International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V71I7P233

Volume 71 Issue 7, 349-356, July 2023

Original Article

Predictive Models for Dropout Rates Affected by COVID-19 Using Classification and Feature Selection Techniques

Patchara Nasa-Ngium[1], Wongpanya S. Nuankaew[2], Kanakarn Phanniphong[3], Phaisarn Jeefoo[4], Pratya Nuankaew[5]

[1]Rajabhat Maha Sarakham University, Maha Sarakham, Thailand.
[2,4,5]University of Phayao, Phayao, Thailand
[3]Rajamangala University of Technology Tawan-ok, Bangkok, Thailand

[5]Corresponding Author : pratya.nu@up.ac.th

Received: 16 May 2023 Revised: 27 June 2023 Accepted: 29 June 2023 Published: 21 July 2023

Abstract *- The COVID-19 outbreak in Thailand has severely affected students in higher education. This research, therefore, had three significant objectives 1) to develop a model to predict the dropout rate and graduation rate of students in tertiary education, 2) to assess the effectiveness of the prediction model, and 3) to determine the contributing factors that affect the dropout rate and the graduation rate of students in higher education. This research highlighted the students' academic achievement in the Bachelor of Business Administration Program in Management program at the Faculty of Business Administration and Information Technology at the Rajamangala University of Technology Tawan-ok: Chakrabongse Bhuvanarth Campus, Bangkok, Thailand. The data collected was 547 students in this educational program during the academic year 2015 – 2022. It contains 14,402 transactions. The research tool used educational data mining development and applied machine learning techniques for factor analysis. The research results found that the COVID-19 pandemic had significant implications for the reduced graduation of students in the Business Administration Program in Management programs. It was also found that factors related to 21st-century skills influenced the student's termination.*

Keywords *- Academic achievement model, COVID-19, Dropout analysis, Educational data mining, Feature selection.*

1. Introduction

Educational Data Mining (EDM) analytics are gaining popularity and interest among educational and data science researchers. It is about taking the information that most educators ignore and analysing it with machine learning techniques and artificial intelligence. EDM [1]–[6] is a growing area of research in new educational technologies. It uses the essence of data mining in education to extract useful information about student behaviour in the learning process. EDM has been used in a wide range of fields, such as predicting student performance and detecting student behaviours [2], grouping similar materials or students based on their learning and interaction patterns [3], identifying relationships in learner behaviour patterns, and diagnosing student difficulties [4], identification of relationships among student behaviours and characteristics or contextual variables [5], interpretation of the structure and relations in collaborative activities and interactions with communication tools [6], and so on. EDM has become the basis for applying big data and machine learning techniques to improve education quality.

The global pandemic of Coronavirus 2019 (COVID-19) has forced many countries to close their cities. Thailand was also affected, which caused the educational process in Thailand to be halted. All universities and educational institutions are compulsory and must provide instruction on online platforms only, which affects students in many contexts. It became an impetus for researchers to study the impact on academic achievement of university students. The research problem is to find the critical factors that affect the predictive models and find suitable methods for students' dropout and pass rates in our dataset.

As a result, the research has three primary research goals and objectives. The first objective was to develop a predictive model for forecasting dropout and graduation rates of tertiary students. The second objective was to assess the effectiveness of the predictive model, and the third objective was to determine the factors affecting the dropout and graduation rates of students in higher. Research tools use data mining development processes and machine learning techniques. It consists of Decision Tree, Random Forest, Gradient Boosted Trees, Support Vector Machine, Feature

Selection techniques, a 10-fold cross-validation method, and confusion matrix techniques. Lastly, researchers firmly believe that knowing the factors that affect learning achievement will motivate them to manage further and promote quality learning.

2. Literature Reviews and Related Works

Education aims to develop learners' potential and support learners to develop a body of knowledge for sustainable use in their lives [30]. Although all countries and organizations are aware of and essential to this, the problem of non-graduation or student dropout has emerged in many research studies [8]–[11]. The impact of dropout burdens students and parents with the cost, waste of time, and opportunity lost, according to research by Nuankaew [8], [11].

