*Jurnal Hurriah: Jurnal Evaluasi Pendidikan dan Penelitian*
*Hurriah Journal: Journal of Educational Evaluation and Research*
*Volume* 6, No. 4 (2025) PP 1440-1448

Jurnal Hurriah: Jurnal Evaluasi Pendidikan dan Penelitian
Hurriah Journal: of Educational Evaluation and Research

Journal homepage: https://academicareview.com/index.php/jh
E-ISSN: 2774-8472 I P-ISSN: 2774-8480 I DOI : 10.56806

# Longitudinal Analysis of Knowledge and Skills Scores in Vocational Multimedia Education: An AI-Driven Predictive Approach in Indonesian High Schools

Rosita[1,*], Wahyu Sasongko Putro[2], Rina Dwijuliani[3]

[1] Department of Informatics Management, Universitas Negeri Surabaya, 60231, Surabaya, East Java, Indonesia.
[2] Department of Electrical Engineering, Universitas Negeri Surabaya, 60231, Surabaya, East Java, Indonesia
[3] Department of Beauty Education, Faculty of Engineering, Surabaya State University, Surabaya, Indonesia.

**ARTICLE INFO**

***Article history:***
Received: October 2025
Received in revised from: December 2025
Accepted: January 2026
Available online: January 2026

***Keywords:*** *AI in vocational education, Longitudinal Analysis, Student performance prediction.*

**ABSTRACT**

This study aimed to examine the progression of students' knowledge (P) and skills (K) scores in an Indonesian Vocational High School on Multimedia department using longitudinal artificial intelligence (AI) driven analysis. In this study, the data observation taken from academic scores from 32 students gathered over five semesters and examined via data preprocessing, descriptive statistical analysis, and machine learning modelling employing the Random Forest algorithm. The results show that both knowledge and skills scores have been going up steadily over the semesters. The predictive model based on Random Forest works very well, with a high level of accuracy and a low level of prediction error. Additionally, Pearson correlation analysis and simple linear regression demonstrate that knowledge significantly and positively influences students' skills ($p < 0.05$), suggesting that proficiency in cognitive dimensions directly facilitates the enhancement of practical skills in vocational education. These results validate that the amalgamation of longitudinal analysis and artificial intelligence can enhance data-driven learning assessment and promote more precise academic decision-making in vocational education.

## 1. Introduction

The rapid advancement of artificial intelligence (AI) technology has changed the way people learn all over the world in education especially over vocational education. AI has become an important part of changing digital learning in vocational education, analyzing academic data, and making predictions-based evaluations in the last few years (Muttaqin et al., 2024). Systematic reviews show that using AI in vocational education makes learning more efficient, customizes learning experiences, and gives students feedback that is tailored to their needs (Suparyati et al., 2024). In this context, vocational education in Indonesia's multimedia sector possesses considerable potential to utilize AI, especially for the analysis of students' academic progress through the continuous generation of assessment data over time.

[*] *Corresponding author.*
*E-mail address: rositarosita@unesa.ac.id*
https://doi.org/10.56806// *jh.v6i4.388*

Indonesian vocational education is strategically important for getting people ready for Industry 4.0, but it still has a "competency gap." Empirical evidence reveals a substantial gap between students' theoretical knowledge (P) and practical skills (K). Kemendikbudristek (2023) says that more than 40% of vocational graduates have trouble meeting industry standards because their scores are not balanced. High theoretical scores don't always mean that they are good at technical skills. In multimedia programs, which need a careful balance of technical skills like animation and theoretical visual communication, traditional teaching methods often don't address this gap (Setiawan et al., 2023; Widayanthi, 2024).

A longitudinal is useful for figuring out how students' knowledge and skills change from first semester to last semester. Saleh (2018) showed that longitudinal analysis is impactful in academic performance that get better or worse over certain time semester periods. This method can help teachers figure out what kinds of help are needed. Furthermore, by combining longitudinal analysis with AI the descriptive analysis can be develop predictive modelling that can guess how students will learn in the future (Azizah et al., 2024). AI also makes it easier to find hidden patterns in large amounts of educational data that are hard to find with traditional statistical methods.

