IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-6, June 2019 E-ISSN: 2347-2693

Performance Prediction In Educational Data Mining using Neural Network

Shankar G. Mundhe[1*], Shital Y. Gaikwad[2]

[1,2]Dept. of CSE, Matoshri Pratishtans Grop of Instititions Nanded. Maharashra, India

Corresponding Author: mundhe.shankar@gmail.com

DOI: https://doi.org/10.26438/ijcse/v7i6.761764 | Available online at: www.ijcseonline.org

Accepted: 08/Jun/2019, Published: 30/Jun/2019

***Abstract*—** Educational system is facing the challenges to improve the quality of education and to make managerial decisions to achieve the quality. Student's academic performance is the reflection of previous academic background. Performance prediction helps educational institute to support decision making procedures and to formulate better management plans. This performance record is important for the educational institution because they can learn from this to improve their quality by knowing the performance of the student. Educational data mining analyze these data and extract information from it. We can determine the status of student's academic performance. For achieving this we can use techniques like decision tree, classification, data clustering and neural network and so on. In this paper, we will predict student's semester wise performance.

***Keywords*—**Educational Data Mining, Neural Network, Normalization, Knowledge Discovery in Databases, Prediction, Backpropagation algorithm.

I. INTRODUCTION

There is valuable information hidden in data. Since the underlying data is generated much faster than it can be processed and made sense of, this information often remains buried and untapped. It becomes virtually impossible for individuals or groups with limited resources specifically technological to find and gain any insight from the data.

Data Mining encompasses tools and techniques for the extraction or mining knowledge from large amounts of data. There are many other terms carrying a similar or slightly different leaning to data mining, such as knowledge mining from databases, knowledge extraction, data pattern analysis, data archaeology, and data dredging. Another popularly used term, Knowledge Discovery in Databases" , or KDD.

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. EDM is both a learning science, as well as a rich application area for data mining, due to the growing availability of educational data. EDM contributes to the study of how students learn, and the settings in which they learn. It enables data-driven decision making for improving the current educational practice and learning material.

Educational Data Mining (called EDM) is an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand students, and the settings which they learn in. A key area of EDM is mining computer logs of student performance [5]. Another key area is mining enrollment data. Key uses of EDM include predicting student performance, and studying learning in order to recommend improvements to current educational practice today[5].

Educational data mining (EDM) develops methods and applies techniques from statistics, machine learning and data mining to analyze data collected during teaching and learning. EDM tests learning theories and improves educational practices [10].

The Artificial Neural Networks (ANN's) represent an alternative for endowing to the computers one of the characteristics that makes the difference between humans and other alive beings, the intelligence. An artificial neural network is an abstract simulation of a real nervous system and its study corresponds to a growing interdisciplinary field which considers the systems as adaptive, distributed and mostly nonlinear, three of the elements found in the real applications. The ANNs are used in many important engineering and scientific applications, some of these are, signal enhancement, noise cancellation, pattern classification, system identification, prediction, and control. Besides, they are used in many commercial products, such as modems,

image processing, recognition systems, and biomedical instrumentation, among others.

In this paper section I contains introduction of the educational data mining (EDM), Section II contains the existing algorithm definitions for Educational Data Mining (EDM), Section III contains the implementation challenges, Section IV contains the brief about Neural Network, Section V contains the algorithm implemented i.e. Multilayer Back-propagation Algorithm , Section VI Result and Discussion .

II. RELATED WORK

Existing algorithms used for EDM are Naive Bayes, Multilayer Perceptron and Decision tree etc. and many more. A decision tree is a decision support tool that uses a tree-like model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. It is one way to display an algorithm that only contains conditional control statements.

Decision trees are commonly used in operations research, specifically in decision analysis, to help identify a strategy most likely to reach a goal.

Classification and Prediction

Classification is the processing of finding a set of models which describe and distinguish data classes or concepts, for the purposes of predicting the class of objects whose class label is unknown. The derived model is based on the analysis of a set of training data [3]. The derived model may be represented in various forms, such as classification (IF-THEN) rules, decision trees, and mathematical formulae. A decision tree is a flow-chart-like tree structure, where each node denotes a test on an attribute value, each branch represents an outcome of the test, and tree leaves represent classes. Decision trees can be easily converted to classification rules. In many applications, one may like to predict some missing or unavailable data values rather than class labels. This is usually the case when the predicted values are numerical data, and is often specifically referred to as prediction. Prediction can be viewed as a type of classification. Classification and prediction may need to be preceded by relevance analysis which attempts to identify attributes that do not contribute to the classification or prediction process.

Data classification

Classification is the forms of data analysis which can be used to extract models describing important data classes or to predict future data trends. It is perhaps the most familiar and most popular data mining technique. Estimation and prediction can be viewed as types of classification. Definition: Given a database D = {t1, t2... tn} of tuples and a set classes C = {C1, C2... Cm}, the classification problem is to define a mapping f: D → C where each ti is assigned to one class. A class, Cj, contains precisely those tuples mapped to it; that is, Cj = {ti |f(ti) = Cj , $1 \le i \le n$, and ti ∈ D}. This definition views classification as a mapping from the database to the set of classes. All classes are predefined, nonoverlapping and partition the entire database. Each tuple in the database is assigned to exactly one class.

