Global Scientific JOURNALS

GSJ: Volume 9, Issue 4, April 2021, Online: ISSN 2320-9186
www.globalscientificjournal.com

LEARNER RECOMMENDATIO SYSTEM USING EDM FOR ELEMENTARY PUPIL

TIMES NEW ROMAN 24

Fatade, Oluwayemisi Boye[1], Kuyoro Afolashade(PhD)[2]

Computer Science Department,
Babcock University/College Name
fatade0306@pg.babcock.edu.ng

*Abstract***—** The use of data mining (DM) in education is a fast growing interdisciplinary research field which is also known as educational data mining (EDM). About 14% of children in the classroom between the ages of 4 to 12 are slow learner; this is a major set-back in the current education system which does not pay close attention to this category of learners especially in elementary schools. A well-designed educational system that will cater for slow learners can be achieved with the introduction of Education Data Mining. Hence this research focuses on implementing a Learner Recommender System that will detect pupil's learning status and recommend better learning techniques. Different factors that can influence pupil learning where obtained and applied on 8 classification algorithms using WEKA tool to build a predictive model.

The result shows that J48 performed well with 10-fold validation and hold-out as 66.7% and 67.5% respectively. J48, Reptree and Naïve Bayes performed well in the time taken to build the model but OneR and ZeroR had the optimal result. However, ZeroR has the lowest number of correctly classified instances. Hence, OneR can be said to perform optimally with respect to time taken to build the model. In comparison based on accuracy based on classifier and confusion matrices using 10-fold cross validation and hold-out j48 has optimal performance.

***Keywords*—Education data Mining, Recommender, Algorithm, validation, Slow Learner**

I. INTRODUCTION

In recent times, there have been significant advancement of information mining innovations to combat difficulties related to Knowledge Discovery Database through methods of clever mining in large databases. According to Gorunescu, (2011), Data Mining is predominantly centered on getting understanding of obscure valuable data from information. The utilization of DM in the field of Educational is referred to as Educational Data Mining (EDM). EDM accumulates crude information from educational sector and converts them into valuable data that has the capability of influencing training research (Parneet et al., 2015).

When a child is slow in learning academically or low in achieving academic skills, they are said to be slow learners (Priyamvada & Subodh, 2017). A few students are recognized as having uncommon necessities when they are at Early Year Stage; however, most youngsters are not distinguished until they enter the educational system (Haneesh et al., 2013).

There are different researches into how to improve the academic performance of high school students. With various factors affecting education system like the student performances, student family background, teaching techniques and more, the EDM analysis are performed on these factors and find out the factors which improve the higher educational system. The focus of this research is early detection of learning status for elementary pupil.

Classification techniques was used to analyze pupil family background and previous academic performance to predict the learning status of pupil as slow or fast learner and design a recommender system to recommend better study techniques in other to achieve a better education system that put all categories of learner into consideration.

A. Related Work

Mukesh Kumar, Shambhu & Aggarwal, (2016) worked on the concept Educational Data Mining and Knowledge Discovery. With the use of EDM, they identified slow learners in a classroom. This is done by analysing students' performance, using the best selected data mining technique. Scholars' performance data was collected and WEKA software was used in analysing the data in which high potential attributes were selected from the dataset for the implementation of EDM techniques using different attribute evaluators and search methods. Rankers were used to select the high potential attributes in the dataset, and a test was carried out on the dataset with analysis utilizing five different algorithms. Afterwards, a test on the entire algorithm was done. 3 algorithms – J48, SMO and ZeroR performed better than other algorithms.

Also, Deepa & Priya, (2016) worked on identifying slow learners using categorization DM methods. Their main concept is Prediction Process Model (PPM), with a focus on the model on discovering different indicators affecting the academic performance of the students. Their aim was to gain insight and compare different school slow learners, utilizing DM methods with existing data sources with J48 algorithm. With the use of accessible attributes, selection of topmost attributes is done and rebalancing of data is achieved with the use of cost sensitive classification.

Varsha, (2018) Naïve Bayes theorem was implemented in his work: Classification Technique for predicting Learning Behaviour in Students in Higher Education with the use of C# to predict categories of learners; Slow Learners, Average Learners and Fast Learners. This will enable instructors know the best way to teach each group of students for optimal performance. Naïve Bayes algorithm was implemented on pre-processed data set, using C# for stepwise implementation of the algorithm and predicting data for unknown record. The attributes include Gender, Area, SSC_Medium, SSC_Percentage, HSC_faculty, Math_At_HSC, Graduation_Marks, Admission_Type, Entrance_Rank, Parents' Income, Attendance, communication Skill, Learning-Behavior (Class Label). For database, Sqlserver was used for the imported data.

