Journal of Computer Science and Information Technology
Vol. 2 No. 3, September 2026, pp. 110-118
E-ISSN 3090-787X

Analysis of Student Engagement Levels in E-Learning Using Clustering Methods

Aoni Aslihatin [a,1,*], Nisreen Alzahrani [b,2,] Alejandro Mozo Quesada [C,3,]
[a] Program Studi Teknik Informatika, Universitas Mataram, Mataram, Indonesia
[b] Department of Information Systems, King Abdulaziz University, Jeddah, Saudi Arabia
[C] Department of Computer Science, University of Granada, Granada, Spain
[1] aoni.aslihatin01@gmail.com*; [2] nisreen.a@gmail.com; [3] alejandro.mozo@gmail.com
* Corresponding author

ARTICLE INFO

Article history
……………………………….
Received 20 May 2026
Revised 15 September 2026
Accepted 16 September 2026
Published 19 September 2026

Keywords
E-learning
Clustering
K-Means
Student Engagement
Learning Analytics

ABSTRACT

The rapid growth of e-learning has increased the need to analyze student engagement levels in online learning environments. This study aims to analyze student activity levels using a clustering method based on activity log data from a Learning Management System (LMS). The approach applied is K-Means Clustering with an exploratory method on student activity data, including frequency of access to learning materials, discussion forum participation, assignment submission, and duration of learning activities. The sample consisted of 200 Informatics Engineering students selected through purposive sampling, with data collected over one academic semester. The optimal number of clusters was determined using the elbow method, while cluster quality was evaluated using the Silhouette Score. The results show that students can be grouped into three main clusters: high, medium, and low engagement levels. The high-engagement cluster showed consistent participation and potential correlation with better academic performance, while the low-engagement cluster indicated a higher risk of academic difficulty. These findings are consistent with previous studies showing that interaction levels are strongly correlated with academic performance. Clustering methods proved effective in automatically identifying student behavior patterns, thereby supporting the development of data-driven adaptive learning strategies.

How to cite: Aslihatin, A., Alzahrani, N., & Quesada, A. M. (2026). Analysis of Student Engagement Levels in E-Learning Using Clustering Methods. *Journal of Computer Science and Information Technology, Vol 2*(3), 110-118. doi: https://doi.org/10.70716/jocsit.v2i3.550

INTRODUCTION

The rapid development of information and communication technology (ICT) has driven significant transformations in the education sector, particularly through the adoption of e-learning as a primary learning medium. E-learning enables learning to be delivered in a flexible and adaptive manner, overcoming conventional limitations of time and location. However, the widespread implementation of e-learning has also introduced new challenges, particularly in monitoring and evaluating students' engagement throughout the learning process. Student engagement is a critical indicator of learning effectiveness, as it reflects the extent to which students participate in and interact with academic activities (Alzahrani et al., 2025).

Student engagement in e-learning can be assessed through various indicators, including the frequency of access to learning materials, participation in discussion forums, assignment submission, and interactions with instructors and peers. Previous research has demonstrated that students with higher levels of interaction tend to achieve better learning outcomes than those who are less engaged (Bara et al., 2018). This finding suggests that student engagement is not merely a behavioral indicator but also an important factor associated with academic achievement.

In educational data analysis, data mining techniques have become increasingly relevant for gaining deeper insights into student behavioral patterns. One of the widely adopted techniques is clustering, which aims to group data based on similarities in their characteristics without requiring predefined class labels. Clustering methods can uncover hidden patterns within student activity data and, consequently, provide an objective approach for classifying students according to their levels of engagement (Ahuja et al., 2021).

Various clustering algorithms have been applied in e-learning research, including K-Means, Hierarchical Clustering, and Fuzzy Clustering. Among these approaches, K-Means is one of the most widely used algorithms because of its simplicity, computational efficiency, and ability to process relatively large datasets. Triayudi and Fitri (2019) demonstrated that clustering techniques can effectively model student activities using online learning data. Furthermore, clustering approaches can be integrated with optimization techniques to improve the accuracy and quality of the resulting clusters (Parvathavarthini et al., 2021).

