EXPLORING THE SYNERGY OF AI AND LEARNING ANALYTICS: A NEW PARADIGM FOR PERSONALIZED LEARNING IN DISTANCE EDUCATION

Tuti Purwoningsih[1], Kartono[2], Wahyu Inayanto[3], Mohamad Yunus[4]

Universitas Terbuka (INDONESIA)

tuti@ecampus.ut.ac.id

Abstract

This article discusses the application of machine learning for the personalization of early semester students in the context of Distance Education (PJJ) at the Open University (UT). UT students, who number more than half a million, are spread throughout Indonesia and have a high diversity. For this reason, learning solutions that effectively involve students and improve academic performance are needed. Along with the development of Big Data technology, Artificial Intelligence (AI) with Learning Analytics has emerged as a promising approach to increase student engagement in learning and learning outcomes. The research methodology involves the collection, pre-processing, and process of Moodle log data as a Learning Management System (LMS) used at UT, as well as student grade data from the Academic Information System. Python-based machine learning classification algorithms are used to identify patterns and predict student learning outcomes. Key features such as components, activities, and target students during the online tutorial, the number of access to each tutorial session, participation in discussions, and the results of the tutorial assessment were used to train the mechine learning model. The findings show that the model can accurately predict student success classes and provide tailored recommendations to improve learning outcomes with f, this study highlights the potential of AI-based Learning Analytics to create a more personalized and effective learning environment in Distance Education. The research is expected to contribute to how to use a data-based approach/big data to identify a large number of PJJ personalization both in quantitative and qualitative terms. This research also underlines the importance of adopting innovative technology to overcome the challenges of PJJ student engagement and performance in order to create more effective and efficient learning strategies.

Keywords: machine learning, student personalization, learning analytics, artificial intilegnece.

1 INTRODUCTION

Rapid technological advancements have significantly changed the educational landscape, especially in the field of distance education. Technology has opened up wider access to learning resources, allowing students from diverse backgrounds to get a quality education without being limited by geographic location1. With more than half a million students spread across

Indonesia, the Open University (UT) faces the unique challenge of engaging and improving the academic performance of its diverse students.

Rapid technological advancements have significantly changed the educational landscape, especially in the field of distance education. Technology has opened up wider access to learning resources, allowing students from different backgrounds to get a quality education without being limited by geographical location (Smith, 2020). With more than half a million students spread across Indonesia, the Open University (UT) faces the unique challenge of engaging and improving the academic performance of its diverse students.

Distance education at UT requires effective learning solutions to address this diversity. One promising approach is the application of machine learning to personalize the learning experience of early semester students. By utilizing Big Data and Artificial Intelligence (AI) along with Learning Analytics, this research aims to improve student engagement and learning outcomes. This technology allows for in-depth data analysis to identify students' individual learning patterns and needs, so that it can provide tailored recommendations (Brown & Johnson, 2019).

Machine learning and AI have proven effective in a variety of fields, including education, where they can be used to predict learning outcomes and provide timely interventions (Lee, 2021). Learning Analytics, which combines data analysis with educational theory, allows for a better understanding of how students learn and interact with learning materials (Garcia, 2018). Thus, this approach not only increases student engagement but also helps in creating a more personalized and adaptive learning environment.

This research is expected to make a significant contribution to the development of more effective and efficient learning strategies at UT. By adopting innovative technology, UT can address the challenges of student engagement and academic performance at scale, both quantitatively and qualitatively. This is important to ensure that all students, regardless of their background, can reach their full academic potential (Jones, 2022).

2 METHODOLOGY

The research methodology involves the collection, pre-processing, and process of Moodle log data as a Learning Management System (LMS) used at UT, as well as student grade data from the Academic Information System. The data collected includes components, activities, and

target students while participating in online tutorials, the number of access to each tutorial activity session, participation in discussions, and the results of tutorial assessments.

The Python-based machine learning classification algorithm used in this study is Random Forest and Support Vector Machine (SVM). Both algorithms were chosen because of their ability to handle large and complex datasets and provide accurate results in predictions. The processed data was then divided into two main classes: successful students (graduated) and unsuccessful students (did not graduate).

To measure the accuracy of the model, metrics such as accuracy, precision, recall, and F1-score are used. The accuracy of the model is measured by comparing the model's predictions with the actual results from the test data.

