorithm is simple because of basic IF-THEN principles made utilizing tree. Neural Networks creates high level of accuracy, yet the translation is troublesome because of their mind boggling structure. Both the algorithms have points of interest and disadvantages. In any case, because of attributes, for example, robustness, adaptability, self-arranging structure, parallel preparing, disseminated capacity and high level of adaptation to internal failure, I can say that Neural System based algorithm is superior as compared to other algorithms.

5.Conclusion

As the time passes the amount of data generated in any educational institute grows exponentially. In order to extract the knowledge and meaningful information from such a hugeamount of data and to correctly predict the performance of students we used the seven parameters. The experimental result shows that the prediction accuracy of a Multilayer Back-propagation Neural Network based algorithm is commendable as compared to other algorithms for the dataset under consideration. But Back propagation neural network algorithm takes more time for learning than other algorithms.This model will assist the instructor to identify the weak students initially and take the necessary actions to improve the students performance.

References

[1] Jian-Hua Zhao, Bo Liu,"Non-linear correlation techniques in educational data mining", Sixth International Conference on Fuzzy Systems and Knowledge Discovery, IEEE,2009.

[2] M. Wook, Norshahriah Wahab, Yuhanim Hani Yahaya,Mohd Rizal Mohd Isa, Nor Fatimah Awang, Hoo Yann Seong, "Predicting NDUM Student's Academic Performance Using Data Mining Techniques", Second International Conference on Computer and Electrical Engineering, IEEE ,2009.

[3] Vo Ngoc Chau, "A Knowledge-Driven Educational Decision Support System", IEEE, 2009.

[4] K. Abhishek, Rajeev Ranjan, Abhay Kumar, Sarthak Kumar "A Rainfall Prediction Model using Artificial Neural Network", Control and System Graduate Research Colloquium IEEE, 2012.

[5] Mohd Sharifuddin Ahmad, Alan Cheah Kah Hoe, Tan Chin Hooi, Mohana Shanmugam, Saraswathy Shamini Gunasekaran, Zaihisma Che Cob, Ammuthavali Ramasamy, " Analyzing Students Records to Identify Patterns of Students' Performance", International Conference on Research and Innovation in information , IEEE, 2013.

[6] J. Jacob, Paarth Kotak, Kavya Jha, Shubha Puthran, "Educational Data Mining techniques and their Applications", International Conference on Green Computing and Internet of Things, IEEE, 2015.

[7] S. Nordbotten "Data Mining with Neural Networks", Svein Nordbotten and Associates Bergen 2006.