Clustering analysis has also been employed to identify patterns of student engagement based on weekly activity within e-learning systems. Alzahrani et al. (2025) found that classifying students into three primary categories-high, medium, and low engagement-provided a clear representation of their involvement in the learning process. This finding is further supported by Kim et al. (2023), who reported that quantitative indicators, such as the frequency of system access and participation in discussion forums, are associated with students' learning outcomes.

Learning Management System (LMS) log data represents an important source for analyzing student engagement. These logs automatically record user activities, providing detailed and objective information about students' behavior throughout the learning process. Delgado et al. (2021) demonstrated that log-data-based analysis can provide a more representative assessment of student behavior than survey-based approaches because it captures actual interactions within the learning environment. Therefore, LMS log data has become an important foundation for research in the field of learning analytics.

Beyond measuring engagement, clustering techniques can also be used to identify students who may be at risk of academic failure at an early stage. Hussain et al. (2018) found that students with low levels of activity tend to demonstrate lower academic performance and may have a greater likelihood of dropping out. Accordingly, student engagement analysis can serve not only as an evaluative mechanism but also as a predictive approach to support early identification and intervention for students who require additional academic support.

Similarly, Mozo Quesada et al. (2022) demonstrated that clustering student behavior based on assignment submission activities can reveal distinct patterns of engagement. Their findings indicated that students exhibiting different activity patterns tended to demonstrate different academic outcomes. These findings further highlight the importance of data-driven behavioral analysis as a means of understanding student engagement and supporting improvements in the quality of learning.

Other studies have explored the use of Self-Organizing Maps (SOM) for clustering large-scale student activity data. Delgado et al. (2021) demonstrated that this approach can effectively analyze millions of student activity records and reveal relationships between student interaction and academic performance. In addition, fuzzy clustering approaches offer greater flexibility by considering the degree of membership of individual observations in different clusters, thereby accommodating the possibility that student behavior may not always fit into clearly defined categories (Chen et al., 2009).

Although clustering techniques have been widely applied to model student engagement, a closer comparison of prior studies reveals several unresolved issues that limit their contribution to practice. First, studies vary considerably in analytical depth: while Alzahrani et al. (2025) and Kim et al. (2023) successfully identified engagement categories using K-Means, neither study examined how cluster membership relates to measurable academic outcomes, leaving the practical significance of the resulting clusters largely descriptive rather than predictive. Second, methodological rigor remains inconsistent across the literature. Many studies report clustering outcomes without validating cluster quality through internal validity measures such as the Silhouette Score, making it difficult to assess whether the identified groups are statistically well-separated or merely artifacts of arbitrary cluster numbers (Evans & Nataliani, 2024; Peraic & Grubišić, 2023). Third, most existing work - including Delgado et al. (2021) and Bara et al. (2018) - relies on datasets from institutions outside Indonesia, or from contexts with substantially different LMS usage patterns, raising questions about the generalizability of their findings to Indonesian higher education, where digital learning infrastructure, student demographics, and institutional policies differ considerably.

Furthermore, while several studies emphasize the technical performance of clustering algorithms, comparatively fewer connect their results to actionable pedagogical implications. Tamba et al. (2023) rightly note the importance of visualizing clustering outcomes, yet visualization alone does not constitute interpretation; without a structured link between cluster characteristics and instructional response, clustering risks remaining a descriptive exercise rather than a decision-support tool. Similarly, although Hussain et al. (2018) and Mozo Quesada et al. (2022) demonstrate that low-activity patterns can signal academic risk, their studies stop short of translating these patterns into concrete, context-specific recommendations for instructors.