The solution to dropout is to encourage learners to continue their daily life, known as Lifelong Learning, which has urged stakeholders to develop their organizations to become professional [12], [13]. But their models were not considered in the feature selection process for analysing the factors necessary [8]– [13]. Elevating research from pure science to cross-scientific integration has become increasingly popular. Part of the focus is on using artificial intelligence technologies and machine learning in conjunction with improving the quality of education. Feature scaling adjusted the boundary range of features' numeric data type to be in the same range and suitable for processing.

The principal component analysis was used to reduce the dimensionality of the data and compared multiple linear regression and support vector regression. The result showed that the best prediction model successfully had an accuracy of 83.44%. The researchers found a relationship between student behaviour attributes and academic performance [14]. Researchers applied statistical analysis and machine learning to see the students who dropped out and did not drop out in the first year. They used the correlation-based feature selection method to choose the crucial features, reduce the dimensionality of data, and minim overfitting. After, compared the five different base classifiers and analysed the predictive model dropout. The best result of the predictive was achieved with the Logistic Regression. The researchers found that they did not have an excellent only finding of dropouts, but their model improved performance after the first-semester result. They did not consider the differences between the students' programs [32].

Researchers used big data to analyse and synthesize learner behaviour and use learning achievement to design educational program recommendations [16], [17]. Researchers predict success or solve students dropping out with data mining techniques [9], [10], [11], [33]. Researchers construct learning paths and simulate student success with machine learning techniques [12], [14].

Two critical points involved and motivated this research. The first aspect is educational research on predicting academic achievement. Occasionally, it is referred to as learning achievement and academic success [19]. The main contributions of this section aim to classify student achievement, which occurs at all educational levels, for example, secondary school achievement predictions, university achievement predictions [13], [17], and analysis of learners' learning styles according to learning attitudes [20], [21]. These researchers aim to develop a model for preventing dropout among students based on factors such as academic performance, classroom environment, time spent in the curriculum, and context concerning the learner [22], [23].

The second aspect is to predict course achievement or course outcomes. A large part of this research group aimed to study learner behaviour using educational theory combined with appropriate analysis and design to monitor learners' learning behaviour [24]. The data used often included classroom activity data, including pre-test scores, post-test scores, group activity scores, midterm test scores, final exam scores, classroom participation scores, and so on. The nature of learner analysis is to group or cluster learners primarily to categorize and reinforce learning to meet learners' needs [25]. Their information was used for consideration with a few students, a few subjects, and not for a long time.

The previous work, coupled with the unusual situation affected by the COVID-19 pandemic in Thailand, has dramatically affected the students. Educational technologists and educators were trying to study problems that reflect the impact of COVID-19 on the Thai education system [34]. It reinforces that learners were affected, making researchers need to look at and research in more diverse dimensions. That is the meaning and importance of this research.

3. Materials and Methods

The research materials and methods of this research consist of three main parts: population and sample, research instrument design and selection, and factor analysis and interpretation.

3.1. Population and Sample

The research population was 547 undergraduate students who enrolled as students in the Bachelor of Business Administration Program in Management Program at the Faculty of Business Administration and Information Technology at the Rajamangala University of Technology Tawan-ok: Chakrabongse Bhuvanarth Campus, Bangkok, Thailand, from all students, and had academic achievement during the academic year 2015 – 2022.

3.2. Research Instrument Design and Selection

The research tool uses data mining and a machine learning technique called CRISP-DM: A Cross-Industry Standard Process for Data Mining. It consists of 6 parts as follows:

Table 1. Data collection

Students Status	Academic Year					
	2015 – 2018	2016 – 2019	2017 – 2020	2018 – 2021	2019 – 2022	Total
Termination	9	10	4	8	8	39
Resignation	5	6	2	6	4	23
Graduation	121	183	105	45	31	485
Grand Total	135	199	111	59	43	547

3.2.1. Business Understanding

The goal of Understanding Business is the study and understanding of research problems. This business understanding is to study the impact of the COVID-19 pandemic situation in Thailand on the education system and the student's academic achievement in the Bachelor of Business Administration Program in Management Program at the Faculty of Business Administration and Information Technology at the Rajamangala University of Technology Tawan-ok: Chakrabongse Bhuvanarth Campus, Bangkok Thailand.

Basically, the researchers found that 62 students who did not complete the course of study accounted for 11.33% of the total students in the program.

3.2.2. Data Understanding

Understanding the data is to manage and prepare the collection properly. The preliminary data analysis found three statuses of students in this study program: graduation, resignation, and termination of student status. The issue of student resignation and student termination is known as the "student drop-out problem" [8]. From the data gathered, the status of students in this educational program can be summarized as shown in Table 1.

