International Journal of Innovative Research in Advanced Engineering (IJIRAE) ISSN: 2349-2163
Issue 08, Volume 6 (August 2019) www.ijirae.com

PERFORMANCE PREDICTION IN EDUCATIONAL DATA MINING USING NEURAL NETWORK

Shankar G. Mundhe
M.Tech II, Dept. Of CSE,
Matoshri Pratishtans Group of Institutions, Nanded. INDIA
mundhe.shankar@gmail.com

Shital Y. Gaikwad
Asst.Professor, Dept Of CSE
Matoshri Pratishtans Group of Instititions Nanded.INDIA
shitalygaikwad@gmail.com

Manuscript History
Number: **IJIRAE/RS/Vol.06/Issue08/AUAE10080**
Received: 01, August 2019
Final Correction: 07, August 2019
Final Accepted: 09, August 2019
Published: **August 2019**
Citation: Shankar & Shital (2019). Performance Prediction in Educational Data Mining Using Neural Network. IJIRAE:: International Journal of Innovative Research in Advanced Engineering, Volume IV, 582-585. **doi://10.26562/IJIRAE.2019.AUAE10080**
Editor: Dr.A.Arul L.S, Chief Editor, IJIRAE, AM Publications, India

Abstract— The students' 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 analyse these data and extract information from it. We can determine the status of student's academic performance. For determining the students performance techniques like decision tree, Naïve Bayes and neural network algorithms are compared and considering accuracy the neural network -Multilayer Perceptron Backpropagation Algorithm used.

Keywords—EDM- Educational Data Mining, NN-Neural Network, Normalization, KDD-Knowledge Discovery in Databases, Prediction, Backpropagation algorithm.

I. INTRODUCTION

The valuable information is hidden in EDM. 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.

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.

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.

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 live 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 and recognition systems, speech recognition, 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 brief about Neural Network, and Section IV contains the implementation of Multilayer Back-propagation Algorithm

II. EXISTING ALGORITHM DEIFINTIONS

Existing algorithms that I have considered for my experiment in EDM are Naive Bayes, Multilayer Perceptron Backpropagation algorithm and Decision tree but there exists many more algorithms.

In machine learning, naive Bayes classifiers are a family of simple "probabilistic classifiers" based on applying Bayes' theorem with strong (naive) independence assumptions between the features.

Naive Bayes has been studied extensively since the 1960s. It was introduced (though not under that name) into the text retrieval community in the early 1960s, and remains a popular (baseline) method for text categorization, the problem of judging documents as belonging to one category or the other (such as spam or legitimate, sports or politics, etc.) with word frequencies as the features. With appropriate pre-processing, it is competitive in this domain with more advanced methods including support vector machines. It also finds application in automatic medical diagnosis.

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, but are also a popular tool in machine learning.

The Back propagation neural network is multilayered, feedforward neural network and is by far the most extensively used. It is also considered one of the simplest and most general methods used for supervised training of multilayered neural networks. Backpropagation works by approximating the non-linear relationship between the input and the output by adjusting the weight values internally. It can further be generalized for the input that is not included in the training patterns (predictive abilities).

Generally, the Back propagation network has two stages, training and testing. During the training phase, the network is "shown" sample inputs and the correct classifications. For example, the input might be an encoded picture of a face, and the output could be represented by a code that corresponds to the name of the person.

Measuring Performance of Algorithms:

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.

III. WHAT IS AN ARTIFICIAL NEURAL NETWORK?

A neural network is a network or circuit of neurons, or in a modern sense, an artificial neural network, composed of artificial neurons or nodes. Thus a neural network is either a biological neural network, made up of real biological neurons, or an artificial neural network, for solving artificial intelligence (AI) problems. The connections of the biological neuron are modeled as weights. A positive weight reflects an excitatory connection, while negative values mean inhibitory connections. All inputs are modified by a weight and summed.

This activity is referred as a linear combination. Finally, an activation function controls the amplitude of the output. For example, an acceptable range of output is usually between 0 and 1, or it could be −1 and 1.

"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".

Neural networks are a set of algorithms, modeled loosely after the human brain, that are designed to recognize patterns. They interpret sensory data through a kind of machine perception, labeling or clustering raw input. The patterns they recognize are numerical, contained in vectors, into which all real-world data, are it images, sound, text or time series, and must be translated.

Neural networks help us cluster and classify. You can think of them as a clustering and classification layer on top of the data you store and manage. They help to group unlabeled data according to similarities among the example inputs, and they classify data when they have a labeled dataset to train on. (Neural networks can also extract features that are fed to other algorithms for clustering and classification; so you can think of deep neural networks as components of larger machine-learning applications involving algorithms for reinforcement learning, classification and regression.)

Benefits of neural network:

• No linearity: ANNs have ability to learn and model non-linear and complex relationships, which is really important because in real life, many of the relationships between inputs and outputs are non-linear as well as complex.
• 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.
• ANNs can generalize: After learning from the initial inputs and their relationship, it can infer unseen relationships on unseen data as well, thus making the model generalize and predict unseen data.
• 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.
• ANN does not impose any restrictions on input variables. Many studies have shown that ANN can better model data with high volatility and non-constant variance. This is something very useful in financial time series forecasting e.g. stock prices where data volatility is very high.

IV. ALGORITHM SELECTION FOR PERFORMANCE PREDICTION

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: Error in Percentage

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

Table 2: Build Time in Percentage

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

Table 3: Accuracy in Percentage

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

V. IMPLEMENTATION OF MULTILAYER PERCEPTRONS BACK-PROPAGATION

Actual input values given are in the range of 0 to 100. The values given as input are scaled in the range of -1 to +1. The scaling factor is 0.2 which is calculated using formula as below.

$$SF = (SR\ max - SR\ min) / (X\ max - X\ min)$$
$$Xp = SR\ min + (X - X\ min) * SF$$

Where,

X -Actual value of numeric column.
X min -Minimum Actual Value of Column
X max -Maximum Actual Value of Column
SR min -Lower scaling range limit
SR max -Upper scaling range limit
SF - Scaling Factor
Xp -Processed Value

For the input column scaling range is [-1 ,1]
For implementation purpose of the algorithm, various parameters for the ANN are as follows:

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

VI. CONCLUSION

Comparing Naive Bayes, Tree J-48 and Multilayer Perceptron- Back propagation Algorithms for performance prediction in EDM neural networks Multilayer Perceptrons –Back propagation algorithm gives more accuracy.

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.