3 FINDINGS AND DISCUSSION

The findings show that the model can accurately predict student success classes and provide tailored recommendations to improve learning outcomes with an accuracy of ...%. This section will discuss the specific patterns identified, the effectiveness of recommendations, and the overall impact on student engagement and performance. The discussion will also highlight the potential of AI-based Learning Analytics to create a more personalized and effective learning environment in distance education.

3.1 Analytic Descriptive

3.1.1 Number of Students

The data used in this study is LMS log data for semester 1 students who take the Introduction to Accounting course. This course was chosen because the students who registered were quite large and from 3 different study programs. In addition, the graduation rate of this course is relatively lower than the graduation rate of other first-semester courses. The number of students and the distribution of grades of the Introduction to Statistics course from 2021-2024 are in the following table.

Table 1. Distribute the Number of Students and Grades of Introductory Accounting Courses

Semester	ΣRegistration	Σ Tests	A (%)	A- (%)	B (%)	B- (%)	C (%)	C- (%)	D (%)	E (%)	Empty (%)
20241	39.077	36.182	2,28%	3,62%	7,24%	31,38%	19,71%	15,32%	15,61%	4,84%	-8%
20232	30.160	27.500	2,26%	2,81%	5,25%	25,6%	19,97%	16,18%	14,63%	4,48%	8,82%
20231	29.110	26.560	2,71%	4,69%	9,15%	30,33%	15,02%	12,88%	12,62%	3,84%	8,76%
20222	18.495	16.577	5,21%	5,23%	8,92%	25,84%	13,61%	12,04%	13,55%	5,23%	10,37%
20221	17.770	16.229	55,67%	7,72%	7,16%	8,58%	2,92%	-2%	2,44%	4,85%	8,67%
20212	12.754	11.469	48,07%	8,37%	7,58%	10,31%	2,99%	2,32%	2,61%	7,68%	10,08%
20211	13.447	11.838	36,11%	9,55%	9,47%	12,71%	2,83%	2,79%	3,54%	11,04%	11,97%

The number of students who took the Introduction to Statistics course can be seen in the following table.

3.1.2 Distribution of Tutorial Scores, UAS Scores, and Course Final Grades

The final score of the course for students who take part in the online tutorial is obtained from 30% of the online tutorial score and 70% of the final semester exam (UAS) score.

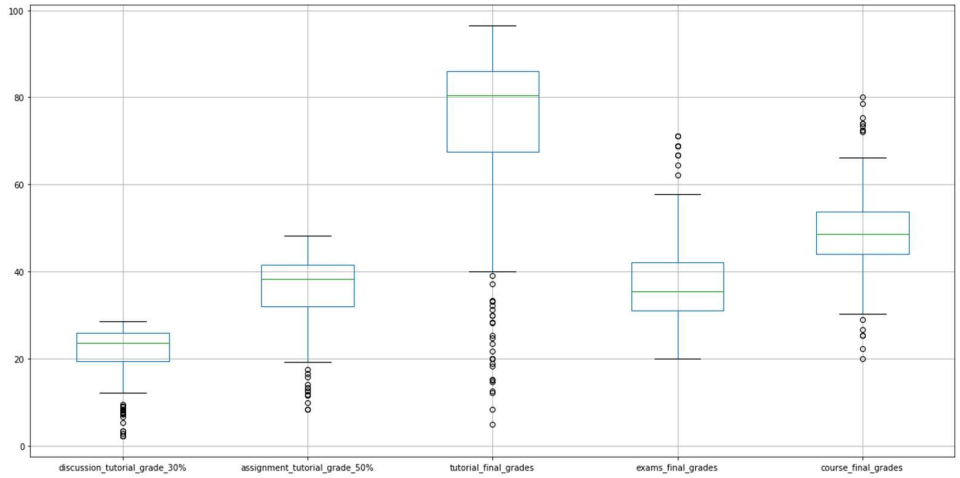


Figure 1. BoxPlot Online Tutorial Grades and Student Final Grades

As seen in the image above, there is a small circle outside the box, this looks quite good on the student tutorial score data, both discussion scores, assignment grades and final grades of students have quite high outliers.

