Journal of Engineering Science and Technology
Special Issue on SIMASE, August (2021) 30 - 39

MONITORING LEARNERS' PERFORMANCE BY MODELING LEARNING PROGRESS USING MACHINE LEARNING

RIA ARAFIYAH[1,2,*], HARRY BUDI SANTOSO[1], ZAINAL A. HASIBUAN[3]

[1]Faculty of Computer Science, Universitas Indonesia,
Depok, Indonesia
[2]Computer Science Department, Universitas Negeri Jakarta,
Jakarta, Indonesia
[3]Faculty of Computer Science, Universitas Dian Nuswantoro,
Semarang, Indonesia
*Corresponding Author: ria.arafiyah@ui.ac.id

Abstract

In this globalization era, the involvement of machines is inevitable, including one within the context of monitoring learners' performance. This study investigates learners' performance monitoring to predict learning progress using Machine Learning techniques. This study uses the results of a one-semester assessment of Indonesian junior high school students in mathematics. Aspects of the assessment include cognitive and psychomotor domains while the forms of assessment comprise formative and summative. This study shows that learning progress modeling can also be carried out throughout learning and model accuracy is fairly good in four learning periods. The learning progress model generated in the last period of the course is better than that in the early period of the course. Out of 7 machine learning algorithms being compared, Random Forest shows the best accuracy as it reaches 92% in the whole period. It is indicated that the model being used in this study is proven to be effective in one subject being taught in one semester.

Keywords: Learning modeling, Learning progress, Machine learning, Monitoring learners' performance, Predicting learning progress.

1. Introduction

Monitoring is necessary in a variety of fields in order to achieve goals as well as to avoid failure. In the environmental field, for instance, monitoring the environment from air pollution is conducted to reduce system power consumption [1]; monitoring carbon dioxide levels in closed houses for optimization growth of broiler chickens [2]; and monitoring temperature and humidity is used to improve the performance of baby incubators [3]. In the meantime, in education, monitoring basic skills in reading, writing, and mathematics is used to determine the learning progressions of children [4]. In Indonesia, not all students go to class when recorded at the end of the year. This causes to the rate of staying in the same grade; one of the cases is that 5.34% of them stay in grade 8 and 4.44% of them stay in grade 9 [5]. To reduce this number of student unsuccessfulness, monitoring learners' performance is needed. Monitoring learners' performance can be used to avoid learner failure. Ideally, learners' performance is carried out by observing the learning process of each learner and conducting periodic assessments (evaluations). However, it is often constrained by a large number of students and limited time. Education data are very large (volume), varied, and scattered everywhere starting from the value data recorded by teachers, the background of students, and others [6]. This data can be used to overcome the problem of monitoring learners' performance by using various methods and techniques in information technology.

Monitoring learners is very necessary to prevent unsuccessful learning [7], as well as to support and make improvements to the learning process [8]. The limitation of teachers to monitor learners' performance continuously is the main reason for the need for automated learner monitoring. Furthermore, the ability to predict learning progress will also greatly help teachers so that they have more time to manage classes and achieve learning goals. Learning progress modeling can be used to automate learner monitoring so that it can be a solution to avoid learners' unsuccessfulness. Mathematical model of a phenomenon is to transfer the phenomenon into the language of mathematics. A valid variable is needed to construct it. The mathematical model related to learning achievement is the cups model. In order to create a mathematical model, valid variables such as such as Intelligence, Emotion, and Adversity taken from valid psychological measuring instruments, such as instruments IST (Intelligent Structure Test), Pauli test, and EPPS (Edwards Personal Preference Schedule) Test [9]. To get a statistical model, it must meet several requirements such as the number and method of sampling and fulfil certain assumptions. In making the academic achievement prediction model, a stratified random sample technique is used and must meet the normality test and others to get a multiple regression model from the independent variables of motivation and attitude towards school [10]. To take advantage of large Education data without the hassle of testing several assumptions, machine learning methods can be used that can automatically detect patterns in the data, then use those patterns to predict future data, or to perform other types of decision making under uncertainty [11].