Table 1 provides a summary of the data collected. It found that 23 students had resigned (representing 4.20%) and 39 had terminated (representing 7.13%). It seems that only a small number of students fail to achieve academic achievement, but the Rajamangala University of Technology Tawan-ok has an essential goal of ensuring that all students achieve academic achievement. Therefore, factors that affect learners' learning achievements must be identified.

Table 2 shows the dropout and pass rate data collected during the academic year 2015 – 2022. It found that the dropout rate grew explicitly from the academic year 2018 – 2021 and 2019 – 2022 during the COVID-19 pandemic.

Table 2. Dropout rate and pass rate

Status	Academic Year				
	2015 – 2018	2016 – 2019	2017 – 2020	2018 – 2021	2019 – 2022
Dropout	10.37%	8.04%	5.41%	23.73%	27.91%
Pass Rates	89.63%	91.96%	94.59%	76.27%	72.09%

3.2.3. Data Preparation

Data preparation is the most prolonged process and aims to prepare data for prototype development. The data provided in this section consists of student code, educational program code, educational program name, course code, course name, grades for each course, grade point average (GPA), grade mode, course unit, and student status. This process is cleaning, transforming, and selecting data to be appropriate for the modelling process.

The data collection consisted of 14,402 transaction records. The data used herein is standard and is not published to the damage to the Rajamangala University of Technology Tawan-ok and its owners.

3.2.4. Modeling

Modelling is the process of analysing data using data science or machine learning techniques. In this process, machine learning techniques were applied to model students' achievement predictions in the Bachelor of Business Administration Program in Management Program at the Faculty of Business Administration and Information Technology at the Rajamangala University of Technology Tawan-ok: Chakrabongse Bhuvanarth Campus, Bangkok, Thailand. The tools used at this stage belong under the category of predictive modelling, known as "supervised learning techniques" or "classification techniques". It comprises Decision Tree, Random Forest, Gradient Boosted Trees (GBT), and Support Vector Machine (SVM).

3.2.5. Evaluation

At this stage, the results analysed from the prototype model were considered. In the case of predictive analytics analysis, it usually uses a measure of accuracy as a criterion for determining the appropriate model.

The current predictive model evaluation test tool uses the 10-fold cross-validation method (90% of the data was used for the training data, and 10% was used for the testing data) and the Confusion-Matrix technique with seven key indicators: accuracy, precision, recall, F-measure, the area under the curve (AUC), sensitivity, and specificity.

The researchers determine the nature of the preparation and divide the data for testing the prototype into two parts: the first part is provided to develop the model called "training dataset". The rest were prepared to test the prototypes that have been developed as called "testing datasets". After testing the model in each technique, the prototype with the highest accuracy was selected.

3.2.6. Deployment

After analysing the data, the results must be interpreted for easy understanding or further use. This research uses relevant parts and elements of the model to analyse factors affecting learning achievement.

The technique and tool used for factor analysis is the feature selection technique for classificatory analysis, described in the next section.

3.3. Factor Analysis and Interpretation

Factor analysis and interpretation applied the methods and techniques of variable selection based on data mining techniques. It has become famous for extracting variables and multicomponent data, known as the "wrapper approach" [27]. In essence, there are two techniques for the wrapper approach: forward selection and backward elimination techniques. Both styles are the sequential selection algorithm, which is a method that selects a feature by considering the incrementing or decreasing attributes until an appropriate set of attributes is obtained.

The researchers used the forward selection technique to analyse the correlations between courses in the program and feature selection. The result is a prototype with increased accuracy attributes or variables that can be analysed to optimize the prototype further.

4. Research Results

The research report focuses on research objectives which consist of three main parts: a summary of data collection, an overview of model development, and the factors affecting university students' academic achievement.

4.1. A Summary of Data Collection

Research data was collected and summarized in Table 2. What it turns out to be questioning the researchers is why the number of graduates has declined. During the academic year 2018-2021, researchers found that 59 students graduated, 52 students (53.15 percent) decrease from the previous year.

Moreover, the graduation rate declined during the academic year 2019-2022, with 43 graduates. However, the Bachelor of Business Administration Program in Management Program graduate remained at a high-quality level with an average GPA of 2.80.