In vocational education, AI application for forecasting and evaluating student performance has yielded encouraging outcomes. Trujillo et al. (2025) discovered that machine learning models, including Random Forest and XGBoost, can precisely forecast students' career trajectories and performance utilizing historical academic data. Tao (2025) similarly illustrated that deep learning technologies can be utilized to suggest tailored learning strategies for vocational students. Other research has shown that using fuzzy logic algorithms and AI-based predictive systems can make learning management in vocational schools more effective (Liang et al., 2024). These results indicate that AI can provide a basis for intelligent, adaptive, and evidence-based educational systems.

Although there is an increasing amount of international research on AI in vocational education, there are still not many empirical studies that look at how AI is being used in Indonesian vocational high schools with real vocational high school student data. Most of the prior research is conceptual or predominantly concentrates on the adoption of educational technologies rather than on longitudinal analyses of students' academic performance. Consequently, this study seeks to examine the progression of knowledge (P) and skills (K) scores among vocational high school students enrolled in the multimedia program in Indonesia over five semesters, employing an AI-based methodology. This study aims to provide empirical evidence from Indonesian vocational high school students, contributing significantly to the development of data-driven vocational learning models and the enhancement of adaptive educational informatics management in the age of artificial intelligence.

## 2. Methodology

### *2.1 Research Design*

This research utilizes a quantitative methodology featuring an artificial intelligence (AI)–driven longitudinal research framework. The longitudinal design was chosen to examine the ongoing variations in students' knowledge (P) and skills (K) scores across five semesters. The AI-based method was used to find patterns in students' academic progress and to guess how their performance would change in the future based on past academic data.

### *2.2 Data and location*

The participants in this study were Multimedia major students at an Indonesian Vocational High School who had completed five semesters of coursework. The main source of data was official academic transcripts for 32 students. These secondary data points, taken from faculty grade records from Semesters 1 to 5, show how well students did on tests of their knowledge (P) and skills (K) in both general and vocational subjects.

Before using machine learning models, the raw transcript data went through a strict preprocessing pipeline to make sure the data was of high quality by Taking care of Mean Imputation was used to fill in missing scores (less than 3% of the total dataset) because the data was found to be missing completely at random (MCAR). In addition, the normalization data is used to bring all values into the range of 0 to 1 because the scores for P (Knowledge) and K (Skill) can change depending on the weight of the curriculum. Furthermore, Outlier Detection is used to obtain Z-score analysis to find and check extreme values to make sure they were real academic performance and not mistakes in the records.

### *2.3 Research Variables*

Each semester, both research variables were measured and analyzed to find patterns of improvement, decline, or stability in how well students did over time. These are the two variables:

a) Knowledge Variable (P): The cognitive score that shows how well students understand the concepts in each subject.
b) Skills Variable (K): The psychomotor score that shows how well students can use multimedia concepts in real life, like graphic design, 2D/3D animation, and creative production.

The knowledge (P) and skills (K) scores were put in a long table, with each row showing the scores for one student and each column showing the scores for a certain semester. In this study, the preprocessing data included cleaning the data (getting rid of missing values or outliers), normalizing the scores (scaling them from 0 to 100 to make sure they are the same across semesters), and arranging the data sets (putting them in chronological order from Semester 1 to 5).

In addition, a descriptive statistical analysis was conducted to analyze the distribution of mean scores, standard deviations, and the range of score improvements for both P and K variables in each semester. The results were used to show how students' academic progress usually goes.

### *2.4 AI-Based Longitudinal Analysis*

This research utilizes a quantitative longitudinal design, employing Supervised Machine Learning (ML) to evaluate and predict student academic performance. The selection of ensemble models namely Random Forest (RF) and Extreme Gradient Boosting (XGBoost) is substantiated by their demonstrated effectiveness in Educational Data Mining (EDM). Here, the Random Forest is Used because it is strong against overfitting in small to medium datasets through bagging (bootstrap aggregating). It works especially well with the non-linear relationships that often exist between theoretical (P) and practical (K) scores while XGBoost was chosen because it is better at making predictions and can work with sparse data using gradient boosting. Its regularization parameters help keep the root mean square error (RMSE) low while dealing with the fact that longitudinal semester data is collected over time.

