M-Learners Performance Using Intelligence and Adaptive E-Learning Classify the Deep Learning Approaches

Rajalakshmi T[1], Kaleeswari S[2], Dr. Jai Ruby[3], Mrs. E. Julie Ruth[4]

[1, 2]PG Scholar, Department of Computer Applications and Research Centre, Sarah Tucker College (Autonomous), Tirunelveli, Tamil Nadu, India
[3]Associate Professor, Department of Computer Applications and Research Centre, Sarah Tucker College (Autonomous), Tirunelveli, Tamil Nadu, India
[4]Assistant Professor, Department of Computer Applications and Research Centre, Sarah Tucker College (Autonomous), Tirunelveli, Tamil Nadu, India

---***---

Abstract - *Data mining techniques may assist in closing the knowledge gap in higher education. The data mining process aids in the improvement of educational efficiency. To increase student accomplishment, data mining techniques such as classification, association rule mining, clustering, prediction, and so on are applied. It aids in the management of their life cycle and the course selection. Classification is a crucial data mining process that may be used to great advantage in educational data. The implementation of a classification algorithm in education data mining is the topic of this research. The comparison research was carried out in order to forecast a student's academic achievement based on socioeconomic variables, previous test marks, and other factors connected to student performance. The experiment used the J48, Nave Bayes, Bayes Net, Back Propagation Network, and Radial Basis Function Network classification algorithms. The Radial Basis Function Network properly classified 100% of the instances, which is a high percentage when compared to other classifiers.*

Key Words: RBF Network, Naïve Bayes, Multilayer Perceptron, J48 algorithm, Educational Data Mining, Classification, WEKA

1. INTRODUCTION

The application of the data mining method to educational data is known as educational data mining (EDM). An EDM's goal is to examine educational data in order to enhance the performance of teachers, students, and educational institutions. For the benefit of learners, EDM blends computational theory, database management, and machine learning. Because education is so vital in every community, data mining researchers concentrate on EDM, which has evolved as a study subject in recent years. Data on education has been gathered from numerous educational surveys and school records, and data mining techniques like categorization may be used to enhance academic achievement. One of the most important requirements for successful education is student performance. Data mining is required in education for the benefit of students and academics. Educational data mining is a set of approaches for extracting new information from educational data, which may be used to better anticipate student behaviour, academic achievement, and topic interest, among other things [1-3]. Figure 1 illustrates the educational data mining system.

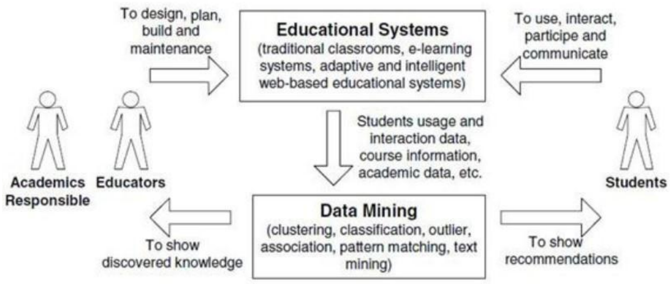

Fig -1: Educational Data Mining System

The figure above depicts the requirement for EDM. In EDM, all forms of education are considered educational systems, including conventional classrooms, E-learning systems, intelligent and adaptable web-based educational systems, and so on. As an input to the data mining process, task-relevant data is supplied. During the data mining process, you may choose a task-specific data-mining approach. The information and patterns generated by data mining are used by students, academics, and educators. This mechanism may be used to make student recommendations. Academics and educators sought to enhance the educational system in order to increase student performance. The information gathered may be used to enhance the educational system through organizing courses, academic activities, and student use [2-5].

2. RELATED RESEARCH

The academic institution must be able to forecast student academic achievement in order to increase student performance. Educational data mining has the potential to help predict student success. As a result, educational data mining researchers assisted in the development of supervised learning approaches for predicting student performance. This section will provide a quick overview of the categorization techniques used to forecast student performance.

For the prediction of student performance, Kabakchieva [6] used the decision tree, Bayesian classification, closest neighbour, and two rule learners (OneR and JRip). The decision tree classifier (J48) is the most accurate (with the greatest overall accuracy), followed by the rule learner (JRip) and the k-NN classifier. The Bayes classifiers aren't as good as the others. However, all of the examined classifiers have an overall accuracy of less than 70%, indicating that the error rate is large and the predictions are unreliable.

Norlida et al. [7] used a data mining approach to predict engineering student performance. The student's success was predicted and classified using the Cumulative Grade Point (CGPA). This research provided an overview of Neuro-fuzzy categorization.

Al-Saleem et al. [8] built a performance prediction model based on prior students' academic records and established categorization procedures. The model was constructed using decision tree classifications such as ID3 and J48. With the integration of this model and recommender system, it may assist students in course selection based on their graduating students' grades.