The model produced by machine learning can be used to accurately predict academic achievement [12]. Learner progress prediction to prevent student failure can use linear and non-linear classification algorithms [13], random forest [7], Linear Regression [14], decision tree [15], and single and ensemble classifiers [16]. Formative tests are used to monitor learning progress [17]. Other ways to monitor the cognitive status of learning include using brain sensors [18] or a predictive model. This study proposed to monitor the progress of learners by making models with machine learning techniques. The updates of this research study are formative and summative tests on cognitive and psychomotor aspects that can be used to model learning progress in several learning situations, so that learner failure can be avoided beforehand. To show that monitoring can be carried out as early as possible and as often as possible, the prediction model of student progress was carried out in 4 periods throughout the course. To get the best model, several single algorithms are used, like Logistics Regression (LogReg), Decision Tree (DT), Random Forest (RF), K-Neural Network (KNN), Linear Discriminator Analysis (LDA), Gaussian Naïve Bayes (GNB), and SVM.

The objective of this study is to show that the success of learners can be predicted by knowing the results of the tests that have been carried out previously and the psychomotor aspect is quite influential on cognitive assessment and can be used to predict student graduation/failure. Machine learning techniques will be used to make predictive models of learning progress and measure the performance of the resulting models. The research question that will be answered is whether it is possible to model learning progress using only the results of the previous assessment. Can learning progress be predicted as early as possible? So that there is enough time to avoid failure and make improvements for students and teachers? Is there a relationship between aspects of cognitive and psychomotor assessment with learning outcomes? Models are developed using an algorithmic technique in machine learning [11]. The model generated from the algorithm in each period was compared for accuracy to choose the best one based on the level of accuracy using training and testing data.

2. Methods

This study uses research data derived from the results of teacher assessments for one semester, in a formative and summative manner which is used to determine student graduation. Formative assessment is carried out after the teacher completes one competency in both cognitive and psychomotor aspects. While the summative assessment is carried out in the middle of the semester and at the end of the semester. The course material tested on the formative test is one competency that has been studied, while the summative test is the overall competency that students must achieve to pass the mathematics course.

Participants in this research are students at private junior high schools in South Jakarta, Indonesia. The number of parallel classes at each level sequentially is 6, 7, and 6 classes or totalling 19 classes. These classes are followed by totalling 659 students. Participants were students of mathematics courses in classes.

Collecting data using an instrument made by the teacher according to the standards set by the Indonesian government. The formative assessment instrument or daily assessment is made by the course instructor while the summative assessment instrument is made by several teachers teaching the same course. Aspects of formative assessment include cognitive and psychomotor, while summative only cognitive aspects. Students who get scores below the lowest allowable score limit or commonly termed KKM (Kriteria Ketuntasan Minimum or minimum completeness criteria) must take remedial tests until their scores are above the KKM. The value of students who take the remedial test is the KKM score. In this research, each level was taught by different teachers with different course materials. So the assessment instruments for levels 7, 8, and 9 are also different.

During one semester, summative assessments were carried out 2 times on cognitive aspects, namely the mid-test (SC1) and the final test (SC2). Formative assessment is carried out 5 times on the cognitive aspect (FC1, .., FC5) and 4 times on the psychomotor aspect (FP1, .., FP4). In the psychomotor aspect, no summative assessment was carried out. If a student gets a formative or summative test score on the cognitive aspect under the KKM, then he must take a remedial cognitive formative test (RE) or a cognitive summative test (RS). The final score on the cognitive and psychomotor aspects is the average of the results of all the assessments that have been carried out. The description of the time of taking values is shown in Fig. 1.

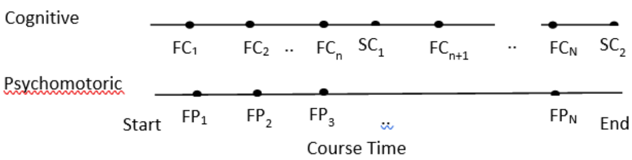

Fig. 1. Formative and summative assessment times on cognitive and psychomotor aspects.

The coding of the type of assessment uses two capitals: F being a formative assessment, S representing a summative assessment, C for a cognitive aspect, and P for a psychomotor aspect. The index shows the rating counter. Formative assessment of cognitive and psychomotor aspects can be done in the same week or even time. Data collection was carried out at all levels of junior high school (7, 8, and 9) for 1 semester (odd 2019) in Mathematics. The data collected are unified in the dataset in Table 1.

Table 1. Dataset description.