B. Data collection and Processing

Dataset was obtained purposively from a public elementary school, a total of 1730 datasets obtained and about 832 were used; this implies that about 75% of students enrolled in the school between 2014 and 2019. Identified background factors as shown in … together with pupils' grades were obtained from five basic subjects done at the elementary levels which include English, Mathematics, Cultural and Creative Arts (CCA), Pre-Vocational Studies (PVS) and Basic Science and Technology (BST). Tools that were used for data collection and sorting includes Microsoft Excel which was imported as ARFF into WEKA for processing.

C. Classification Algorithm Selection

Therefore, it is a usual practice in machine learning to engage the use of multiple models and determine the one that works for a particular problem (Cai, 2014). It is also a good practice to train models that work well using multiple algorithms. Hence, for this work, eight classification algorithms were chosen to train the models and these include Decision trees, Gradient Boosting, Naive Bayes, Multilayer Perception (MLP), Sequential Minimal Optimization (SMO), J48, OneR and ZeroR respectively.

D. Framework for Learner Status Recommender System

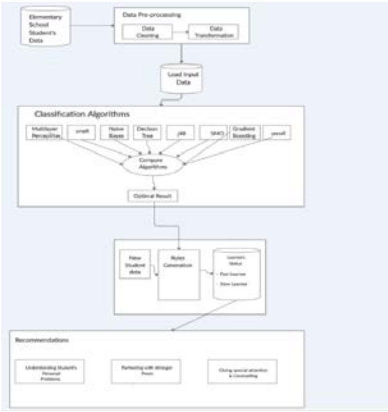

Fig.1 Framework for Learner Status Recommender System

The model validation was done utilizing the k-fold cross validation technique. The LRS would be in charge of mapping the patterns in the rules that would be generated from the optimally performing model with new pupils' information to predict the type of learners they are. The LRS is divided into three parts: rules generated from the optimal model from model building stage, data of new pupils and prediction of the type of learners. Rules are generated based on the pattern obtained from the optimal model from pupils' historical input data; the new pupils' data comprise of pupils' records that were not used during the development of the model. The LRS is in charge of mapping the pattern in the rules generated with the latest pupils' data to predict what type of learners they are. The Recommendation phase is in charge of recommending required steps to be taken on individual pupils based on the LRS prediction.

E. Results Of The Classification Models Performance

a. Performance of OneR

Table 1

OneR Algorithm Performance

Variable	10-fold	Hold-Out
Correctly classified instances	64.6032%	66.242%
Incorrect classified instances	35.3968 %	33.758%
Kappa statistics	0.2736	0.2898
Mean absolute error	0.354	0.3376
Root mean squared error	0.595	0.581

Performance of ZeroR

Table 2

ZeroR Algorithm Performance

Variable	10-fold	Hold-Out
Correctly classified instances	54.127%	54.7771%
Incorrect classified instances	45.873%	45.2229%
Kappa statistics	0	0
Mean absolute error	0.4966	0.4963
Root mean squared error	0.4983	0.4978

b. Performance of Naïve Bayes

Table 3

Naïve Bayes Algorithm Performance

Variable	10-fold	Hold-Out
Correctly classified instances	66.3492%	61.7834%
Incorrect classified instances	33.6508%	38.216%
Kappa statistics	0.3148	0.2286
Mean absolute error	0.3502	0.3617
Root mean squared error	0.5116	0.5233

c. Performance of Random Forest

Table 4

Random Forest Algorithm Performance

Variable	10-fold	Hold-Out
Correctly classified instances	63.0159%	61.7834%
Incorrect classified instances	36.9841%	38.2166%
Kappa statistics	0.2495	0.2267
Mean absolute error	0.4052	0.4009
Root mean squared error	0.5624	0.5569

d. Performance of Multiple Layer Perceptron (MLP)

Table 5

MLP Algorithm Performance

Variable	10-fold	Hold-Out
Correctly classified instances	61.9048 %	64.9682 %
Incorrect classified instances	38.0952 %	35.0318 %
Kappa statistics	0.2341	0.309
Mean absolute error	0.4051	0.3881
Root mean squared error	0.5373	0.5231

e. Performance of J48

Table 6

J48 Algorithm Performance

Variable	10-fold	Hold-Out
Correctly classified instances	64.2857 %	66.879 %
Incorrect classified instances	35.7143 %	33.121 %
Kappa statistics	0.2714	0.3298
Mean absolute error	0.3571	0.3312
Root mean squared error	0.5976	0.5755

f. Performance of RepTree

Table 4. 8

RepTree Algorithm Performance

Variable	10-fold	Hold-Out
Correctly classified instances	66.1905 %	65.6051%
Incorrect classified instances	33.8095 %	34.3949 %
Kappa statistics	0.3099	0.2901
Mean absolute error	0.414	0.407
Root mean squared error	0.4732	0.4642

i. Performance of SMO

Table 7

SMO Algorithm Performance

Variable	10-fold	Hold-Out
Correctly classified instances	66.6667 %	67.5159 %
Incorrect classified instances	33.3333 %	32.4841 %
Kappa statistics	0.322	0.3451
Mean absolute error	0.4109	0.4097
Root mean squared error	0.4752	0.4755

F. Comparison Of The Classification Models Performance

This section compares the performance of each model developed using measures which include classification accuracy, time taken to build the model and the confusion matrices.