For the extraction of valuable information, Devasia et al. [9] used the Nave Bayesian mining approach. The experiment used a database of 700 students with 19 characteristics. Nave Bayesian classification was shown to be more accurate than Regression, Decision Trees, and Neural Networks.

For the prediction of a student's academic achievement, Hamsa et al. [10] used a decision tree and a fuzzy genetic algorithm. These models were evaluated using internal marks, sessional marks, and admission scores. The concept divides pupils into two groups: safe and risk. When compared to a fuzzy genetic algorithm, the data demonstrates that decision trees identify more students in the danger group.

Daud et al. [11] provided a strategy for forecasting student performance using advanced learning analytics. The expenditures of the family and the personal information of the students were evaluated in this research. The experimental study used the support vector machine, C4.5, Classification and Regression Tree (CART), Bayes Network, and Nave Bayes methods. The results reveal that the support vector outperforms the other feature sets in use. The results also show that Bayes Network and Nave Bayes classifiers outperform C4.5 and CART in most cases.

3. METHODOLOGY

This document is template. We ask that authors follow some simple guidelines. In essence, we ask you to make your paper look exactly like this document. The easiest way to do this is simply to download the template, and replace(copy-paste) the content with your own material. Number the reference items consecutively in square brackets (e.g. [1]). However, the authors name can be used along with the reference number in the running text. The order of reference in the running text should match with the list of references at the end of the paper.

3.1 Data Set

The UCI machine learning repository provided the student performance data set. School reports and questionnaires were used to compile the data collection. There are 32 characteristics in total in the data set. The data attributes include a student's first, second, and final grades, demographic information such as age and gender, student address type (urban or rural), social information such as mother's and father's educations, and school-related information such as study time, extra educational support, and extra paid classes.

3.2 Method of Supervised Learning

The machine learning task of supervised learning is to learn from a previously known class data set, often known as labelled training data set. The training data set includes a collection of input qualities as well as their corresponding output values. The classifier model is built by analysing the training data set and then used to categorize fresh samples with unknown class labels or desired output values. The approach for determining class labels for unknown examples is enabled by an ideal scenario. For the learning process to simplify from the training data to unseen scenarios, a realistic approach is necessary. J48, Bayes Net, Nave Bayes, Multilayer Perceptron, and Radial Basis Function classification algorithms were utilized in this research to compare and evaluate these approaches for predicting student performance.

J48 is an ID3 extension. Accounting for missing values, decision tree pruning, continuous attribute value ranges, rule generation, and other features are included in J48. J48 is an open-source Java implementation of the WEKA data mining tool.

Bayesian Networks (BN) are a probabilistic classification approach also known as belief networks. A directed acyclic graph or tree plus a collection of conditional probability distributions make up this system. Given the observable evidences, the purpose is to determine the posterior conditional probability distribution of each of the potential unseen causes. The provisional chance on each node is computed first, followed by the formation of a BN. The best assumption in Bayes Net is that all attributes are nominal, that no missing values exist, and that such values are replaced globally.

Models that give class labels to issue occurrences, represented as vectors of feature values, using the Nave Bayes (NB) approach. It is a family of techniques for training such classifiers based on the same principle: all naive Bayes classifiers assume that the value of one feature is independent of the value of any other feature, given the class variable.

A Multilayer Perceptron (MLP) is a kind of artificial neural network that uses layers to process information. There are at least three layers of nodes in a Multilayer Perceptron. Each node, with the exception of the input nodes, is a neuron with a nonlinear activation function. Back propagation is a supervised learning approach used by MLP during training. A Multilayer Perceptron is distinguished from a Linear Perceptron by its numerous layers and non-linear activation. It can differentiate non-linearly separable data. Back propagation network (BPN), a multilayer perceptron design, was employed in this investigation.

Each hidden unit implements a radial activation function, and each output unit implements a weighted sum of hidden unit outputs. RBF Network (RBFN) was also constructed, in which the process is based on a normalized Gaussian radial basis function network [3].

3.3 Convolutional Neural Networks (CNN)

CNNs are powerful image processing, artificial intelligence (AI) that use deep learning to perform both generative and descriptive tasks, often using machine vison that includes image and video recognition, along with recommender systems and natural language processing (NLP). A neural network is a system of hardware and/or software patterned after the operation of neurons in the human brain. Traditional neural networks are not ideal for image processing and must be fed images in reduced-resolution pieces. CNN have their “neurons” arranged more like those of the frontal lobe, the area responsible for processing visual stimuli in humans and other animals. The layers of neurons are arranged in such a way as to cover the entire visual field avoiding the piecemeal image processing problem of traditional neural networks. A CNN uses a system much like a multilayer perception that has been designed for reduced processing requirements. The layers of a CNN consist of an input layer, an output layer and a hidden layer that includes multiple convolutional layers, pooling layers, fully connected layers and normalization layers. The removal of limitations and increase in efficiency for image processing results in a system that is far more effective, simpler to trains limited for image processing and natural language processing.