From the data collection shows the status of students in the Bachelor of Business Administration Program in Management Program in Table 1. Table 3 shows an overview of the student's academic achievement. It appears that the number of dropouts has increased, and the number of graduates has decreased. Therefore, the factors affecting the decrease in learning achievement should be studied.

The researchers used the forward selection technique to consider the attributes appropriate and found five courses from thirty-seven courses in this program. It included C_001201, C_205200, C_231252, C_231322 and C_231350 and was an important factor for the prediction model. This research presented predictive and analytical models for factor selection in the following sections.

4.2. Model Development

The parameter of the forward selections model was defined as the maximal number of ten attributes. The parameter setting configured to the classification models in the rapid miner studio program is shown in Table 4.

Table 3. Context of data collection

Academic Year	N. of students	Academic Achievement Results				
		Max	Min	Mean	Mode	Median
2015 – 2018	135	3.80	2.02	2.77	2.54	2.68
2016 – 2019	199	3.92	2.00	2.82	2.63	2.76
2017 – 2020	111	3.80	2.02	2.77	2.67	2.74
2018 – 2021	59	3.63	2.06	2.77	2.21	2.68
2019 – 2022	43	3.65	2.31	2.92	3.13	2.85
Average	547	3.92	2.00	2.80	2.67	2.74

Table 4. Parameter summary

Classifiers	Parameter Setting
DT	Criterion=gain ratio, maximal depth=10, apply pruning=true, confidence=0.5, apply prepruning=true, minimal gain=0.01, minimal leaf size=2, minimal size for split=4, number of prepruning alternative=3
RF	number of trees = 100, criterion = gain ratio, maximal depth =20, guess subset ratio = true, voting strategy = confidence vote
GBT	number of trees=50, maximal depth=10, min rows=10, min spilt improvement= 1.0E-5, learning rate=0.01, sample rate=1.0
SVM	kernel type=dot, kernel cache=200, convergence epsilon=0.001, max iterations=100000, scale=true, L pos=1.0, L neg=1.0, epsilon=0.0, epsilon plus=0.0, epsilon minus=0.0

Table 5. Comparison of model analysis results

Criterion	Classifier			
	DT	RF*	GBT	SVM
Accuracy	95.78% ± 2.68%	**97.73% ± 1.56%**	94.49% ± 3.42%	90.97% ± 2.94%
Classification Error	4.22% ± 2.68%	2.27% ± 1.57%	5.51% ± 3.42%	9.05% ± 2.94%
AUC	0.846 ± 0.18	0.942 ± 0.109	0.957 ± 0.065	0.738 ± 0.127
Precision	98.24% ± 1.86%	98.25% ± 1.81%	97.00% ± 3.24%	91.83% ± 1.62%
Recall	97.17% ± 2.25%	99.29% ± 1.48%	97.20% ± 4.31%	98.60% ± 3.27%
F-measure	97.68% ± 1.47%	98.77% ± 0.85%	96.99% ± 1.92%	95.01% ± 1.65%
Sensitivity	97.17% ± 2.25%	99.29% ± 1.48%	97.20% ± 4.31%	98.60% ± 3.27%
Specificity	81.67% ±19.95%	80.00% ± 21.94%	63.33% ± 39.13%	1.40% ± 98.02%
Time (s)	2	5	15	5

DT = Decision Tree, RF = Random Forest, GBT = Gradient Boosted Trees, SVM = Support Vector Machine, and ± = Standard Deviation (S.D.)

The model for predicting the learning achievement of learners in this study was summarized into two essential points: those who graduated and those who did not graduate due to termination of student status. In model selection, the researchers selected the models from the validity values, which summarized the validity values per model by various techniques, as shown in Table 5.

Table 5 shows that the Random Forest model has the highest accuracy (97.73%), precision (98.25%), recall (99.29%), F-measure (98.77%), sensitivity (99.29%), the lowest classification error (2.27%), lower the time processing than Gradient Boosted Trees, and the standard deviation of the accuracy, classification error, AUC, precision, recall, F-measure, and sensitivity values are less than in other models. It has little diffusion, which shows the high accuracy of this model. Although the Decision Tree uses the runtime processing less than Random Forest, all the standard deviation values of the Decision Tree have diffusion more than Random Forest.