Normally, classification problem is implemented in two phases: 1. Create a specific model by evaluating the training data. This step has as input the training data and as output a model developed. 2. Apply the model developed in step 1 for classifying the tuples from the target database.

Issues in Classification:

• Missing Data: Missing data values cause problems during both the training phase and to the classification process itself. Missing values in the training data must be handled carefully and may produce an inaccurate result. There are various approaches for handling missing data:

1. Ignore the missing data.
2. Assume a value for missing data.
3. Use prediction methods-Assume a special value for missing data.

• Measuring Performance: Classification results of two classification techniques are never same. Accuracy of classification is the most important performance measure for classification problem.

Classification accuracy is usually calculated by determining the percentage of tuples placed in the correct class. Confusion matrix can be used for calculating accuracy of classification.

Features of educational data

An important and unique feature of educational data is it is hierarchical. Data at the concept level, the answer level, the session level, the student level, the classroom level, the teacher level, and the institute level are nested inside one another. Other important features are time, sequence, and context. Time is important to capture data such as length of practice sessions or time to learn. Sequence represents how concepts build on one another and how practice and tutoring should be ordered. Context is important for explaining results and knowing where a model may or may not work. Educational data mining researchers have the following goals for their research:

1. Predicting student's future learning behavior by creating student models that incorporate such detailed information as students' knowledge, motivation, performance and attributes [5].
2. Discovering or improving domain models that characterize the content to be learned and optimal instructional sequences.
3. Studying the effects of different kinds of pedagogical support that can be provided by learning software.

Implementation Challenges for EDM

Successful application of educational data mining will not come without effort, cost, and a change in educational culture to more frequent use of data to make decisions. Implementation challenges for EDM includes technical challenges, institutional capacity, and ethical issues.

Technical challenges usually, educational data is collected all over the institutional settings; hence it is big in size and hierarchical in structure. A challenge for successful implementation of educational data mining techniques is having sufficient technical resources for using big data and incurring the expenses associated with software services and storage in either remote servers or local servers. Lack of data interoperability among different data systems used in educational institutes imposes a challenge to educational data mining and analytics that rely on diverse and distributed data. Lack of consistency among government data systems requires data translation and alignment.

Limitations in institutional capacity Technical challenges can be overcome through research, development, and testing, and computing and storage can be budgeted as part of an institutions infrastructure costs. However, implementing educational data mining and learning analytics in gigantic institutions have costs that go beyond simply computing and storage. Significant human resources also are needed for data preparation, processing, and analysis.

Ethical issues Educational data mining make predictions and recommend actions based on increased visibility into student actions, and these give rise to a number of social and ethical concerns. Policymakers bear an ethical responsibility to investigate the validity of any predictive model that is used to make consequential decisions about students.

What is an Artificial Neural

"Artificial Neural Networks are massively interconnected networks in parallel of simple elements (usually adaptable), with hierarchic organization, which try to interact with the objects of the real world in the same way that the biological nervous system does".

Benefits of Artificial Neural Networks, It is evident that the ANN obtains their efficacy from:

1. Its structure which is massively distributed in parallel. The information processing takes place through the iteration of a great amount of computational neurons, each one of them send exciting or inhibiting signals to other nodes in the network. Differing from other classic Artificial Intelligence methods where the information processing can be considered sequential – this is step by step even when there is not a predetermined order, in the Artificial Neural Networks this process is essentially in parallel, which is the origin of its flexibility. Because the calculations are divided in many nodes, if any of them gets astray from the expected behavior it does not affect the behavior of the network.

2 Its ability to learn and generalize. The ANN has the capability to acquire knowledge from its surroundings by the adaptation of its internal parameters, which is produced as a response to the presence of an external stimulus. The network learns from the examples which are presented to it, and generalizes knowledge from them. The generalization can be interpreted as the property of artificial neural networks to produce an adequate response to unknown stimulus which are related to the acquired knowledge. These two characteristics for information processing make an ANN able to give solution to complex problems normally difficult to manage by the traditional ways of approximation.

Additionally, using them gives the following benefit:

• No linearity: The answer from the computational neuron can be linear or not. A neural network formed by the interconnection of nonlinear neurons, is in itself nonlinear, a trait which is distributed to the entire network. No linearity is important over all in the cases where the task to develop presents a behavior removed from linearity, which is presented in most of real situations.

• Adaptive learning: The ANN is capable of determine the relationship between the different examples which are presented to it, or to identify the kind to which belong, without requiring a previous model.

• Self–organization: This property allows the ANN to distribute the knowledge in the entire network structure; there is no element with specific stored information.

• Fault tolerance: This characteristics is shown in two senses: The first is related to the samples shown to the network, in which case it answers correctly even when the examples exhibit variability or noise; the second, appears when in any of the elements of the network occurs a failure, which does not impossibilities its functioning due to the way in which it stores information.