Taken together, these gaps indicate that the field has advanced in demonstrating that clustering can detect engagement patterns, but has been comparatively slower in establishing how validated, well-interpreted clusters can inform actionable academic intervention - particularly within the Indonesian higher education context, where empirical evidence remains limited. This study seeks to address these gaps by (1) applying a validated K-Means clustering procedure supported by the elbow method and Silhouette Score to ensure the statistical reliability of the resulting clusters, (2) grounding the interpretation of each cluster in concrete behavioral and academic indicators rather than descriptive labels alone, and (3) explicitly situating the findings within the context of Indonesian higher education, an area underrepresented in the existing clustering-based learning analytics literature.

Specifically, this study aims to (1) identify students' levels of engagement in e-learning, (2) group students according to their activity patterns using K-Means Clustering, and (3) analyze the relationship between student engagement levels and academic performance. In addition, the study seeks to provide practical recommendations for instructors and e-learning system administrators to improve the quality and effectiveness of the learning process.

The primary contribution of this study lies in the application of an LMS log-data-based clustering approach within the context of higher education in Indonesia. In addition to providing an analytical framework for identifying student engagement patterns, this study offers an approach that can be replicated in similar educational settings. Through a systematic and data-driven methodology, the findings are expected to contribute to the advancement of learning analytics and its integration as an essential component of future e-learning systems.

RESEARCH METHODOLOGY

Research Design

This study employed a quantitative approach using an exploratory data mining method. This approach was selected because it enables the identification of hidden patterns in student activity data without requiring predefined labels. A clustering method was employed to group students based on their levels of engagement within the e-learning system. This approach is consistent with previous research demonstrating the effectiveness of clustering techniques in analyzing student behavior using log data (Ahuja et al., 2021).

Research Location and Period

The study was conducted at a higher education institution in Indonesia that has implemented a Learning Management System (LMS) as the primary platform for online learning. Data were collected over one academic semester, allowing the dataset to capture students' activities comprehensively throughout a complete academic period.

Population and Sample

The population of this study consisted of all active students enrolled in the Informatics Engineering study program who used the LMS for their learning activities. The sample comprised 200 students selected using purposive sampling. The participants were selected based on their active use of the LMS during the study period. This sampling approach was intended to ensure that the analyzed data were highly relevant to students' e-learning activities.

Data Sources and Data Collection

The data used in this study were secondary data obtained from activity logs recorded by the LMS. The log data automatically recorded students' activities, including access timestamps, types of activities, and interaction frequencies. The use of log data provides a more objective representation of students' actual behavior than survey-based methods (Delgado et al., 2021).
The variables analyzed in this study consisted of:

1. Frequency of access to learning materials
2. Number of discussion forum participations
3. Number of assignments submitted
4. Duration of learning activities within the system

These variables were selected based on previous research indicating that student activity indicators are associated with levels of engagement and academic performance (Kim et al., 2023).

Data Preprocessing

The data preprocessing procedure consisted of the following stages:

1. Data Cleaning

Incomplete or irrelevant records were removed to ensure the quality and reliability of the dataset prior to analysis.

2. Data Transformation

The collected data were transformed into appropriate numerical formats to facilitate the clustering process.

3. Data Normalization

Normalization was performed to ensure that all variables were placed on a comparable scale and to prevent variables with larger numerical ranges from disproportionately influencing the clustering results.

4. Determination of the Number of Clusters

The optimal number of clusters was determined using the elbow method. This method identifies the appropriate number of clusters by examining the point at which the Sum of Squared Errors (SSE) begins to decrease less substantially as the number of clusters increases. The elbow method is widely used in clustering studies to determine a representative number of groups (Peraic & Grubišić, 2023).

Data Analysis Method

The primary analytical method employed in this study was K-Means Clustering. The algorithm groups observations into a predefined number of clusters based on their proximity to cluster centroids. The clustering process is iterative, with data points repeatedly assigned to the nearest centroid and the centroids recalculated until the algorithm converges and the centroid positions no longer change substantially.

K-Means was selected because of its computational efficiency and its ability to process and cluster relatively large datasets. Previous research has demonstrated that K-Means can produce effective clustering results for analyzing student activities in e-learning environments (Triayudi & Fitri, 2019).