G. Comparison based on classification accuracy

The outcome of both 10-fold cross validation and hold-out method is similar for all the classifiers J48 outperformed all other classifiers on both counts. Random Forest, RepTree, SMO, Naïve Bayes, OneR and Multilayer Perceptron performed well while ZeroR had the lowest classification accuracy for both hold-out and 10-fold cross validation with 54.7771% and 54.127 respectively.

H. Comparison based on time taken to build the models

The difference between the time taken to build Multilayer Perceptron compared with the other classification algorithms is wide. This means that in terms of computational resources usage, MLP used more resources and that is not acceptable. Based on the number of datasets used, SMO and Random Forest also used a considerable number of times. J48, RepTree and Naïve Bayes did not perform so badly while OneR and ZeroR had the optimal result in terms of time taken; however, ZeroR had already been decided to be unreliable based on its lowest number of correctly classified instances; hence it can be deduced that in terms of time taken, OneR performed optimally.

I. Comparison based on detailed accuracy and confusion matrices

This is the comprehensive accuracy level obtained by the eight classification algorithms. The performance of each algorithm based on the True Positive Rate (TPR), False Positive Rate (FPR), Precision, Recall, Receiving Operating Characteristics (ROC), F-measure, Mathews Correlation Coefficient (MCC) which is used to measure the quality of a 2-class or binary classification and Random Forest Classifier (RFC). The confusion matrices which show the numbers of both the correctly and incorrectly classified instances for all the identified algorithms based on classes, using both 10-fold cross validation and hold-out methods. Based on the results it is concluded that J48 is the model that performed optimally when compared to the other algorithms. Hence, it is concluded that J48, based on the result of this analysis, is a very good classifier for predicting elementary pupils' learner's status compared to the other algorithms used. It can thus be said that a system built based on this result is likely to perform efficiently and effectively

II. CONCLUSION

This work made use of different classification techniques to build models to predict slow learners in a classroom. These models were compared and the optimal model was used to develop a framework for a system that will predict and recommend best learning strategy for slow learners, in order to improve their academic performance. The development of the models was based on elementary students' variables obtained from an elementary public school of a local Government Area in Ibadan, Oyo state. Classification algorithms were carefully chosen. Eight (8) classification algorithms were used which are J48, RepTree, Naives Bayes, Multiple layer Perceptron (MLP), Sequential minimal optimization (SMO), OneR, Random Forest, ZeroR. In all the classifiers, it was concluded that the J48 decision tree performed optimally out of the 8 classifiers going by the result from accuracy. This shows that J48 performed well with 10-fold validation and hold-out as 66.7% and 67.5% respectively in comparison to time taken to build the model J48 together with OneR, SMO, REptree optimal performance in comparison. All these together with confusion matrices that measure the quality of a 2-class or binary classification and random forest classifier J48 model, were used for the prediction of elementary pupils' learning rate.

REFERENCES

Gorunescu, F. (2011). *Data mining: Concepts, models and techniques.* USA: Morgan Kaufmann.

Parneet, K., Manpreet, S., & Gurpreet, S. J. (2015). Classification and prediction-based data mining algorithms to predict slow learners in the education sector. *Procedia Computer Science, 57*, 500-508.

Priyamvada, S.,& Subodh, S. P. (2017). *Slow learners: A study of primary and middle school students of Raipur and Gariyaband District of Chhattisgarh.* 45(3), 65-78.

Haneesh, K., Krishnakumar, P., Sowmya, K. S., & Riyaz, A. (2013). Risk factors for scholastic backwardness in children. *Indian Pediatrics, 50*(7), 655-658.

Mukesh Kumar, K., Shambhu, S., & Aggarwal, P. (2016). Recognition of slow learners using classification data mining techniques. *Imperial Journal of Interdisciplinary Research, 2*(12), 741-747.

Deepa, S., & Priya, R. (n.d.). Identification of slow learners with categorization mining techniques. *International Journal of Pure and Applied Mathematics, 119*(10), 169-173.

Varsha, P. D. (2018). Classification technique for predicting learning behavior of students in higher education. *International Journal of Trend in Scientific Research and Development*, 4(2)163-166.

Cai, E. (2014). Machine learning of the day: The "no free lunch" theorem. Retrieved from www.chemicalstatistician.wordpress.com/machine-learning-of-the-day-The-no-free-launch-theorem