Table -1: The performance Comparison for various classifiers

Method/Parameters	**J48**	**NB**	**BN**	**BPN**	**RBF**
Correctly classified Instances (%)	74.94	75.19	78.73	95.19	100
Incorrectly classified Instances (%)	25.06	24.81	21.26	04.81	0
Kappa statistic	0.72	0.72	0.76	0.94	1
Mean absolute error	0.036	0.037	0.0319	0.0069	0.0001
Root mean squared error	0.1341	0.133	0.1235	0.0583	0.0014
Relative absolute error (%)	37.26	38.87	33.03	7.16	0.0832
Root relative squared error (%)	61.09	60.59	56.26	26.55	0.6467
Total Number of Instances	395	395	395	395	395

4. EXPERIMENT AND RESULTS

The comparison research was carried out using Weka and the student performance data set. There are 32 characteristics in the student performance data collection. Age, sex, mother and father education status, study time, extracurricular activities, health, and other performance indicator variables were used to train the model, and the final grade was used as a predictor class. The performance of several classification methods is shown in table 1.

In terms of accurately identified instances, the BPN and RBF classifiers fared better, with 95.19 percent and 100 percent, respectively. J48, NB, and BN properly identified instances are 74.94 percent, 75.19 percent, and 78.73 percent, respectively, and these algorithms perform poorly when compared to BPN and RBF.

```
Classification Report:
----------------------
             precision    recall  f1-score   support

          0       0.66      0.72      0.69        32
          1       0.78      1.00      0.88        21
          2       0.76      0.60      0.68        43

avg / total       0.73      0.73      0.72        96

------------------------------
```

Fig -2: Logistic Regression Accuracy

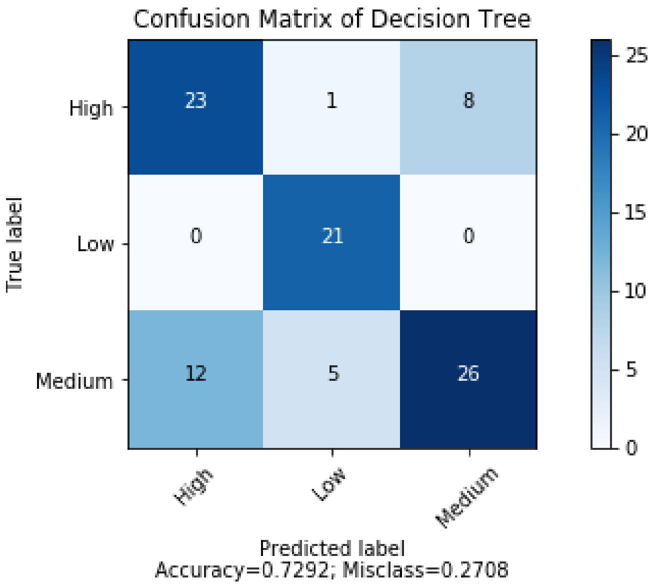

Fig -3: Logistic Regression Confusion Matrix

```
Classification Report:
----------------------
             precision    recall  f1-score   support

          0       0.66      0.91      0.76        32
          1       0.88      1.00      0.93        21
          2       0.89      0.58      0.70        43

avg / total       0.81      0.78      0.77        96

------------------------------
```

Fig -4: Naïve Bayes Accuracy Results

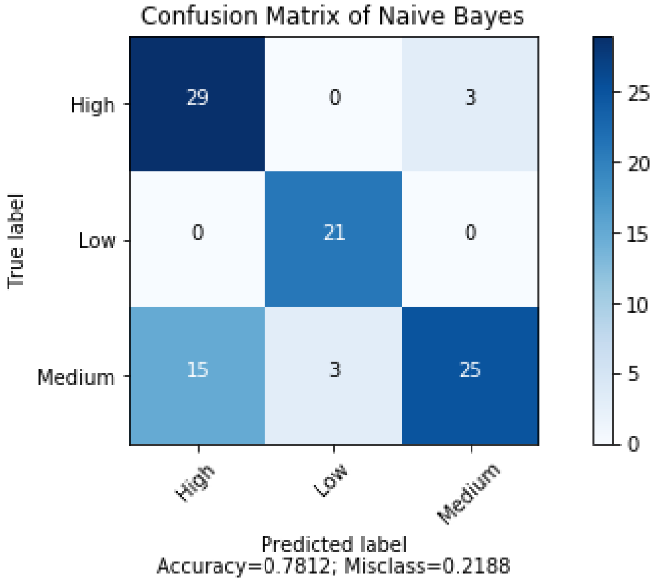

Fig -5: Naïve Bayes Confusion Matrix

```
Classification Report:
----------------------
             precision    recall  f1-score   support

          0       0.68      0.78      0.72        32
          1       0.84      1.00      0.91        21
          2       0.79      0.63      0.70        43

avg / total       0.76      0.76      0.76        96

------------------------------
```