The results of the best model performance analysis have the most recall in the graduation class and high precision for both classes. The specificity (80.00%) has lower in the termination class. It depends on the courses, which are important factors for predicting the model.

4.3. A Summary of Factors Analysis -Order Heading

The final objective of the research was to identify factors that significantly impacted student achievement, with the analysis and simulation, and significant factors for student termination are summarized in Table 6.

Table 6. Attribute weights of each classifier

Attribute (Course)	Weight			
	DT	RF	GBT	SVM
C_001201	0.145	0.220	79.235	0.042
C_205200	0.052	0.152	97.129	0.023
C_231252	0.168	0.253	15.231	-0.043
C_231322	0.350	0.177	0.159	0.008
C_231350	0.285	0.197	256.080	-0.105

Table 6 shows the courses that are significant to the student's termination. It consists of five courses: C_001201: Thai Language Skills, C_205200: English for Communication I, C_231252: Business System Analysis and Design, C_231322: JAVA Programming in Business, and C_231350: Database Development and Management. The five courses influence and relate to the predictive results of the termination and graduation groups and four classifiers. All five subjects correspond to 21st century learner skills that need attention to manage appropriation courses for decreasing student termination rate and increasing graduation rate.

5. Research Discussion

Research discussions focus on research objectives. It was found that the model developed with the Random Forest model was the most suitable as it had the highest accuracy (97.73%), high precision (98.29%), high F-measure (98.77%), the highest sensitivity (99.29%), lowest classification error (2.27%), low the standard deviation (± 1.56%) and low dispersion of data. Gradient Boosted Trees, Random Forest, and Decision Tree models provided high AUC values close to 1. The overall models had effectiveness for predictive dropout and pass rates. In addition, the forward selection to find the significantly impacted student achievement [8], [11]. We found the correlation of grades the course C_231252 was C and c_231322 was C+ which affected dropout and pass rate during the COVID-19 pandemic. The Support Vector Machine model provided a lower specificity value than other classifications. It cannot predict in the termination class, need more data training, or was inappropriate for this dataset, and has fewer attribute weights than other methods. The main concern is the amount of data on academic performance in some subjects that affect the number of dropout students. It may not be sufficient for the analysis during the COVID-19 pandemic.

Moreover, with the addition of feature selection techniques, the model has increased accuracy (97.73%) and lower classification error (2.27%), as shown in Table 5.

Besides that, this model showed even greater efficiency when tested with a confusion matrix. It summarizes the significant factors for the student's termination, as shown in Table 6.

Using data from the academic year 2015–2020 and 2018–2022, built proper models for drop-out learning patterns. If the model uses improper amounts of data, it will be less accurate because of less learning.

6. Conclusion

The dropout rate increase is a crucial problem in higher education during the COVID-19 pandemic. It impacts the number of students and pass rates of students that need to be solved of the problem. This research has achieved all research objectives using data mining and machine learning techniques. The three research objectives consist of (1) developing an academic achievement prediction model for university students, (2) assessing an academic achievement prediction model for university students, and (3) analysing factors affecting the academic achievement of university students.

The research achieved the first objective by developing a model by comparing the accuracy, classification error, and other elements. It was found that the most stable model was the model developed with the Random Forest model, with the highest accuracy, the lowest classification error, and a low run time processing. It has tested the model's performance corresponding to the second objective. Courses are factors to influence the specificity in the termination class that impacts the number of false positives (FP) and the specificity value. For the last objective, the research concluded the factors that are significant for the termination of the student status using the forward selection technique, which consists of five courses based on attribute weight value: C_001201: Thai Language Skills, C_205200: English for Communication I, C_231252: Business System Analysis and Design, C_231322: JAVA Programming in Business, and C_231350: Database Development and Management. It was useful to analyse students who are expected to drop out because of poor grades in courses, which was an important factor that affected the dropout and pass rates of students in programs during the COVID-19 pandemic. The program executive committee and the instructor will plan to improve their courses to decrease the number of student dropouts. The researcher found the best predictive model because it properly determined the parameter, analysed the factor, and applied the feature selection technique appropriately.