No.	Attribute Name	Description	Type	Range
1	Name	Student's name	nominal	-
2	Course_code	Code of level and course	ordinal	{7,8,9}
3	FC1, .., FC5	The results of the formative test on the cognitive aspect (daily test)	numeric	[0..100]
4	RE1,.., RE5	Taking remedial for the cognitive aspect of formative tests, the score is 1 if students have to take remedial, 0 if not.	boolean	{1,0}
5	SC1, .., SC2	The result of the summative test on the cognitive aspect (mid-test, final test)	numeric	[0..100]
6	RS1, RS2	Taking remedial for the cognitive aspect of summative tests, the score is 1 if students have to take remedial, 0 if not.	boolean	{1,0}
7	TC	The Sum of the result is on the cognitive aspect	numeric	[0..100]
8	GC	course grade	ordinal	{A, B, C, D}
9	FP1, .., FP4	The results of the formative test on the psychomotor aspect (daily test)	numeric	[0..100]
10	TP	Sum of the results test on psychomotor aspect	numeric	[0..100]
11	GP	psychomotor grade	ordinal	{A, B, C, D}

After the teacher teaches one competency, a formative test is carried out, during one semester the number of formative tests given to each student at each level is different depending on the number of competencies, the length of teaching, and the teaching conditions. In the data taken, FC was carried out varied from 4 to 10 in each level. The summative test (SC) was carried out 2 times, namely in the middle and at the end of the semester. Assessment of the psychomotor aspect is carried out alternately or simultaneously with cognitive assessment. FP was carried out between 4 to 8 times and no summative test was carried out on the psychomotor aspect. Next, five FC and FP data were selected, based on the data that filled the most numbers. Missing value on FC and FP is filled in based on the previous value.

Total cognitive scores (TC) and total psychomotor scores (TP) were derived from the average test scores. GP is determined from the TC value, at each level the KKM value is different, so the range of GP determination is also different. Therefore, TC and GP are included differently. In Indonesia, the KKM (Minimum eligibility criteria) is applied or the lowest score that must be achieved to pass the course. The KKM is determined by the school based on the results of the deliberation of course teachers.

3. Results and Discussion

To understand the data relating to the learning progress of students in learning mathematics, it is carried out exploring the data using summary statistics and data visualization. A description of the participants' mathematical grades at each level can be seen in Fig. 2.

In Fig. 2, the number of students who got the least A grades at each level. Most students got B grades except at level 8 most students got B grades. In terms of the number of students who got certain grades at each level, they were almost the same.

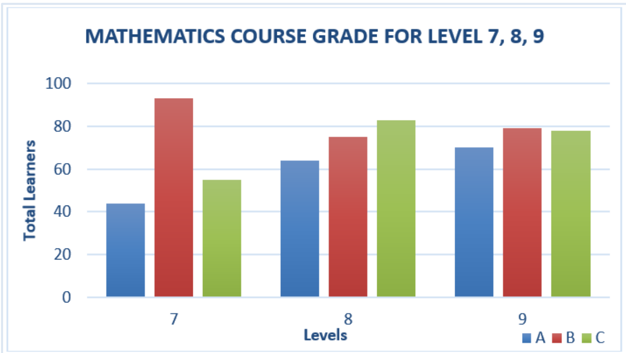

Fig. 2. Learning progress graph, x-axes represent levels, y-axis total learners.

During class learning, the teacher uses formative assessment to assess the learning process. Assessment on cognitive and psychomotor aspects, each done 4 times, before or after the midtest. It is not known whether the material content relationship between the two tests is whether TF1 is a concept and TP1 is a skill using TF1 concepts or something else. By using the Pearson correlation test, the relationship between two features is known. From Fig. 3 it can be seen the correlation between cognitive and psychomotor formative assessments. FC1 has the closest relationship with FP2 and FP4. FC2 has the closest relationship with FP1 and FP4. FC3 has the closest relationship with FP2. FC4 has the closest relationship with FP4. By knowing this correlation, the teacher can make improvements to the content of the cognitive test material related to the content of the psychomotor test material.

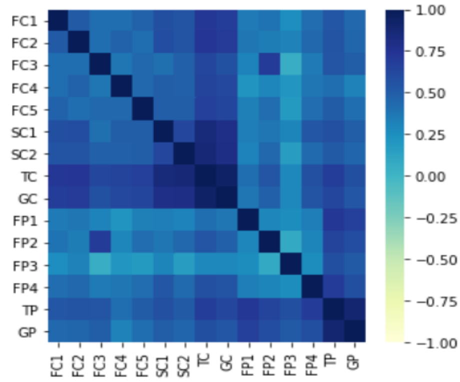

Fig. 3. Correlations between 2 feature.