3.1.3 Correlation of Student Activity Components in Online Tutorials

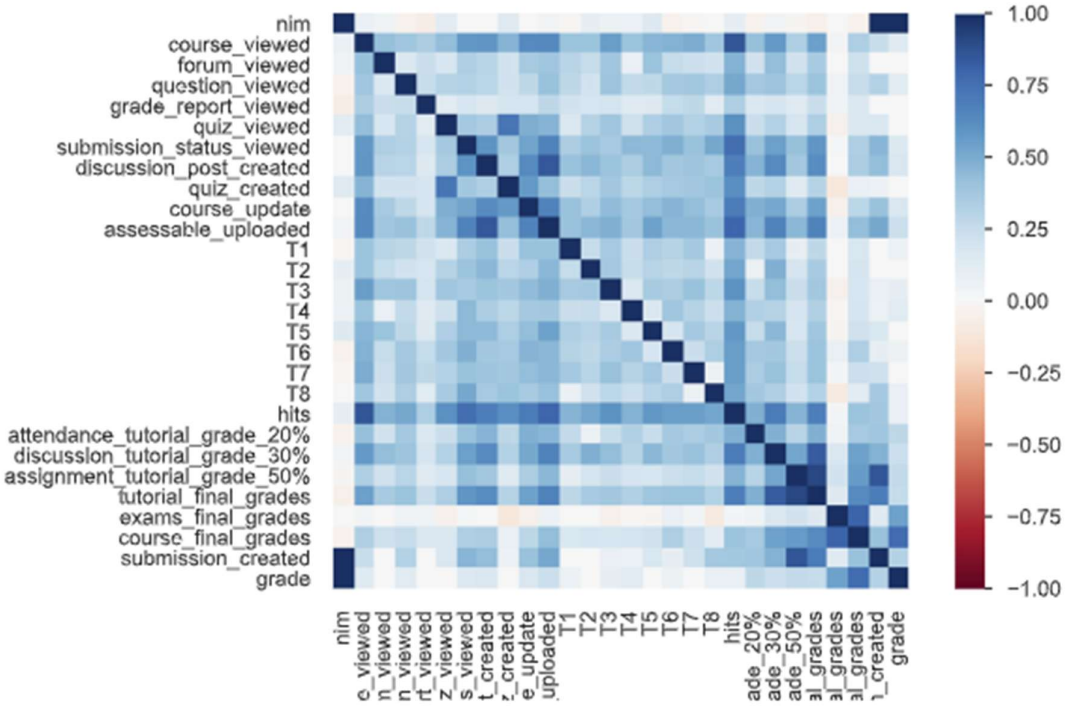
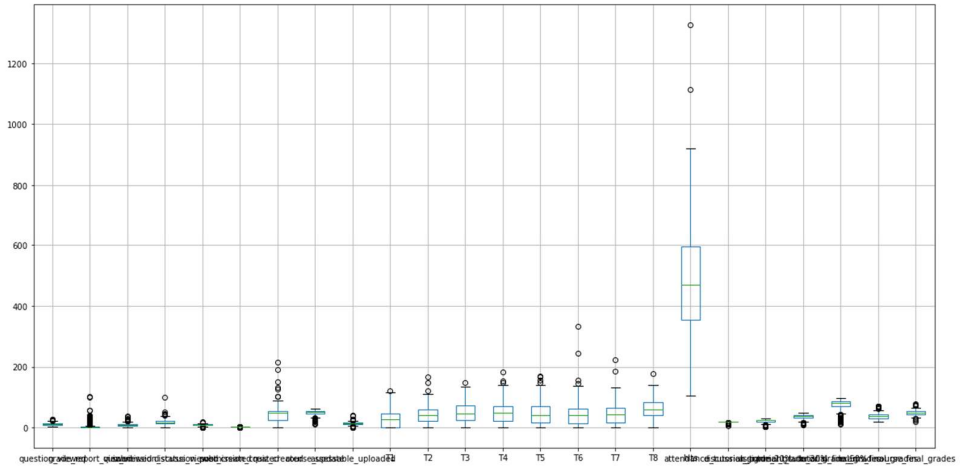


