Educational Administration: Theory and Practice
2024, 30(7), 318- 323
ISSN: 2148-2403
https://kuey.net/

Research Article

Mathematical Applications And Modelling In Higher Education

Mrs. P. Sudha[1*], Mrs. R. Deepa[2], V. Anandkumar[3], Dr. Ravi kumar M[4], Dr. V. Hemalatha[5], Mr. A. Durai Ganesh[6],

[1*]Associate Professor of Mathematics, Excel Engineering College (Autonomous), Namakkal, Komarapalayam, Tamil Nadu. Email: sudhamathsecet@gmail.com
[2]Assistant Professor of Mathematics, Excel Engineering College (Autonomous), Namakkal, Komarapalayam, Tamil Nadu. Email: deepssengo@gmail.com
[3]Faculty Associate, Amrita School of Artificial intelligence, Amrita Vishwa Vidhyapeetham, Coimbatore, Tamil Nadu. Email: v_anandkumar@cb.amrita.edu
[4]Associate Professor, Department of Electronics and Communication Engineering, Cambridge Institute of Technology, Bengaluru, Karnataka. Email: mravikumar.vlsi@gmail.com
[5]Assistant Professorof Mathematics, Hindusthan Institute of Technology, Coimbatore,Tamil Nadu. Email: hema.varathan@gmail.com
[6]Assistant Professor of Mathematics, PET Engineering College, Vallioor, Tirunelveli, Tamil Nadu. Email: aduraiganesh25@gmail.com

Citation: Mrs. P. Sudha , et al (2024), Mathematical Applications And Modelling In Higher Education , *Educational Administration: Theory and Practice*. *30*(7), 318- 323
Doi: 10.53555/kuey.v30i7.6665

ARTICLE INFO

ABSTRACT

In today's educational landscape, student achievement rates are important indicators in institutional efficiency. Forecasting students' academic performance concluded data analysis is growth valuable for teachers and decision-makers. The purpose of this study is to grow and implement an advanced mathematical application & modeling technique to predict students' educational grades accurately using their online behaviors. To train our proposed model, we gathered a dataset that included 112 college students' final grades, as well as records of their online educational interaction and attendance at online lectures. Data normalization is used to pre-process the acquired information. We present an innovative Effective Remora optimization-driven Random Forest (ER-RF) mathematical-based technique for predicting students' educational grades. The established procedure is implemented using Python software. During the result analysis phase, we assess the efficacy of our model across various parameters. Additionally, we conduct comparative analyses with existing methodologies. The obtained findings demonstrate the efficacy and superiority of the suggested framework.

Keywords: *Mathematical Applications and Modelling, Higher Education, Effective Remora optimization-driven Random Forest (ER-RF)*

1. Introduction

Education plays a significant role in passing value, knowledge and essential skills in society to the future generation. Learning enables students to discover through different ways of being taught and also the ability to acquire knowledge in class, which is turned into fun by the instructors. Education in mathematics makes the student critical and innovative and also helps the student reason while solving problems. Education is a process that is active, dynamic and has a continuous nature particularly when it comes to mathematics. Attainments in mathematics education enable the learners to enhance their thinking, systematic, analytical and thorough, and learn to be objective and receptive during the handling of problems [1]. There is increasing evidence that mathematical modeling as a pedagogy may improve student engagement in science, technology, engineering, and mathematics (STEM) while also enhancing mathematics competency and self-efficacy. Emphasized the necessity of having a teaching framework that relates mathematics to the actual world by using real data and context, demonstrating mathematics' application in explaining the real world and developing effective mathematical modeling abilities [2].
These fields have garnered the interest of professionals and scholars as well as other interested stakeholders, as they seek to investigate these technologies grounded in data to improve the learning process and achievement, and address challenges facing higher learning institutions such as, drop out and retention rates.

For the sake of these goals, the use of predictive modeling proves to be rather efficient [3]. Higher education instructors are in charge of creating online content, interacting and discussing with students in primarily nonsynchronous scenarios, and involving online learners. In order to accommodate students who could attend on campus, academic staff at universities had to change their courses to be provided online [4]. The macro obstacles of teaching and learning during the pandemic, such as involving social, cultural and political concerns; nevertheless, a few small elements, such as the teachers' learning procedures while they are teaching online, were mostly disregarded [5].The problem tackled topresent an innovative mathematical approach, called ER-RF for precise prediction of students' educational grades.
The remaining section of thestudy is organisation:Part 2 covers the related study. Part 3 describes the proposed study approach, Part 4 shows the study outcomes, and Part 5 provides the research conclusions.