From the conclusions and findings of the research, the researchers can conclude that the COVID-19 pandemic in Thailand directly impacts the Thai education system. What was reflected was a decrease in the number of graduates, as shown in Table I. It emerged 59 graduates (representing 10.79%) in the academic year 2018 – 2021 and 43 graduates (representing 7.86%) in the academic year 2019 – 2022. It was a much lower number than the previous academic year. In addition, from the findings of the four courses related to the termination of students, of which four skills correspond to the skills of learners in the 21st century, educational institutions and related personnel should be aware of and emphasize the solution to this problem.

For future work, we plan to study the activity learning and scores of the subjects which affect students' dropout. It is divided before and after the COVID-19 pandemic groups to compare the difference between students in both groups [28].

Funding Statement

This research project was supported by the Thailand Science Research and Innovation Fund and the University of Phayao (Grant No. FF66-UoE002)

Acknowledgments

This research was supported by many advisors, academics, researchers, students, and staff from the University of Phayao and the Rajamangala University of Technology Tawan-ok, Thailand. The authors would like to thank all of them for their support and collaboration in making this research possible.

References

[1] Laura Calvet Liñán, and Ángel Alejandro Juan Pérez, "Educational Data Mining and Learning Analytics: Differences, Similarities, and Time Evolution," *RUSC, Universities and Knowledge Society Journal*, vol. 12, no. 3, pp. 98-112, 2015.[CrossRef] [Google Scholar] [Publisher Link]

[2] Surjeet Kumar Yadav, and Saurabh Pal, "Data Mining: A Prediction for Performance Improvement of Engineering Students using Classification," *Machine Learning, arXiv*, vol. 2, no. 2, pp. 51-56, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[3] Pavlo D. Antonenko, Serkan Toy, and Dale S. Niederhauser, "Using Cluster Analysis for Data Mining in Educational Technology Research," *Educational Technology Research and Development*, vol. 60, no. 3, pp. 383–398, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[4] John S Kinnebrew, and Biswas Gautam, "Identifying Learning Behaviors by Contextualizing Differential Sequence Mining with Action Features and Performance Evolution," *International Educational Data Mining Society*, 2012. [Google Scholar] [Publisher Link]

[5] Hogyeong Jeong, and Gautam Biswas, "Mining Student Behavior Models in Learning-by Teaching Environments," *Educational Data Mining*, pp. 127–136, 2008. [Google Scholar] [Publisher Link]

[6] Camilo Palazuelos, Diego García-Saiz, and Marta Zorrilla, "Social Network Analysis and Data Mining: An Application to the E-Learning Context," *Computational Collective Intelligence, Technologies and Applications*, pp. 651–660, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[7] Putu Eka Juliana Jaya, "Increasing Learning Achievement and Student Participation through Independent Learning Design," *SSRG International Journal of Economics and Management Studies*, vol. 8, no. 10, pp. 50-56, 2021. [CrossRef] [Publisher Link]

[8] Pratya Nuankaew, "Dropout Situation of Business Computer Students, University of Phayao," *International Journal of Emerging Technologies in Learning*, vol. 14, no. 19, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[9] Concepción Burgos, "Data Mining for Modeling Students' Performance: A Tutoring Action Plan to Prevent Academic Dropout," *Computers & Electrical Engineering*, vol. 66, pp. 541–556, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[10] Marzieh Karimi-Haghighi, Carlos Castillo, and Davinia Hernández-Leo, "A Causal Inference Study on the Effects of First Year Workload on the Dropout Rate of Undergraduates," *Artificial Intelligence in Education*, pp. 15–27, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[11] Pratya Nuankaew et al., "Analysis Dropout Situation of Business Computer Students at University of Phayao," *The Impact of the 4th Industrial Revolution on Engineering Education*, pp. 419–432, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[12] Sittichai Bussaman et al., "Prediction Models of Learning Strategies and Learning Achievement for Lifelong Learning," *IEEE 6th International Conference on Teaching, Assessment, and Learning for Engineering*, pp. 192–197, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[13] Wongpanya Nuankaew, and Pratya Nuankaew, "Educational Engineering for Models of Academic Success in Thai Universities during the COVID-19 Pandemic: Learning Strategies for Lifelong Learning," *International Journal of Engineering Pedagogy*, vol. 11, no. 4, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[14] Pranav Dabhade et al., "Educational Data Mining for Predicting Students' Academic Performance using Machine Learning Algorithms," *Materials Today: Proceedings*, vol. 47, pp. 5260–5267, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[15] Paniah Emeko Kodzo, "The Determinants of Educational Performance in Togo," *SSRG International Journal of Economics and Management Studies*, vol. 9, no. 11, pp. 6-16, 2022. [CrossRef] [Publisher Link]