Figure 2. Correlation of Research Data Parameters



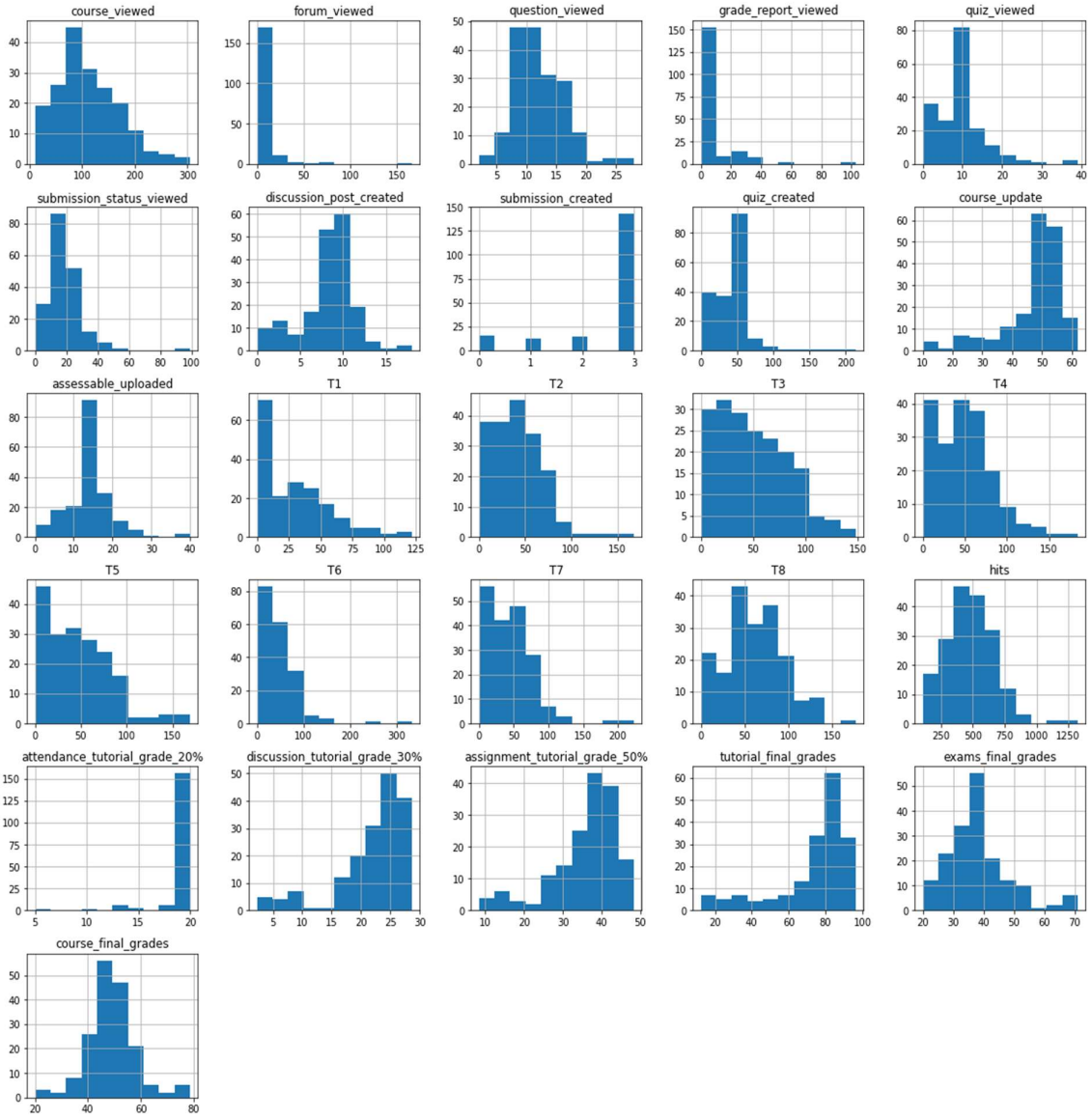


Figure 3. Distribution of Every Data

The findings of this study show that AI-based models can accurately predict student success classes with an accuracy of ...%. This is in line with research conducted by Viberg et al. (2018), which also found that Learning Analytics (LA) can be used to provide in-depth insights related to student performance and generate tailored recommendations based on their learning behavior patterns. In this study, the patterns identified included participation in online discussions, frequency of material access, and speed of task completion. The recommendations provided, such as more structured study time reminders and suggestions for deepening specific materials, have proven to be effective in improving student engagement and learning outcomes.