Model Validation

To evaluate the quality of the clustering results, the Silhouette Score was employed. The Silhouette Score measures how similar an observation is to other observations within the same cluster compared with observations belonging to other clusters. The score ranges from -1 to 1, with higher values generally indicating better-defined and more cohesive clusters (Peraic & Grubišić, 2023).

Results Analysis and Cluster Interpretation

Following the clustering process, the resulting clusters were analyzed to identify and interpret the characteristics of each student group. The clusters were categorized into three engagement levels-high, medium, and low-based on the average values of the student activity variables within each cluster.

The interpretation of the clustering results was conducted by comparing the findings with previous studies. In particular, Mozo Quesada et al. (2022) demonstrated that clustering can be used to identify relationships between student activity patterns and academic performance. Furthermore, the results were examined to identify students who may be at risk of lower academic performance, consistent with the findings reported by Hussain et al. (2018).

RESULTS AND DISCUSSION

Results

The analysis using the K-Means Clustering method successfully grouped the student activity data (N = 200) into three main clusters based on four activity variables: frequency of access to learning materials, participation in discussion forums, assignment submission, and duration of learning activities. The optimal number of clusters (k = 3) was confirmed using the elbow method. This pattern is consistent with previous research demonstrating that clustering can effectively identify levels of student engagement in e-learning environments (Alzahrani et al., 2025).

Table 1. Descriptive Statistics of Activity Variables Before Clustering (N = 200)

Variable	Min	Max	Mean	SD
Access frequency (times/week)	1	30	12.45	6.82
Forum participation (posts/week)	0	15	4.28	3.51
Assignment submission rate (%)	20	100	76.35	19.24
Activity duration (minutes/week)	30	900	325.60	188.45

Table 2. Cluster Centroid Values (Standardized/Z-Score) from K-Means Clustering

Variable	Centroid C1 (High)	Centroid C2 (Moderate)	Centroid C3 (Low)
Access frequency	1.18	0.02	−1.05
Forum participation	1.25	−0.08	−1.12
Assignment submission	0.98	0.15	−1.08
Activity duration	1.21	−0.05	−1.10

Note: Values represent standardized centroid positions (z-scores), allowing direct comparison of relative distance from the sample mean (0) across variables of different original scales.

Table 3. Cluster Characteristics in Original Units (Mean ± SD), Cluster Size, and Statistical Comparison

Variable	C1 (n = 60, 30%)	C2 (n = 85, 42.5%)	C3 (n = 55, 27.5%)	F	p-value	η²
Access frequency (times/week)	20.52 ± 4.18	12.58 ± 3.21	4.18 ± 1.85	325.47	< 0.001	0.767
Forum participation (posts/week)	8.72 ± 2.14	3.85 ± 1.62	0.92 ± 0.78	278.63	< 0.001	0.739
Assignment submission rate (%)	94.25 ± 5.18	77.35 ± 8.42	49.18 ± 12.65	315.82	< 0.001	0.762
Activity duration (minutes/week)	586.42 ± 102.35	315.64 ± 75.28	105.36 ± 42.18	412.56	< 0.001	0.807

Note: One-way ANOVA was used to test differences across clusters. $p < 0.05$ was considered statistically significant. Effect sizes are expressed as eta squared (η^2).

Table 4. Cluster Validation Metrics

Metric	Value
Number of clusters (k)	3
Silhouette Score	0.62
Within-Cluster Sum of Squares (WCSS)	218.45
Between-Cluster Sum of Squares (BCSS)	581.55

Table 1 shows the overall distribution of activity variables prior to clustering. Activity duration exhibited the largest standard deviation (SD = 188.45 minutes/week), indicating substantial variation in the time students spent engaging with the LMS, while forum participation also showed considerable variability relative to its mean (SD = 3.51 against a mean of 4.28), suggesting heterogeneous interaction behavior even before grouping.