The data for each feature is fairly evenly distributed for each feature. The FC3 and FP3 variables have the highest scores, meaning that students have good enough scores for this assessment. Outliers only exist in FC3 and FP4 features (see Fig. 4).

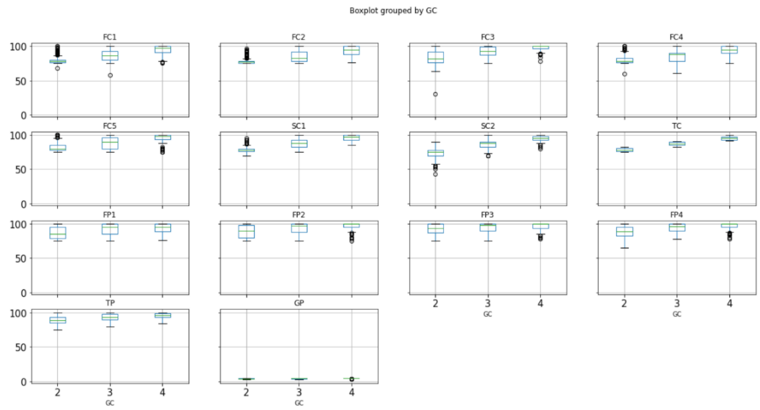

Fig. 4. Feature boxplot by grade.

In classroom learning, the teacher monitors by observing student activities in learning and by giving tests. Observations of student activities are difficult to describe so they are usually not documented or stored by the teacher for further use. Teachers usually use tests, either in the form of formative or summative tests. The results of this test are stored and used to decide whether the student passed the course or not. Based on the test results, the teacher provides feedback to improve teaching. Monitoring in the form of tests can be explored to be used in capturing learning performance. This research data is the result of tests on mathematics courses at levels 7, 8, and 9 (see Fig. 2). There are 24 features for 642 sample students in which the matrix is scattered as shown in Fig. 5.

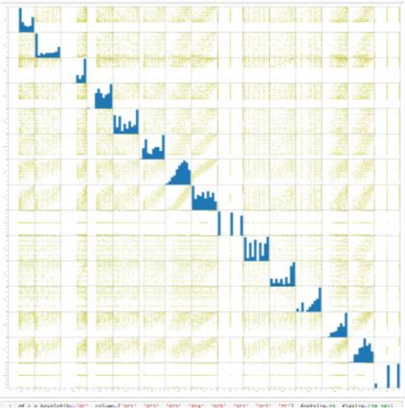

Fig. 5. Scattered matrix.

As shown in Fig. 3, there is a correlation between the results of cognitive and psychomotor formative assessments with the final score. The correlation is quite good between grades with the Formative test of cognitive aspect (FC1, FC2, FC3), and the Formative test of psychomotor aspect (FP1, FP2, and FP4) with grade. Midtest scores have a higher correlation to grade than final test scores. The total psychomotor assessment has a higher correlation with grades than the total cognitive assessment. So that it can be used as a grade improvement suggestion for teachers and students that grade improvement is better done before the middle of the semester. Likewise, efforts to improve psychomotor abilities will further increase grades compared to increasing cognitive abilities. This becomes very reasonable because good psychomotor abilities are built on good cognitive abilities [19]. The results obtained from this study, using psychomotor assessment data, the model obtained has better accuracy and can even reach 100%. while the model that only uses cognitive assessment only achieves 95% accuracy [20].

Machine learning can be used to create predictive models, based on previously obtained data. The fundamental purpose of machine learning is to generalize or induce unknown rules from examples of rule applications [21].

To show that monitoring can be done at any time, the modeling is divided into several periods for one semester. In this study, modeling is carried out in 4 scenarios that reflect the learning period. The distribution of the period and its variables is shown in Table 2. The first scenario is carried out after the semester, the second scenario is carried out after the mid-test, the third scenario is carried out after the semester, and the fourth scenario is after the end of learning. From the 4 scenarios, 4 prediction models were generated. Calculation of model accuracy uses 70% of training data and 30% of testing data. Summary of model accuracy with training and testing data using Logistics Regression (LogReg) algorithms, Decision Tree (DT), Random Forest (RF), K-Neural Network (KNN), Linear Discriminant Analysis (LDA), Gaussian Naïve Bayes (GNB), and SVM are shown in Table 2.