2. Literature review

Study [6] suggested the PBC model and BCEP framework to enhance e-learning performance prediction. The PBC model takes learning into the account, whereas the BCEP improves prediction accuracy through feature fusion and behavior classification. OULAD experiments outperform conventional approaches in terms of prediction performance. By utilizing machine learning algorithms to anticipate midterm test results, study [7] sought to determine final exam marks for undergraduate students. Six algorithms were evaluated such asNB, RF,SVM, KNN and LR. With an accuracy of 70–75%, the suggested approach highlights the opportunity for early at-risk student detection. Study [8] utilized MLR, NBand J48 DT, algorithms to predict students' success in current mathematics. Grades for attendance, quizzes, recitation, midterms, final exams and ratings were shown to have strong positive correlations. The Naïve Bayes method achieved the highest accuracy.
Study [9] used predictive modeling to analyze factors impacting student learning performance in MOOCs. It used RF for grade evaluation and GBM for final performance prediction, with an RSME of 8.131 and an average accuracy of 0.086, respectively. The research focuses on the influence of earlier performance and participation on subsequent assessments. Study [10] was to use a digital electronics education dataset to forecast student performance throughout online interactive sessions. It provides a prediction model based on 86 statistical variables divided into activity type, temporal statistics, & peripheral activity count. The RF classifier fared better than previous techniques in similar experimental conditions, achieving 97.4% classification accuracy. Study [11] improved educational institutions' reputations and rankings by taking practical initiatives with machine learning. The RF classifier, achieves 85% accuracy in grade forecasting and 83% accuracy in engagement prediction, focusing on student profiles and interaction variables.

3. Methodology

In the section, Effective Remora optimization-driven Random Forest (ER-RF) method is suggested to accurately predict students' educational grades. The dataset includes final grades for 112 college students. The data pre-processing using Z-score normalization. Figure 1 illustrates the suggested approach flow.

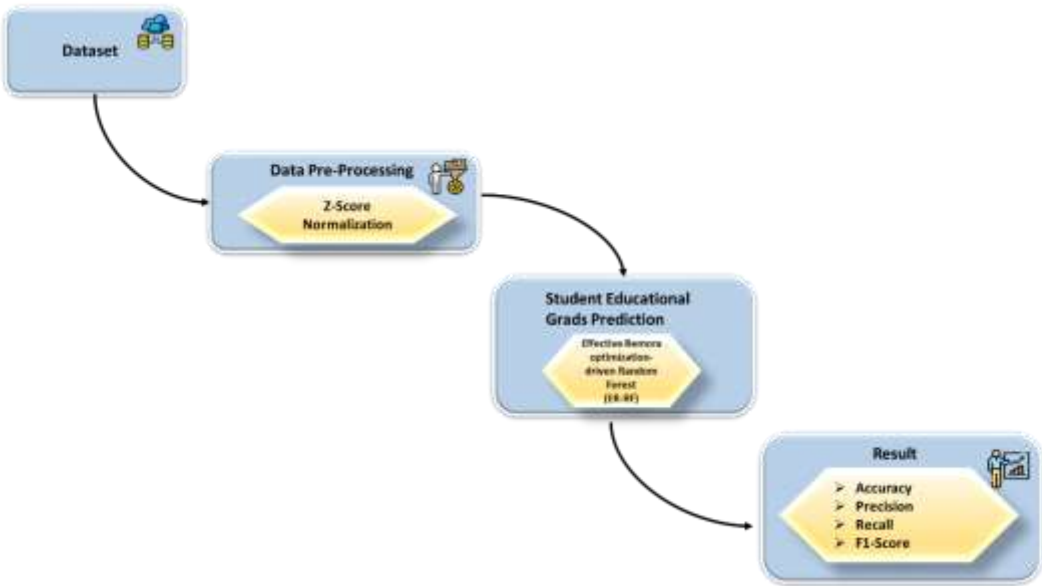

Figure 1: Flow diagram for proposed technique

3.1. Dataset

The collection consists of 7102 records that were taken from Moodle and include information about 112 students' course activities. It has functions including "Assignment (Uploaded/Submitted)," "Exercise (View)," "Lecture (View)," and "Source code (View)." To guarantee that records accurately depict authentic student involvement, cases in which the action on assignments was restricted to "View" were omitted, with an emphasis on actions that demonstrate active involvement or submission.