### *2.5 Hypothesis Testing*

A hypothesis test was performed to examine the impact of theoretical knowledge (P) on practical skills (K) among Multimedia vocational high school students. The data were examined using linear regression analysis. In this instance, hypothesis testing was utilized to investigate the impact of knowledge (P) on skills (K) among vocational high school students enrolled in the Multimedia program. The analysis utilized simple linear regression, designating knowledge as the independent variable (X) and skills as the dependent variable (Y). We chose linear regression because the data had continuous numerical scores and we wanted to see if there was a causal link between mastering theoretical knowledge and practical skills. The significance level for hypothesis testing was 0.05. The research hypotheses were articulated as follows:

*$H_0$: Knowledge (P) has no significant effect on students' skills (K).*
*$H_1$: Knowledge (P) has a significant effect on students' skills (K).*

The decision criteria were predicated on the significance value (p-value). When $p < 0.05$, the null hypothesis ($H_0$) was turned down, and the alternative hypothesis ($H_1$) was accepted. The results of this test were utilized to reinforce the findings from the previously conducted correlation analysis and AI-based longitudinal analysis.

## 3. Results

*3.1 Trends in Average Knowledge (P) and Skills (K) Scores per Semester*

A descriptive analysis of students in the Multimedia program shows that the average scores for knowledge (P) and skills (K) went up steadily from Semester 1 to 5. The average P score was about 84 and the average K score was about 85 in the first semester. By Semester 5, these scores had gone up to about 91 for both variables, and they kept going up over the next semesters.

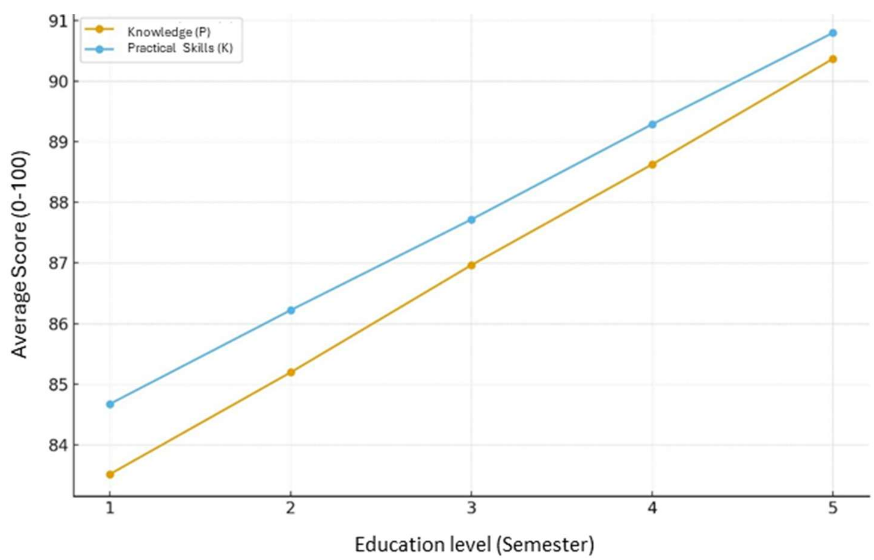


**Fig.1**. Average Knowledge and Skills Scores by Semester

As can be seen in figure 1, a continuous process of vocational learning has helped students do better in both practical and theoretical areas. The observed enhancement indicates that instructional interventions, organized practice, and ongoing assessment are vital in improving student learning outcomes. These results align with longitudinal studies in vocational education that underscore the significance of sustained monitoring of motivation, engagement, and skill acquisition (Yang & St. John, 2023). In the same way, a study by Yağcı (2022) shows that adaptive learning and data-driven personalization can help students do their best in school. So, vocational high schools with multimedia programs should focus on a balanced curriculum that combines theory with practice. They should also set up long-term monitoring systems to keep track of how students are doing over time. It is also important to think about outside factors like student motivation, the quality of teaching, and the availability of learning spaces, which could influence trends in academic performance improvement.