From Table 2, it can be seen that the accuracy of the scenario 4 model is the best compared to the previous scenario. This shows that the closer to the end of the program and the more features related to the target, the better the accuracy of the model. This is a natural thing. The next question is, can we predict learning progress as early as possible? The answer is possible. By using data recorded by the teacher in class. It is best to look for other features that will increase accuracy [22], including features related to internal and external factors that affect student achievements such as learning logs, student discussion activities, student backgrounds, institutions, and curriculum [23]. In a comparison of seven models generated from machine learning algorithms, Random Forest is the most accurate model, with an accuracy of up to 92 percent over the entire period.

Table 2. Model accuracy of each scenario.

Period		T1	T2	T3	T4
Feature		FC1, RE1, FC2, RE2, FP1, CC	FC1, RE1, FC2, RE2, FP1, SC1, RS1, CC	FC1, FC2, FC3, FC4 FC5, RE1, RE2, RE3, RE4, RE5, FC2, RE2, FP1, SC1, RS1, CC	FC1, FC2, FC3, FC4, FC5, RE1, RE2, RE3, RE4, RE5, FC2, RE2, FP1, FP2, FP3, FP4, SC1, SC2, RS1, RS2, CC
Target				GC	
% Accuracy [training, test]	LogReg	[71,72]	[80,78]	[85,88]	[92,90]
	DT	[94,62]	[99,66]	[100,79]	[100,83]
	RF	[94,64]	[99, 76]	[100,90]	[100, 92]
	KNN	[77,64]	[82,74]	[86,81]	[89,85]
	LDA	[70,70]	[79,79]	[86,88]	[96,92]
	GNB	[37,34]	[60,58]	[38,35]	[56,48]
	SVM	[71,69]	[80,79]	[85,87]	[93,90]

4. Conclusion

Only by using the data recorded by the teacher can a model be built to predict learning progress from the beginning of the lecture. This study shows that learning progress modeling can also be carried out throughout learning, and model accuracy is fairly good in 4 learning periods. The result of this research is a learning progress model with machine learning techniques that are used to monitor the learner automatically. From Table 2, accuracy models will be better if there are more features and closer to the end of the course. The better the model at the end of the semester, this is less useful because the opportunities for students and teachers to improve themselves are getting smaller. The accuracy of the model at the beginning of the lecture can be improved in several ways, such as adding features, improving data, and using other more suitable algorithms. Adding other features related to targets such as course grades which are prerequisites, intelligence values, or those related to learning such as learning logs, student discussion activities, student backgrounds, institutions, and curriculum, or those related to student backgrounds such as economic level, parents' education, distance to school. Other features can also be internal and external factors that affect learner progress such as motivation, learning styles, family support, and school facilities. The model generated from this study only comes from one subject in one semester for all students in junior high school at that time, to get a better pattern, you can use student data in 1 or several cohorts. The ensemble model algorithm can also be used to improve accuracies, such as bagging, boosting, or stacking.

Acknowledgment

This research is part of dissertation research entitled Learning Progress Modeling Towards Learning Monitoring Automation, in the doctoral program of the Faculty of Computer Science, Universitas Indonesia. This research was conducted under the auspices of the Digital Library and Distance Learning Lab, the Faculty of Computer Science, Universitas Indonesia, chaired by Harry B.S, Ph.D. The author would like to thank Prof. Sofyan Hanif as head of BPS Lab school YP-UNJ, specifically Yulinda Asnita, M.Pd., Deputy Principal for Academic Affairs at Lab school Kebayoran Middle School for the data provided.