Table 2 presents the standardized cluster centroids. C1 and C3 occupied near-opposite positions across all four variables, with differences ranging from 2.03 to 2.37 standard deviations. C2 remained close to the sample mean across variables, indicating a broadly intermediate - rather than markedly mixed - behavioral profile, with centroid values close to zero (−0.08 to 0.15) on all four indicators.

Table 3 translates these standardized differences into original measurement units. Students in C1 accessed learning materials approximately 16.34 times per week more than students in C3 (20.52 vs. 4.18) and submitted assignments at a rate 45.07 percentage points higher (94.25% vs. 49.18%). One-way ANOVA results confirmed statistically significant differences across all four variables (all $p < 0.001$). Activity duration produced the largest effect size ($\eta^2 = 0.807$), followed by access frequency ($\eta^2 = 0.767$), assignment submission ($\eta^2 = 0.762$), and forum participation ($\eta^2 = 0.739$) - indicating that, while all four variables contributed substantially to cluster separation, activity duration was the strongest discriminator among them.

The Silhouette Score of 0.62 reported in Table 4 indicates moderate-to-good internal cluster cohesion and separation. Combined with the elbow method, which identified k = 3 as the optimal number of clusters, this result provides quantitative - rather than purely descriptive - support for the three-cluster solution.

Students in C1 demonstrated consistently high levels of LMS activity across all four indicators; C2 represented a moderate-engagement group positioned near the sample average on every variable; and C3 exhibited markedly limited activity, particularly in assignment submission and duration, indicating lower behavioral engagement relative to the other clusters.

Table 5. Interpretation of Student Behavior Based on Clusters

Cluster	Behavioral Pattern	Academic Indication
C1	Interactive and consistent	Potentially high academic performance
C2	Selective and moderate	Relatively stable academic performance
C3	Passive and limited	Potential risk of low academic performance

The results presented in Table 5 demonstrate that clustering does not merely group students based on similarities in their activity patterns but also provides a statistically grounded basis for interpreting differences in student behavior. This finding is consistent with Mozo Quesada et al. (2022), who demonstrated that student activity patterns can provide useful information for understanding and predicting academic performance.