[16] Diego A. Forero et al., "Global Perspectives and Recommendations for Curriculum Design in Academic Programs in the Health Sciences," *Educación Médica*, vol. 23, no. 2, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[17] Wongpanya Nuankaew, and Pratya Nuankaew, "The Study of the Factors and Development of Educational Model: The Relationship between the Learner Context and the Curriculum Context in Higher Education," *International Journal of Emerging Technologies in Learning*, vol. 14, no. 21, pp. 205–226, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[18] Fuseini Inusah et al., "Data Mining and Visualisation of Basic Educational Resources for Quality Education," *International Journal of Engineering Trends and Technology*, vol. 70, no. 12, pp. 296-307, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] Rasha M. Abdelrahman, "Metacognitive Awareness and Academic Motivation and their Impact on Academic Achievement of Ajman University Students," *Heliyon*, vol. 6, no. 9, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[20] Pratya Nuankaew et al., "Perception and Attitude toward Self-Regulated Learning in Educational Data Mining," *6th International Conference on Technical Education*, pp. 1–5, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[21] Elham Hussein et al., "Exploring Undergraduate Students' Attitudes towards Emergency Online Learning during COVID-19: A Case from the UAE," *Children and Youth Services Review*, vol. 119, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[22] Myunghee Kang et al., "Developing an Educational Performance Indicator for New Millennium Learners," *Journal of Research on Technology in Education*, vol. 43, no. 2, pp. 157–170, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[23] Kanakarn Phanniphong et al., "Clustering of Learners Performance based on Learning Outcomes for Finding Significant Courses," *Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering*, pp. 192–196, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[24] Élise Lavoué et al., "Analyzing the Relationships between Learners' Motivation and Observable Engaged Behaviors in a Gamified Learning Environment," *International Journal of Human-Computer Studies*, vol. 154, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[25] Lilah F. Morris-Wiseman, and Valentine N. Nfonsam, "Early Detection and Remediation of Problem Learners," *Surgical Clinics of North America*, vol. 101, no. 4, pp. 611–624, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[26] Suman, "Analysis on the Mis in Primary Schools," *International Journal of Computer and Organization Trends*, vol. 11, no. 1, pp. 10-14, 2021. [Publisher Link]

[27] Gang Chen, and Jin Chen, "A Novel Wrapper Method for Feature Selection and its Applications," *Neurocomputing*, vol. 159, pp. 219–226, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[28] Gazi Mahabubul Alam, "Does Online Technology Provide Sustainable HE or Aggravate Diploma Disease? Evidence from Bangladesh—a Comparison of Conditions before and during COVID-19," *Technology in Society*, vol. 66, pp. 1–10, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[29] B.Ajay Babu, and K. Nimala, "A Study on Techniques for Online Feature Choice Method," *SSRG International Journal of Computer Science and Engineering*, vol. 2, no. 3, pp. 1-5, 2015. [CrossRef] [Publisher Link]

[30] Cristina Hava Muntean, "Improving Learner Quality of Experience by Content Adaptation Based on Network Conditions," *Computers in Human Behavior*, vol. 24, no. 4, pp. 1452–1472, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[31] Mitali Sonowal, and Mun Kalita, "Metacognitive Awareness and Academic Achievement of Higher Secondary Level Arts Stream Students of Dibrugarh District, Assam," *SSRG International Journal of Humanities and Social Science*, vol. 6, no. 4, pp. 17-23, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[32] Marina Segura, Jorge Mello, and Adolfo Hernández, "Machine Learning Prediction of University Student Dropout: Does Preference Play a Key Role?," *Mathematics*, vol. 10, no. 18, p. 3359, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[33] Joana R. Casanova et al., "Dimensionality and Reliability of a Screening Instrument for Students At-Risk of Dropping out From Higher Education," *Studies in Educational Evaluation*, vol. 68, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[34] Sinsuda Dechsupa, and Suvichada Assawakosri, "Positive Impact of Lockdown on COVID-19 Outbreak in Thailand," *Travel Medicine and Infectious Disease*, vol. 36, 2020. [CrossRef] [Google Scholar] [Publisher Link]