References

1. Gaikwad, N.; and Mistry, Y. (2015). Review on environment monitoring system and energy efficiency. *International Journal of Engineering Research and Applications*, 5(7), 90-92.
2. Alimuddin, A.; Arafiyah, R.; Sirajuddin, S.; Lady, N.; Susiono, W.; and Masjudin, M. (2021). System of monitoring on carbon dioxide of broiler chicken growth in closed house. *IOP Conference Series: Materials Science and Engineering*, 1098(4), 042016.
3. Alimuddin, A.; Arafiyah, R.; Saraswati, I.; Alfanz, R.; Hasudungan, P.; and Taufik, T. (2021). Development and Performance study of temperature and humidity regulator in baby incubator using Fuzzy-PID hybrid controller. *Energies*, 14(20), 1-21.
4. Mühling, A.; Jungjohann, J.; and Gebhardt, M. (2019). Progress monitoring in primary education using Levumi: A case study. *Proceedings of the* 11*th International Conference on Computer Supported Education*. Crete, Greece, 137-144.
5. Badan Pusat Statistik. (2020). *Potret pendidikan Indonesia, Statistik Pendidikan* 2020. Jakarta: Badan Pusat Statistik.
6. Hadioui, A.; El Faddouli, N.E.; Benjelloun Touimi, Y.; and Mohammed, S.B. (2017). Machine learning based on big data extraction of massive educational knowledge. *International Journal of Emerging Technologies in Learning*, 12(11), 151-167.
7. Arafiyah, R.; Hasibuan, Z.A.; and Santoso, H.B. (2021). Monitoring online learners' performance based on learning progress prediction. *AIP Conference Proceedings*, 2331(1), 060012.
8. Almohammadi, K.; Hagras, H.; Alghazzawi, D.; and Aldabbagh, G. (2017). A survey of artificial intelligence techniques employed for adaptive educational systems within e-learning platforms. Journal of Artificial Intelligence and Soft Computing Research, 7(1), 47 - 64.
9. Meiza, A.; Darwis, S.; Gunawan, A. Y.; and Fitriana, E. (2017). Quantitative analysis of the relationship between three psychological parameters based on swallowtail catastrophe model. *International Journal of Mathematics and Mathematical Sciences*, 2017, 1-6.
10. Kpolovie, P. J.; Joe, A. I.; and Okoto, T. (2014). Academic achievement prediction: Role of interest in learning and attitude towards school. *International Journal of Humanities Social Sciences and Education* (*IJHSSE*), 1(11), 73-100.

11. Murphy, K.P. (2012). *Machine learning: A probabilistic perspective*. London: The MIT Press.
12. Golino, H. F.; and Gomes, C. M. A. (2014). Four machine learning methods to predict academic achievement of college students: A comparison study. *Revista E-Psi*, 4(1), 68-101.
13. Gray, G.; McGuinness, C.; and Owende, P. (2014). An application of classification models to predict learner progression in tertiary education. *Proceedings of the* 4*th IEEE International Advance Computing Conference*. Gurgaon, India, 549-554.
14. Widyahastuti, F.; and Tjhin, V.U. (2017). Predicting students performance in final examination using linear regression and multilayer perceptron. *Proceedings of the* 10*th International Conference on Human System Interactions* (*HSI*). Ulsan, Korea (South), 188-192.
15. Sunday, K.; Ocheja, P.; Hussain, S.; Oyelere, S. S.; Balogun, O.S.; and Agbo, F.J. (2020). Analyzing student performance in programming education using classification techniques. *International Journal of Emerging Technologies in Learning*, 15(2), 127-144.
16. Evangelista, E. (2021). A hybrid machine learning framework for predicting students' performance in virtual learning environment. *International Journal of Emerging Technologies in Learning* (*iJET*), 16(24), 255-272.
17. Hsu, J.L.; Chou, H.W.; and Chang, H.H. (2011). EduMiner: Using text mining for automatic formative assessment. *Expert Systems with Applications*, 38(4), 3431-3439.
18. Mwambe, O.; and Kamioka, E. (2019). Utilization of learners' metacognitive experiences to monitor learners' cognition states in e-learning platforms. *International Journal of Information and Education Technology*, 9(5), 362-365.
19. Krathwohl, D.R.; Anderson, L. W.; Airasian, P. W.; Cruikshank, K.A.; Mayer, R. E.; Pintrich, P. R.; Raths, J.; and Wittrock, M. C. (2002). A Taxonomy For Learning, Teaching, And Assessing: A Revision Of Bloom's Taxonomy Of Educational Objectives. *New York Longman*, 41(4), 302.
20. Arafiyah, R, Hasibuan, Z. A.; and Santoso, H.B. (2021). Development of Early Warning Systems to monitor e-learning progress. *IOP Conference Series: Materials Science and Engineering*, 1098(3), 032062.
21. Hackeling, G. (2017). *Mastering Machine Learning with scikit-learn* (2nd ed.). Birmingham-Mumbai: Packt Publishing.
22. Fauszt, T.; Bognár, L.; and Sándor, Á. (2021). Increasing the Prediction Power of Moodle Machine Learning Models with Self-defined Indicators. *International Journal of Emerging Technologies in Learning* (*iJET*), 16(24), 23-39.
23. Arafiyah, R.; Hasibuan, Z. A.; and Santoso, H. B. (2018). Predicting Learner's Achievement Using Step Function. *Proceedings of the International Conference on Computer Applications and Information Security*. *Riyadh, Saudi Arabia*, 1-5.