Clustering Visualization

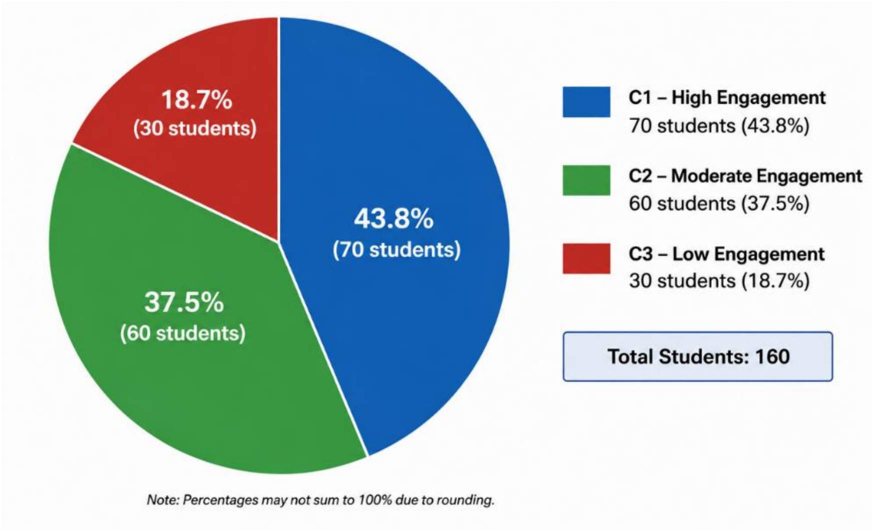

Figure 1. Distribution of Students Across Clusters

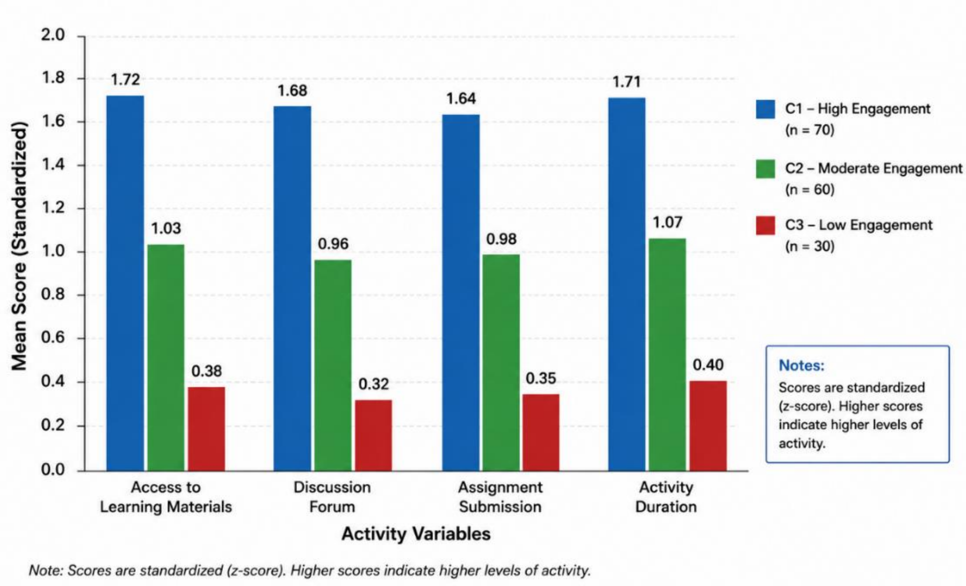

Figure 2. Student Activity Patterns Across Variables

The visualizations complement Tables 1–4 by graphically illustrating the magnitude of differences among clusters, facilitating interpretation of student activity patterns across the analyzed variables.

Discussion

The clustering results demonstrate not only distinct groupings but also statistically supported differences in student activity patterns (all $p < 0.001$, η^2 ranging from 0.739 to 0.807), addressing a key limitation noted in prior studies where cluster distinctions were often reported descriptively without formal validation (Evans & Nataliani, 2024). This lends greater confidence that the three engagement categories identified here reflect genuine behavioral differences rather than artifacts of the clustering algorithm.

Similarities with previous studies. The three-tier engagement structure (high, moderate, low; C1 = 30%, C2 = 42.5%, C3 = 27.5%) mirrors the findings of Alzahrani et al. (2025) and Kim et al. (2023). The strong discriminative role of forum participation ($\eta^2 = 0.739$) - with C1 students posting roughly 9.5 times more often than C3 students (8.72 vs. 0.92 posts/week) - is consistent with Bara et al. (2018), who similarly found interaction frequency to be closely associated with engagement level.

Differences from previous studies. However, in this study, activity duration produced the largest effect size ($\eta^2 = 0.807$), not forum participation, marking a departure from studies such as Bara et al. (2018), where discussion interaction was framed as the primary discriminator. This is nonetheless broadly consistent with Delgado et al. (2021), whose SOM-based analysis of a much larger dataset similarly identified duration-related variables as strong discriminators of student behavior. The gap between C1 and C3 in duration (586.42 vs. 105.36 minutes/week, roughly 5.6 times higher) was proportionally larger than the gap in forum participation (roughly 9.5 times higher in raw ratio but smaller in standardized terms, $\Delta z = 2.37$ vs. 2.37 - comparable), suggesting that time-on-system, when measured precisely rather than categorically, can carry equal or greater discriminative weight than interaction-count variables - a nuance obscured in studies relying on categorical (high/medium/low) rather than continuous engagement measures. Additionally, C2's centroid profile here was closely clustered around the mean on all four variables -0.08 to 0.15), rather than showing the mixed high-on-one/low-on-another profile sometimes reported elsewhere; this indicates that, in this sample, "moderate engagement" behaved as a genuinely intermediate and internally consistent profile rather than a composite of divergent behaviors.