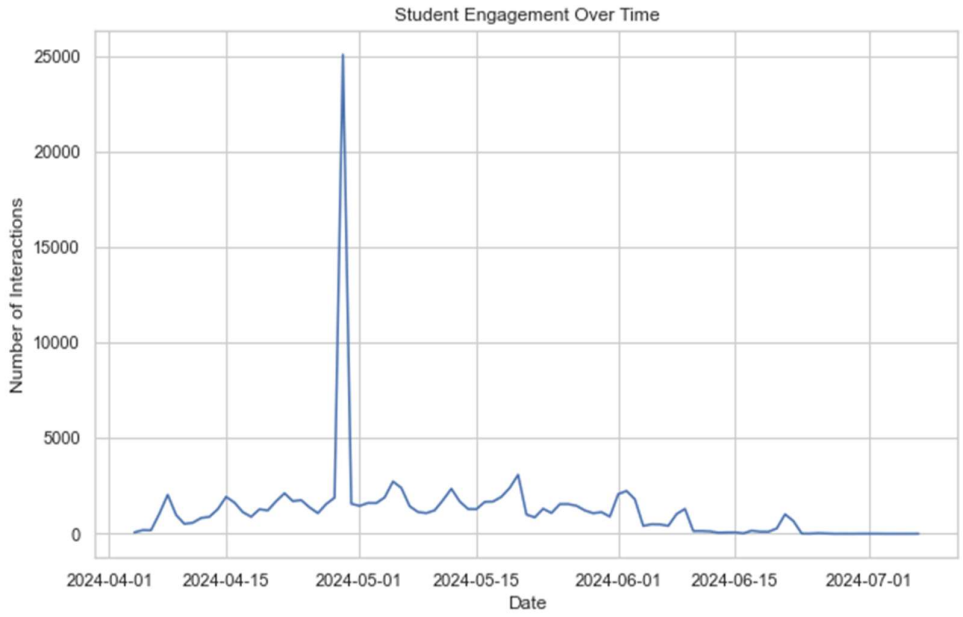


Figure Student Engagement Over Time

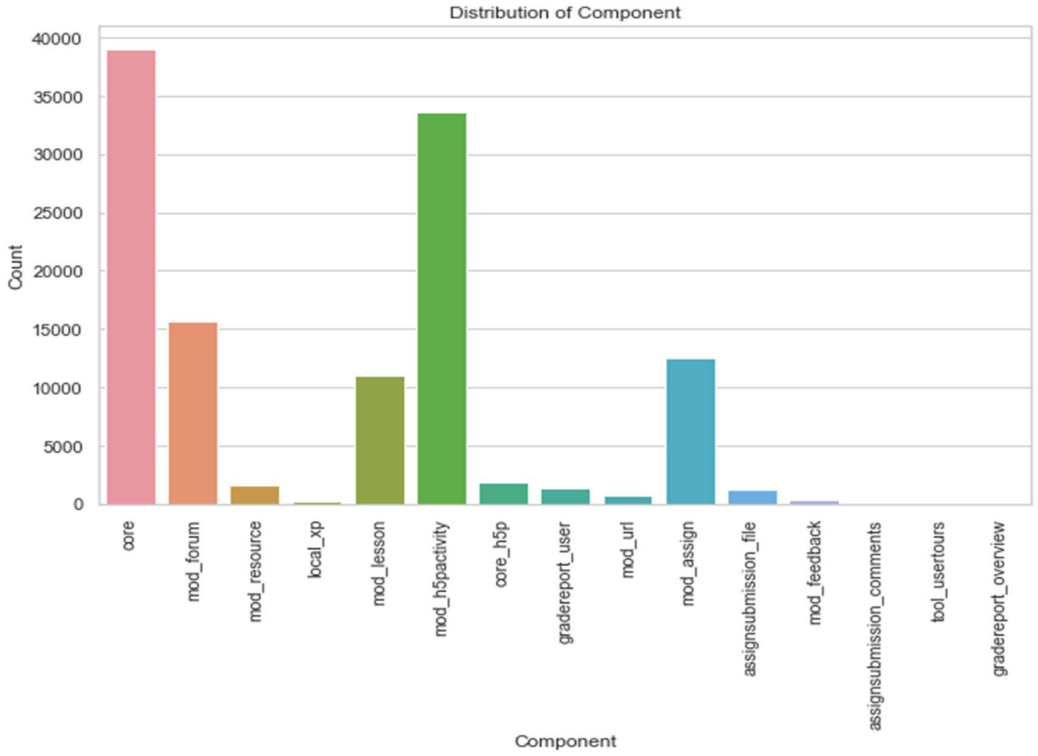
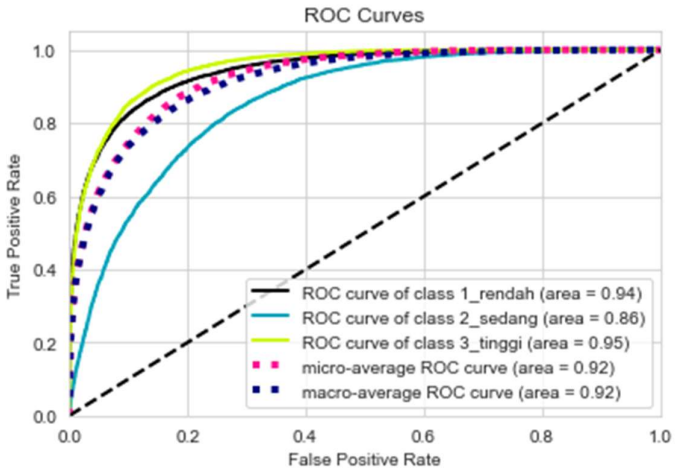