3.2. Data pre-processing

Z-score normalization, also referred as standardization, move average to zero and the standard deviation of the data set to one so that all the variables are placed on the same level of measurement, thus minimizing the effects of outliers in the modeling process and improving the prediction accuracy of students' educational grades using statistical modeling techniques. This method assists in identifying trends and outliers that are important for decision-making. The mean, as well as the standard deviation of each characteristic throughout a series of learning data, are used to normalize the vector of each characteristic included in the input data. For each attribute, the mean & standard deviation are calculated. The equality used in the approach is represented by the following equation: v', which displays the normalized data for the input factor v_i, its average value μ_i, and its standard deviation, equation (1).

$$v' = \frac{v_i - \mu_i}{\sigma_i} \qquad (1)$$

This approach standardizes the data set by setting the mean and normal deviation to zero and one, respectively. The technique begins by normalizing the feature vectors of the information collection. The average and standard deviation of each characteristic are calculated from the training data and utilized as weight in the ultimate system design.

3.3. Prediction of students' educational grades using Effective Remora optimization-driven Random Forest (ER-RF)

A hybrid strategy that combines Random Forest (RF) and the Effective Remora Optimization Algorithm (EROA) might be suggested to improve the prediction of students' academic results. By integrating Sailfish Optimization (SFO) with Levy flight for better global search, adaptive dynamic probability for maintaining the exploration/exploitation multiple and Restart Strategy (RS) to avoid local optima, EROA is considered as a refined metaheuristic algorithm. Meantime, RF is a strong ensemble learning method which applies DT to the classification task.

The hybrid approach might require adjusting the parameters of the random forest model like depth of trees, selected features, and the number of trees with the help of EROA. Introducing the random forest hyper parameters to EROA might make the exploration-exploitation forces enhance the resilience and accuracy of disease predictions. This has the following equation (2-4) representation.

$$X(t-1) = X_{best}(t) - (rand.\left(X_{best(t)} + X_{rand}(t)^2\right) - X_{rand}(t)) \qquad (2)$$

$$X_{att}(t+1) = X(t) + \left(X(t) - X_{pre}(t)\right).randn \ (3)$$

$$X(t+1) = Dist.e^{\propto}.cos\ cos\ (2\pi\alpha)\ + X(t)\ (4)$$

Where $X(t)$reflects the current position, $X_{best}(t)$is the greatest solution recognized thus far, $X_{rand}(t)$is a position decided at random, are random numbers of $rand$ and $randn$, $and\ Dist$ is the separation between one's actual position and ideal position.Then the random number is α. The hybrid approach can entail optimizing the random forest model's parameters using EROA, such as adjusting the number of trees (k), feature selection (m), and splitting criteria based on the Gini Index. The calculation of the Gini Index $(Gini)$ is as follows equation (5).

$$Gini\ (S) - \sum_{i=1}^{K} \left(\frac{|S_i|}{|S_i|}\right)^2 \qquad (5)$$

When$S_{\ i}$ represents the S partition that resulted.

4. Performance evaluation and discussion

In our experimental Python programming environment, we utilized a computer with a Core i7 processor from Intel (3.5 GHz) and 16 GB of RAM. We used the ER-RF method to predict students' academic outcomes. In this part, we compared the suggested approach to current methods. The performance evaluates various parameters such as F1 score %, recall%, precision%, and accuracy%. The existing methods are RF [12], XGB [12], and KNN [12].

Accuracy (ACC) is described as the percentage of the real that is true positive as well as true negatives. Precision (PRE) is calculated as the number of properly categorized samples that are positive divided by the entire amount of positive samples. It provides the accuracy evaluation of positive predictions. The accuracy and precision results are presented in Figure 2. Finally we compared the results of the proposed method with other methods.The proposed method ER-RF (accuracy (85%) and precision (86%)) is better than the existing methods such as RF-accuracy (78%) and precision (77%), XGB- accuracy (76%) and precision (76%) and KNN

- accuracy (72%) and precision (73%). The proposed ER-RF method demonstrates effective outcomes for predicting students' educational grades.