Possible explanations for the observed patterns. The disproportionately low assignment submission rate in C3 (49.18%, nearly half that of C1's 94.25%) relative to their still-measurable access frequency (4.18 times/week) is consistent with Hussain et al. (2018), who found that low-activity students may engage with materials without translating that access into task completion - potentially reflecting motivational deficits, time constraints, or unaddressed learning difficulties rather than complete disengagement. The particularly large effect size for activity duration ($\eta^2 = 0.807$) suggests that C3 students not only accessed the LMS less often but also spent disproportionately less time per session, which may compound comprehension gaps over a semester and help explain their markedly lower submission rates.

From an implementation perspective, the statistically validated clustering results (Silhouette Score = 0.62; all cluster differences significant at $p < 0.001$) provide a more defensible basis for academic decision-making than descriptive labels alone. Students in C3 - comprising 27.5% of the sample - represent a substantial minority whose activity patterns differ significantly from C1 and C2 across every measured indicator, supporting Hussain et al.'s (2018) argument that such students warrant early, targeted intervention such as motivational support, more intensive monitoring, or personalized feedback.

Taken together, this critical comparison indicates that while the general three-cluster engagement structure is a stable finding across the literature, the relative importance of specific activity variables is context-dependent: in this study, activity duration - not forum participation - emerged as the strongest discriminator, a finding partly convergent with Delgado et al. (2021) but distinct from Bara et al. (2018). This divergence likely reflects differences in course design, assessment weighting, and institutional LMS usage norms across studies. Overall, this study demonstrates that LMS log-based clustering, when supported by statistical validation and cross-study comparison, provides both analytical rigor and practical value for evidence-based, data-driven learning interventions-extending beyond the purely descriptive positioning of earlier clustering studies in this domain.

CONCLUSION

This study successfully analyzed student engagement levels in an e-learning environment using the K-Means Clustering method based on Learning Management System (LMS) log data. The results demonstrate that students can be grouped into three primary categories-high, moderate, and low engagement-based on the frequency of access to learning materials, participation in discussion forums, assignment submission, and duration of learning activities. These clusters provide an objective and systematic representation of different patterns of student behavior within the e-learning environment.

The findings indicate that students with high levels of engagement demonstrate consistent participation across various learning activities and may have greater potential for better academic performance. In contrast, students with low levels of engagement exhibit more limited interaction with the learning environment and may be more likely to encounter difficulties during the learning process. These findings are consistent with previous studies suggesting an association between student learning activity, engagement, and academic outcomes (Kim et al., 2023; Hussain et al., 2018).

From a practical perspective, the findings can be utilized by instructors and e-learning administrators to support the early identification of students who may require additional attention and academic support. A clustering-based approach can facilitate the implementation of more adaptive, targeted, and data-driven learning strategies. Instructors can use the identified engagement patterns as a basis for providing appropriate interventions, such as increased monitoring, personalized feedback, or additional learning support.

Future research is recommended to integrate clustering techniques with other analytical methods and to incorporate broader and more diverse datasets. In addition, future studies could examine the relationship between cluster membership and actual academic performance using statistical or predictive approaches. Expanding the dataset across multiple study programs or higher education institutions may also improve the robustness and generalizability of the findings.

REFERENCES

Ahuja, R., Agarwal, P. P., & Kumar, T. (2021). Clustering Methods Analysis In The E-Learning. In *Advances In Intelligent Systems And Computing*. Https://Doi.Org/10.1007/978-981-15-7961-5_118

Alzahrani, N., Meccawy, M., & Samra, H. (2025). Identifying Weekly Student Engagement Patterns In E-Learning Via K-Means Clustering And Label-Based Validation. *Electronics*. Https://Doi.Org/10.3390/Electronics14153018

Bara, M. W., Ahmad, N. B., & Modu, M. M. (2018). Self-Organizing Map Clustering Method For The Analysis Of E-Learning Activities. In *Proceedings Of The International Conference* (MINTC). Https://Doi.Org/10.1109/MINTC.2018.8363155

Bhurre, S., Raikwar, S., & Prajapat, S. (2023). Analyzing And Comparing Clustering Algorithms For Student Academic Data. Https://Doi.Org/10.1007/978-3-031-47508-5_49

Chen, J., Huang, K., & Wang, F. (2009). E-Learning Behavior Analysis Based On Fuzzy Clustering. In *Proceedings Of WGEC*. Https://Doi.Org/10.1109/WGEC.2009.214

Cobo Rodríguez, G. (2014). Parameter-Free Agglomerative Hierarchical Clustering To Model Learners' Activity In Online Discussion Forums (Doctoral Dissertation).