### *3.2 Random Forest Prediction vs. Actual Values*

In order to achieve the result, the Random Forest Regressor algorithm used data from Semesters 1 to 4 to guess the average scores in Semester 5. The model performed well, as shown by a low Root Mean Square Error (RMSE) and a high coefficient of determination ($R^2 > 0.9$). This means that the model was able to explain most of the differences in students' final scores. Figure 2 shows the predicted values are very close to the actual values, which means that the model is very accurate.

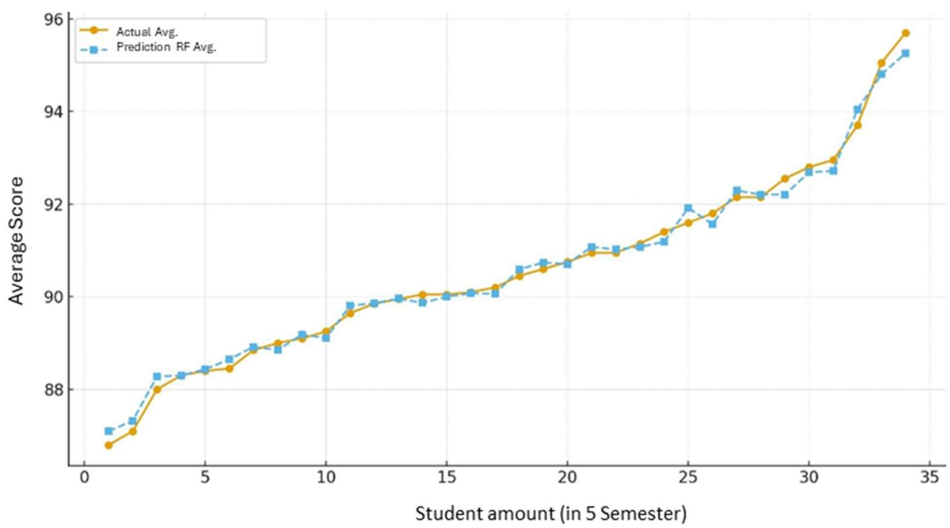


**Fig.2**. Random Forest Prediction vs. Actual Values

Figure 2 shows the predicted and actual average score values for the forest. The results align with Yağcı's (2022) study, which demonstrated the efficacy of Random Forest in forecasting students' academic performance using historical data. Research conducted by Kocakoyun-Aydoğan et al. (2024) similarly highlighted that this algorithm is especially proficient in identifying potential declines in academic performance and facilitating AI-driven learning recommendation systems. Consequently, the utilization of machine learning models, including Random Forest, has demonstrated efficacy in aiding educational institutions to discern student achievement patterns in a predictive rather than solely descriptive capacity. AI models can be incorporated into academic management systems to facilitate the early identification of potential declines in student performance, thereby enabling teachers to execute proactive interventions through academic counseling or tailored instructional planning. In this study, we find that the RMSE and $R^2$ values are 0.19 and 0.991, respectively.

### *3.3 Correlation Between Knowledge (P) and Skills (K)*

Pearson correlation analysis indicated a robust positive correlation between knowledge (P) and skills (K) scores in Semester 5 ($r \approx 0.9$). This shows that students who do better on knowledge tests also tend to do better on skills tests. The scatter plot in Figure 3 shows that the two variables have a positive linear relationship.