III. METHODOLOGY

While doing experiments for all three methods are validated using 10 Cross Fold validation. Car dataset contains total 1728 instances. Iris dataset contains total 10 instances. Weather dataset contains total 14 instances.

Implementation of Multilayer Perceptron's Backpropagation algorithm:

1. Initialize the weights in the network (often randomly)

2. repeat

* For each example e in the training set do

(a) O = *neural_net_ output (network, e);*

forward pass
(b) T = *teacher output for e*
(c) Calculate error (T - O) at the output units
(d) Compute *delta_wi* for all weights from hidden layer to output layer;
backward pass
(e) Compute *delta_wi* for all weights from input layer to hidden layer;
backward pass continued
(f) Update the weights in the network
* End
3. Until all examples classified correctly or stopping criterion satisfied
4. *return (network);*

For implementation purpose of above algorithm, various parameters for the ANN are as follows [3]:
Default learning rate of the network = 0.2
Default momentum = 0.7
Transfer function: SIGMOID function
Number of input neuron = 7
Number of hidden neurons = 11
Number of output neurons = 7

IV. RESULTS AND DISCUSSION

For experimentation purpose three methods are chosen Naive Bayes, Tree J-48 and Multilayer Perceptron-Backpropagation Algorithm. Three datasets are considered for each method to calculate accuracy, error and build time required for each algorithm.

Table 1. Accuracy in Percentage

Dataset \ Classifier	Naive Bayes	Tree J-48	Multilayer Perceptron-Backpropagation
Car	85.5324	92.3611	99.537
Iris	96	96	97.333
Weather	64.2857	64.2857	78.5714

Table 2. Error in Percentage

Dataset \ Classifier	Naive Bayes	Tree J-48	Multilayer Perceptron-Backpropagation
Car	14.4676	7.6389	0.463
Iris	4	4	2.667
Weather	35.7143	35.7143	21.4286

Table 3. Error in Percentage

Dataset \ Classifier	Naive Bayes	Tree J-48	Multilayer Perceptron-Backpropagation
Car	0.01	0.11	18.36
Iris	0	0.02	0.28
Weather	0	0	0.03

V. CONCLUSION AND FUTURE SCOPE

Comparing Naïve Bayes, Decision Tree and Multilayer Perceptron-Backpropagation methods for performance prediction Multilayer Perceptron-Backpropagation algorithm gives more accuracy. Hence Multilayer Perceptron-Backpropagation algorithm is used for Performance Prediction In Educational Data Mining.
Multilayer Perceptron-Backpropagation algorithm can be used for prediction system where accuracy required is more.

REFERENCES.

[1] Prof.Sonal kadu , Prof.Sheetal Dhande "Effective Data Mining Through Neural Net- works", IJARCSSE - Volume 2, Issue 3, March 2012 ISSN: 2277 128X
[2] Dr.Yashpal Singh, Alok Singh Chauhan "Neural Networks in Data Mining" Journal of Theoretical and Applied Information Technology 2009.
[3] Md. Fahim Sikder Md. Jamal Uddin and Sajal Halder "redicting Students Yearly Performance usingNeural Network: A Case Study of BSMRSTU" IEEE 5th International Conference on Informatics, Electronics and Vision (ICIEV) 2016
[4] Hongjun Lu, Rudy Setiono, and Huan Liu"Effective Data Mining Using Neural Networks" IEEE Trans. Knowledge and Data Eng., vol. 8, no. 6.
[5] Permphan Dharmasaroja, Nicha Kingkaew "Application of Artificial Neural Networks for Prediction of Learning Performances" IEEE 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD) 2016.
[6] M.H.Dunham, S.Sridhar "Data Mining Introductory and Advanced Topics", Pearson Education 2007 ISBN 81-7758-785-4.
[7] Han J, Kamber, "Data Mining Concepts and Techniques", Morgan Kaufmann, M 2001.
[8] Witten, Ian H. and Frank, Eibe. "Data mining: Practical machine learning tools and techniques", Academic Press, 2012.
[9] Mrs. Bharati M. Ramageri, "Data Mining Techniques and Applications", IJCSE Vol. 1 No. 4301-305
[10]R. Baker, K. Yacef, "The State of Educational Data Mining in 2009: A Review and Future Visions", Journal of Educational Data Mining, Volume 1, Issue 11-3-17.

Authors Profile

Mr.Shankar.G.Mundhe. pursed Bachelor of Engineerng from Dr. Babasaheb Ambedkar Marathwada University, Aurangabad in 2005 and he is currently pursing Master of Engineering from Babasaheb Ambedkar Technical University, Lonere. He has 6 years of teaching experience.

Miss.Shital.Y.Gaikwad.. pursed Bachelor of Engineerng from Dr. Babasaheb Ambedkar Marathwada University, Aurangabad in 2004 and pursed Master Of Technology in Computer Network and Information Security from Swami Ramanand Tirt h Marathwada University, Nanded. she is currently working as Asst. Professor in Department of Compueter Science And Engineering in Matoshri Pratishtans Grop Of Instititions, Nanded . He has 13 years of teaching experience.