Fig -6: CNN Accuracy

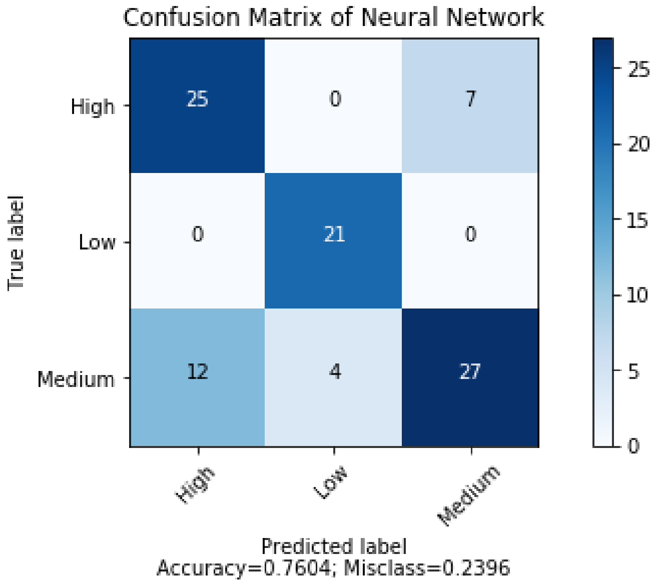

Fig -7: CNN Confusion Matrix

5. CONCLUSIONS

This study describes education data mining and applies classification approaches to student performance. There are several classification approaches, however it is crucial to choose which classification technique will be used on the data in order to improve student academic performance. Various categorization techniques were investigated in this research. RBF and BPN classification were shown to be superior algorithms for predicting student performance in a comparative study based on accuracy % in this area. To summarize, this article will give an insightful look at current solutions for student performance categorization. This may give students with self-assistance and anticipate achievement based on social, educational, and previous performance. Furthermore, such a method assists teachers and academic institutions in assessing student performance prior to the final test and taking required remedial action.

REFERENCES

[1] Romero, C. and Ventura, S., "Educational Data Mining: A Review of the State of the Art," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 40, 2010, pp. 601 - 618, doi: 10.1109/TSMCC.2010.2053532.

[2] Dutt, A., Ismail, M. A. and Herawan, T., "A Systematic Review on Educational Data Mining," IEEE Access, vol. 5, 2017, pp. 15991 - 16005, doi: 10.1109/ACCESS.2017.2654247.

[3] Hota, H. S., Sharma, L. K. and Pavani, S., "Fuzzy topsis method applied for ranking of teacher in higher education," Springer Intelligent Computing, Networking and Informatics, 2014, pp. 1225 - 1232. doi: https://doi.org/10.1007/978-81-322-1665-0_127.

[4] David, L. M. and Carlos E. G., "Data Mining to Study Academic Performance of Students of a Tertiary Institute," American Journal of Educational Research, vol. 2, 2014, pp.713-726, doi:10.12691/education-2-9-3.

[5] Romero, C. and Ventura, S., "Educational data mining: A survey from 1995 to 2005," Expert Systems with Applications, vol. 33, Dec. 2007, pp. 135 - 146.

[6] Kabakchieva, D., "Predicting student performance by using data mining methods for classification, Cybernetics and Information Technologies," Cybernetics and Information Technologies, vol. 13, 2013, pp. 61 - 72, doi: 10.2478/cait-2013-0006.

[7] Norlida, B., Usamah, M. and Pauziah, M. A., "Educational data mining for prediction and classification of engineering students' achievement," IEEE 7th International Conference on Engineering Education (ICEED), 2015.

[8] Kachwala, T. and Sharma, L.K., "Comparative Study of supervised learning in customer relationship management," International Journal of Computer Engineering and Technology, vol. 8, 2017, pp. 77 - 82.

[9] Kaur, P., Singh, M. and Singh, G., "Classification and Prediction Based Data Mining Algorithms to Predict Slow Learners in Education Sector," Procedia Computer Science, vol. 57, 2015, pp. 500 - 508, doi: https://doi.org/10.1016/j.procs.2015.07.372.

[10] Kotwiantis, S. B., "Supervised machine learning: A review of classification," Informatica, vol. 31, 2007, pp. 249 - 268.

[11] Frank E., Hall M. A., and Witten I. H., "The WEKA Workbench. Online Appendix for Data Mining: Practical Machine Learning Tools and Techniques," Morgan Kaufmann, Fourth Edition, 2016.