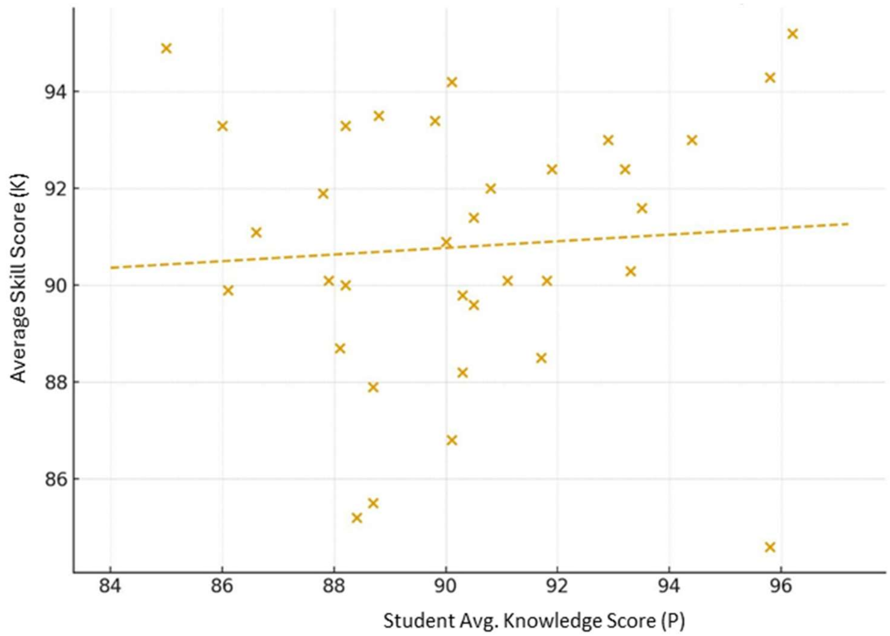


**Fig. 3**. Correlation Between Knowledge (P) and Skills (K)

Figure 3 shows the relationship between knowledge (P) and skill (K). The results here are in line with what Angeioplastis et al. (2024) found, which showed that project-based learning and integrative curricula are necessary for balancing theoretical understanding and practical application in vocational education. Schmid et al. (2022) also said that teaching methods that combine theory and practice make students more resilient and ready for the workforce. So, vocational high school classes shouldn't teach theory and practice separately. Project-based learning, studio-based learning, and working with people in the industry should all be important parts of the curriculum to help students connect their cognitive and psychomotor skills. In this study, the correlation (P and K) value is 0.071 and the Pearson value is 0.689, respectively.

### *3.4 Hypothesis Testing Results on the Effect of Knowledge on Skills*

The results of hypothesis testing with simple linear regression analysis show that knowledge (P) has a big impact on the skills (K) of vocational high school students in the Multimedia program. The regression coefficient indicates a positive correlation, suggesting that an elevation in knowledge scores corresponds with an enhancement in students' skills performance. The statistical test showed a significance value (p-value) of less than 0.05, which meant that the null hypothesis ($H_0$) was not true and the research hypothesis ($H_1$) was true. These results indicate that proficiency in theoretical knowledge plays a crucial role in enhancing students' practical skills within vocational education.

The regression results corroborate the Pearson correlation analysis, which indicated a robust positive correlation between knowledge and skills ($r \approx 0.9$). Consequently, both inferential statistical findings and AI-driven longitudinal analyses consistently demonstrate that cognitive and psychomotor dimensions are interconnected and evolve concurrently within the framework of vocational education.A summary of the linear regression test results is presented in Table 1, showing that knowledge has a significant effect on students' skills.

**Tabel 1**. Summary of Hypothesis Testing Results

| Statistic | Value |
|---|---|
| Test Method | Simple Linear Regression |
| Independent Variable (X) | Knowledge (P) |
| Dependent Variable (Y) | Skills (K) |
| Relationship Direction | Positive |
| Significance | $p < 0,05$ |
| Decision | $H_0$ Rejected |

As can be seen in Table 1, Utilizing a Simple Linear Regression model to analyze the correlation between student Knowledge (P) as the independent variable and Skills (K) as the dependent variable. The results show a positive relationship direction, which means that students' practical skills get better as their theoretical knowledge does. This alignment indicates that the two competencies reinforce each other within the Multimedia curriculum. The analysis produced a significance value of $p < 0.05$, which is very important. In academic research, this threshold verifies that the observed correlation is statistically significant and improbable to have arisen by chance. Consequently, the null hypothesis ($H_0$), which asserts the absence of an effect or relationship, was rejected. This statistical finding verifies that theoretical knowledge substantially predicts or affects the development of practical skills in the examined students.