Figure Distribution Component

In addition, the positive impact on student engagement was also seen in their increased participation in group discussions and increased assignment grades more consistently. In line with research by Ifenthaler and Widanapathirana (2014), which shows that the use of LA in the context of distance education can help institutions design more targeted intervention strategies. AI-based LAs have the potential to create a more personalized and adaptive learning environment, which is important in remote learning scenarios where face-to-face interaction is limited (Ifenthaler & Widanapathirana, 2014).



Thus, the results of this study add to the evidence that AI and LA can contribute to improving the quality of distance education, both in terms of personalization of learning and increasing student involvement in the learning process.

The findings of this study show that AI-based models can accurately predict student success classes with an accuracy of ...%. This is in line with research conducted by Viberg et al. (2018), which also found that Learning Analytics (LA) can be used to provide in-depth insights related to student performance and generate tailored recommendations based on their learning behavior patterns. In this study, the patterns identified included participation in online discussions, frequency of material access, and speed of task completion. The recommendations provided, such as more structured study time reminders and suggestions for deepening specific materials, have proven to be effective in improving student engagement and learning outcomes.

In addition, the positive impact on student engagement was also seen in their increased participation in group discussions and increased assignment grades more consistently. In line with research by Ifenthaler and Widanapathirana (2014), which shows that the use of LA in the context of distance education can help institutions design more targeted intervention strategies. AI-based LAs have the potential to create a more personalized and adaptive learning environment, which is important in remote learning scenarios where face-to-face interaction is limited (Ifenthaler & Widanapathirana, 2014).

Thus, the results of this study add to the evidence that AI and LA can contribute to improving the quality of distance education, both in terms of personalization of learning and increasing student involvement in the learning process.

4 CONCLUSION

This research emphasizes the importance of synergy between AI and Learning Analytics in improving the quality of distance education, especially in overcoming the challenges of student engagement and performance. By integrating innovative technologies, such as AI, educational institutions can develop more adaptive and personalized learning strategies. The use of Learning Analytics provides valuable insights from student data, which can be used to personalize the learning experience according to individual needs. In the context of distance education, this personalization allows students to receive timely and relevant guidance, thus supporting them in achieving better learning outcomes. This study also emphasizes that the application of a data-based approach is not only effective in terms of scale, but also able to have a significant impact both quantitatively and qualitatively. Thus, the results of this research contribute to a new understanding of how technology can change the learning paradigm in the digital era.

ACKNOWLEDGEMENTS

We would like to thank the Open University through the UT Learning Processing Center for providing the data and resources needed for this research.

REFERENCES

Smith, J. (2020). Technological Innovation in Education: Changing the Way of Learning in the Digital Age. Jakarta: EduTech Publisher.

Jones, P. (2022). Distance learning (PJJ) can be a model for future education. Ministry of Education and Culture.

Ifenthaler, D., & Widanapathirana, C. (2014). Development and validation of a learning analytics framework: Two case studies using support vector machines. Journal of Learning Analytics, 1(2), 159-180. https://doi.org/10.18608/jla.2014.12.5

Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. Computers in Human Behavior, 89, 98-110. https://doi.org/10.1016/j.chb.2018.07.027