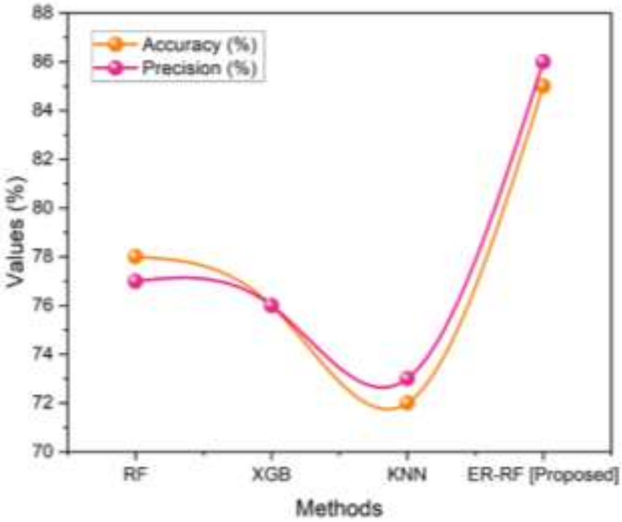

Figure 2: Outcome of accuracy and precision

Recall (R) is the proportion of properly predicted positive finds to all occurrences in the class. It assesses the amount of optimistic prognoses. F1 score is the measure that shows the harmonic mean of the accuracy and the recall values. It returns a single value that maximizes both accuracy and recall. The optimum value for the F1 metric is 1, while the lowest value is zero. Figure 3 shows the recall & F1 score results. Table 1 displays the comparison of the existing approach and the suggested method outcomes. When we compared our suggested ER-RF technique to other methodologies, we detected significant improvements in recall and F1 score. ER-RF achieved a recall of 83% and an F1 score of 84%, outperforming RF (78% recall, 77% F1-score), XGB (76% recall, 75% F1-score), and KNN (73% recall, 73% F1 score). These findings demonstrate the efficacy of ER-RF in predicting students' educational grades, demonstrating its better performance over traditional techniques.

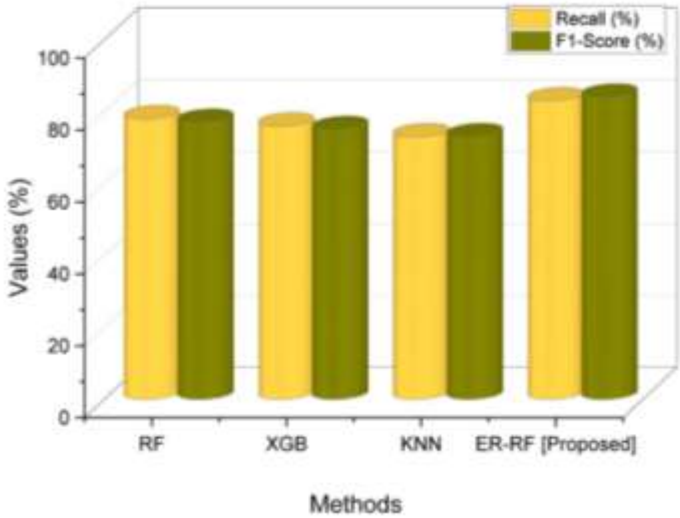

Figure 3: Outcome of recall and F1 score

Table 1: The suggested approach is compared to existing approaches

Methods	**ACC%**	**PRE%**	**R%**	**F1 score%**
RF	78	77	78	77
XGB	76	76	76	75
KNN	72	73	73	73
ER-RF [PROPOSED]	85	86	83	84

4.1. Discussion

Thus, everyapproachhas created its strengths and weaknesses, which may be vital when concerning the possibility of its usage for restoring educational grades' forecast. RF [12] could take a long to train especially with large data sets or complex models and thus has a restrictive scalability. Currently, it could be faced with class imbalance problems if proper methods of handling the imbalance are not employed leading to biased prediction results. XGB [12] is highly sensitive to hyperparameters like learning rate and tree depth implying that great optimization is needed. Without regularization, XGB models can overfit training data, resulting in poor generalization to untested data. KNN's [12] prediction time scales linearly with the quantity of the training data set, rendering it inefficient for big datasets. It could perform harshly if irrelevant features are not filtered away, reducing prediction accuracy.This study ER-RF method improves the prediction accuracy of random forest in educational grading by leveraging remora optimization. ER-RF enhances the reliability and efficiency of grade prediction models by improving feature selection and model parameters. This technique combines remora's adaptive search capabilities with random forest's ensemble learning, resulting in higher performance in managing different student data and improved prediction dependability in educational environments.