Furthermore, based on the results of the AI-based longitudinal analysis of knowledge and skills scores among vocational high school students in the Multimedia program provide several strategic implications for vocational education in Indonesia, as follows:

a) Academic Implications, this study shows that AI-based longitudinal analysis can give a full picture of how students' academic performance changes over time. This method can help schools improve evidence-based learning evaluation by looking at both final learning outcomes and the processes of learning that lead to those outcomes.
b) Teaching and Learning Practices, Vocational teachers are encouraged to use AI-based analytics systems, like Random Forest or Long Short-Term Memory (LSTM) models, to keep an eye on how well their students are doing and plan personalized learning interventions. This approach is in line with the global trend toward data-driven learning design, which is becoming more common in vocational schools in Europe and Asia (Azizah et al., 2024; Fuad Muttaqin et al., 2024).
c) Educational Policy, the results show that there is a need for national policies that encourage the use of AI in vocational high school academic management systems. This is especially important for setting up Learning Analytics Dashboards and early warning systems to keep an eye on student performance.
d) Industry and Employment, the strong link between knowledge and skills shows that graduates who have a good mix of theoretical and practical skills are better prepared to meet the needs of the digital creative industry. These results can help vocational high schools and the multimedia industry work together to make adaptive, industry-relevant curricula.

## 3. Conclusions

This study shows that a longitudinal analysis method based on artificial intelligence (AI) works well to find patterns in the development of knowledge (P) and skills (K) scores among Indonesian vocational high school students in the Multimedia program. The results show that both variables improved steadily from Semester 1 to Semester 5, which shows that continuous and adaptive vocational learning processes work. The predictive model based on Random Forest did very well, with a high coefficient of determination ($R^2 > 0.9$) and a low prediction error. These results imply that machine learning algorithms can be dependably utilized to forecast students' academic performance using historical data, thereby facilitating the prospective establishment of early warning systems in vocational education.

The outcomes of Pearson correlation analysis and simple linear regression consistently indicate that knowledge (P) exerts a significant and positive influence on students' skills (K). The robust correlation coefficient ($r \approx 0.9$) and hypothesis testing outcomes with p-values under 0.05 affirm that proficiency in cognitive domains significantly enhances the acquisition of practical skills among Multimedia vocational students. These results underscore that cognitive and psychomotor domains do not evolve in isolation; rather, they are intricately interconnected and mutually reinforcing within the framework of vocational learning.

This study empirically demonstrates that the amalgamation of longitudinal analysis and artificial intelligence can enhance data-driven learning evaluation and decision-making processes. The utilization of AI models in vocational high school academic systems may augment the precision of student performance assessment, enhance the efficacy of pedagogical strategies, and facilitate the creation of more cohesive curricula that harmonize theoretical and practical elements. It is advisable for future research to incorporate larger sample sizes from various schools and vocational programs to enhance the generalizability of the results. It is also suggested to create AI-based learning analytics dashboards to help keep track of student performance in real time and on an ongoing basis.

## References

1. Angeioplastis, A., Papaioannou, N., Tsimpiris, A., Kamilali, A., & Varsamis, D. (2024). Predicting Student Specializations: A Machine Learning Approach Based on Academic Performance. Journal of e-Learning and Knowledge Society, 20(2), 19–27. https://doi.org/10.20368/1971-8829/1135904
2. Azizah, N., Hanafi, I., & Yusro, M. (2024). Artificial Intelligence in Vocational Education: Perspectives and Practices from a Literature Study. Global Scientific Education Journal, 7(2). https://doi.org/10.61667/w0efrt90
3. Azizah, N., Hanafi, I., & Yusro, M. (2024). Artificial Intelligence in Vocational Education: Perspectives and Practices from a Literature Study. Global Synthesis in Education Journal, 7(2). https://doi.org/10.61667/w0efrt90
4. Fuad Muttaqin, M., Sukrawan, Y., & Rosyadi, M. I. (2024). Exploring the Role of Artificial Intelligence in Learning Media for Vocational Education: A Systematic Literature Review. Journal of Vocational Education Studies, 7(2). https://doi.org/10.12928/joves.v7i2.10364
5. Higgins, N. L., Mercier, E. M., & O'Brien, K. L. (2023). Development of Self-Regulated Learning: A Longitudinal Study of Undergraduates in Science. Studies in Higher Education, 48(8), 1561–1577. https://doi.org/10.1080/03075079.2021.1997978
6. Kocakoyun-Aydoğan, Ş., Pura, T., & Bingül, F. (2024). Predicting Students' Academic Performances Using Machine Learning Algorithms. Malaysian Online Journal of Educational Technology, 12(4), 131–153. https://doi.org/10.52380/mojet.2024.12.4.557
7. Liang, Z., Wang, T., & Xu, Y. (2024). Enhancing the Performance of Vocational Education in the Digital Economy with the Application of Fuzzy Logic Algorithm. International Journal of Cognitive Informatics and Natural Intelligence. 17(185). https://doi.org/10.1007/s44196-024-00591-9
8. Muttaqin, A., Sukrawan, Y., & Rosyadi, M. I. (2024). Exploring the Role of Artificial Intelligence in Learning Media for Vocational Education: A Systematic Literature Review. Journal of Vocational Education Studies (JOVES), 7(2). https://doi.org/10.12928/joves.v7i2.10364

9. Niittylahti, S., Annala, J., & Mäkinen, M. (2016). Students' Learning Performance and Transitions in Different Learning Pathways to Higher Vocational Education. Vocations and Learning, 9, 315–332. https://doi.org/10.1007/s12186-016-9155-6
10. Rosyadi, F., Prasetyo, Y., & Hamzah, A. (2023). Artificial Intelligence Implementation in Indonesian Vocational Learning Systems. Jurnal Pendidikan Teknologi dan Kejuruan. https://doi.org/10.59562/semnasdies.v1i1.794
11. Saleh, A. (2018). Longitudinal Study on Student Performance in Vocational Education and Its Implications for Teaching Strategies. Jurnal Pendidikan Teknologi dan Kejuruan, Universitas Negeri Yogyakarta. https://doi.org/10.21831/jptk.v24i2.20019
12. Schmid, E., Stalder, B. E., & Eberle, F. (2022). Identifying Resilience-Promoting Factors in Vocational Education and Training (VET). Empirical Research in Vocational Education and Training, 14(7). https://doi.org/10.1186/s40461-022-00139-1
13. Setiawan, H., Nugroho, A., & Santoso, D. (2023). The Impact of Indonesia's Decentralized Education on Vocational Skills and Economic Improvement of Students. Jurnal Pendidikan Vokasi, 3(13). https://doi.org/10.21831/jpv.v13i3.68026
14. Suparyati, T., Widiastuti, I., Saputro, I. N., & Pambudi, N. A. (2024). The Role of Artificial Intelligence (AI) in Vocational Education. Jurnal Ilmiah Pendidikan Teknik dan Kejuruan (JIPTEK), 17(1). https://doi.org/10.20961/jiptek.v17i1.75995
15. Tao, Lingyun. (2025). Improving the Quality of Vocational Education Based on Deep Learning Technology: Student Learning Prediction and Personalized Recommendation. Journal of Advanced Computational Intelligence and Intelligent Informatics (JACIII), 29(2). https://doi.org/10.20965/jaciii.2025.p0407
16. Trujillo, S., García-López, E., & Martínez-Arroyo, C. (2025). Artificial Intelligence in Education: A Systematic Literature Review of Machine Learning Approaches in Student Career Prediction. Journal of Technology and Science Education (JOTSE). https://doi.org/10.3926/jotse.3124
17. Widayanthi, D. G. C. (2024). AI in vocational education: Balancing digital skills and humanism. Indonesian Journal of Educational Inquiry, 1(2), 1-10. https://doi.org/10.20965/jaciii.2025.p0401
18. Yağcı, M. (2022). Educational Data Mining: Prediction of Students' Academic Performance Using Machine Learning Algorithms. Smart Learning Environments, 9(11). https://doi.org/10.1186/s40561-022-00192-z
19. Yan, C. (2025). Long-Duration Vocational Education's Effects on Individuals' Vocational Identity, Self-Efficacy, and Job Satisfaction. Behavioral Sciences, 15(9), 1161. https://doi.org/10.3390/bs15091161
20. Yang, L., & St. John, E. P. (2023). Public Investment in Short-Cycle Tertiary Vocational Education: Historical, Longitudinal, and Fixed-Effects Analyses. Education Sciences, 13(6), 573. https://doi.org/10.3390/educsci13060573