5. Conclusion

In modern educational settings, student success rates are crucial indicators of institutional efficacy. Predicting students' academic achievement via data analysis has been more beneficial to both teachers and administrators.The study attempts to innovate by creating and deploying a complex mathematical application and modelling approach for effectively forecasting students' educational grades according to their online customs.Using the final grades of 112 college students together with their records of interactions and attendance from an online course, the obtained dataset undergoes data normalization to prepare the data for the subsequent analysis. This paper's contribution is the Effective Remora Optimization-Driven Random Forest (ER-RF) algorithm to increase prediction precision. Our proposed method is simulated using Python software. The performance parameters include accuracy (85%), precisions (86%), recall (83%), and F1 score (84%) which are compared against previous methods as part of general result analysis and are supportive evidence for the proposed framework as a more effective predictor of students' educational outcomes.

5.1. Limitation and future scope

The Data in educational institutions is frequently inadequate, inconsistent, or biased, which reduces model accuracy and dependability. Future studies should prioritize improving data-gathering techniques and combining varied datasets to increase model accuracy & completeness.

Reference

1. Yuanita, P., Zulnaidi, H. and Zakaria, E., 2018. The effectiveness of Realistic Mathematics Education approach: The role of mathematical representation as a mediator between mathematical belief and problem-solving. PloS one, 13(9), p.e0204847.
2. Spooner, K., Nomani, J. and Cook, S., 2024. Improving high school students' perceptions of mathematics through a mathematical modeling course. Teaching Mathematics and its Applications: An International Journal of the IMA, 43(1), pp.38-50.
3. Sghir, N., Adadi, A. and Lahmer, M., 2023. Recent advances in Predictive Learning Analytics: A Decade systematic review (2012–2022). Education and information technologies, 28(7), pp.8299-8333.
4. Muir, T., Wang, I., Trimble, A., Bainbridge, C. and Douglas, T., 2022. Using interactive online pedagogical approaches to promote student engagement. Education Sciences, 12(6), p.415.
5. Huang, X., Lai, M.Y. and Huang, R., 2022. Teachers' changes when addressing the challenges in unexpected migration to online mathematics teaching during the COVID-19 pandemic: A case study in Shanghai. ZDM–Mathematics Education, 54(2), pp.359-372.
6. Qiu, F., Zhang, G., Sheng, X., Jiang, L., Zhu, L., Xiang, Q., Jiang, B. and Chen, P.K., 2022. Predicting students' performance in e-learning using learning process and behavior data. *Scientific Reports*, *12*(1), p.453.
7. Yağcı, M., 2022. Educational data mining: prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments*, *9*(1), p.11.
8. Morilla, R., Omabe, R., Tolibas, C.J., Cornillez Jr, E.E. and Treceñe, J.K., 2020. Application of machine learning algorithms in predicting the performance of students in mathematics in the modern world. TARAN-AWAN Journal of Educational Research and Technology Management, 1(1), pp.49-57.
9. Alshabandar, R., Hussain, A., Knight, R. and Khan, W., 2020, July. Students' performance prediction in online courses using machine learning algorithms. In 2020 International Joint Conference on Neural Networks (IJCNN) (pp. 1-7). IEEE.
10. Brahim, G.B., 2022. Predicting student performance from online engagement activities using novel statistical features. Arabian Journal for Science and Engineering, 47(8), pp.10225-10243.
11. Badal, Y.T. and Sungkur, R.K., 2023. Predictive modeling and analytics of students' grades using machine learning algorithms. Education and information technologies, 28(3), pp.3027-3057.
12. Gaftandzhieva, S., Talukder, A., Gohain, N., Hussain, S., Theodorou, P., Salal, Y.K. and Doneva, R., 2022. Exploring online activities to predict the final grade of students. Mathematics, 10 (20): 3758.

13. C Ramesh Kumar, Dr. Sabeena Kazi Dr Archana Shukla,v Dr. T. Muthukani Vairavel,Dr. Amol R Madane, Mr. A. Durai Ganesh (2024), Developing Deep Learning Models For Analyzing And Understanding Complex Graph Structures, Educational Administration: Theory and Practice, 30(6), 3946-3950 Doi: 10.53555/kuey.v30i6.6356

Appendix -1

Abbreviation	**Description**
PBC	process-behavior classification
BCEP	behavior classification-based e-learning performance
NB	Naïve Bayes
OULAD	Open University Learning Analytics Dataset
LR	logistic regression
MLR	Multiple logistic regression
SVM	support vector machine
RF	random forests
KNN	k-nearest neighbor
J48 DT	logistic regression
MOOCs	Massive open online courses
GBM	gradient boosting machine
XGB	extreme gradient boost