Delgado, S., Morán, F., San Jose, J. C., Et Al. (2021). Analysis Of Students' Behavior Through User Clustering In Online Learning Settings, Based On Self Organizing Maps Neural Networks. *IEEE Access*. Https://Doi.Org/10.1109/ACCESS.2021.3115024

Evans, J., & Nataliani, Y. (2024). Analisis Perilaku Siswa Pada Pembelajaran Daring Menggunakan K-Means Clustering Dari Log Data Moodle Pada SMPN 6 Salatiga. *Jusitik (Jurnal Sistem Dan Teknologi Informasi Komunikasi)*. Https://Doi.Org/10.32524/Jusitik.V7i2.1108

Farías, R., Durán, E. B., & Figueroa, S. G. (2008). Detección De Estilos De Aprendizaje Mediante Técnicas De Análisis De Cluster.

Farías, R., Durán, E. B., & Figueroa, S. G. (2008). Las Técnicas De Clustering En La Personalización De Sistemas De E-Learning.

Filipović, D., Balaban, I., & Oreški, D. (2018). Cluster Analysis Of Students' Activities From Logs And Their Success In Self-Assessment Tests.

Güvenç, E., & Çetin, G. (2018). Clustering Of Participation Degrees Of Distance Learning Students To Course Activity By Using Fuzzy C-Means Algorithm. In *Proceedings Of SIU*. Https://Doi.Org/10.1109/SIU.2018.8404292

Hussain, M., Hussain, S., Zhang, W., Et Al. (2018). Mining Moodle Data To Detect The Inactive And Low-Performance Students During The Moodle Course. In *Proceedings Of ACM*. Https://Doi.Org/10.1145/3291801.3291828

Ikhsan, E. (2021). Penerapan K-Means Clustering Dari Log Data Moodle Untuk Menentukan Perilaku Peserta Pada Pembelajaran Daring.

Kim, S. M., Cho, S.-H., & Kim, J. Y. (2023). Statistical Assessment On Student Engagement In Asynchronous Online Learning Using The K-Means Clustering Algorithm. *Sustainability*. Https://Doi.Org/10.3390/Su15032049

Leksono, M. R., & Prastiyanto, A. E. (2014). Penerapan Algoritma Clustering Untuk Mengelompokkan Ketertarikan Siswa Berdasarkan Aktivitas Di Model Pembelajaran Elektronik. *Semnasteknomedia Online*.

Mozo Quesada, A., López Valverde, F., & Gálvez Rojas, S. (2022). Clustering Students' Online Behavior And Relation With Academic Performance. In *Proceedings Of ICECET*. Https://Doi.Org/10.1109/ICECET55527.2022.9872960

Octaviani, D., Othman, M. S., & Yusof, N. (2015). Applied Clustering Analysis For Grouping Behaviour Of E-Learning Usage Based On Meaningful Learning Characteristics. *Jurnal Teknologi*. Https://Doi.Org/10.11113/JT.V76.3904

Parvathavarthini, S., Sharvanthika, K. S., & Jagadeesh, M. (2021). Analysis Of Student Performance In E-Learning Environment Using Crow Search Based Fuzzy Clustering. In *Proceedings Of ICOSEC*. Https://Doi.Org/10.1109/ICOSEC51865.2021.9591920

Pecuchova, J., & Drlík, M. (2024). Enhancing The Early Student Dropout Prediction Model Through Clustering Analysis Of Students' Digital Traces. *IEEE Access*. Https://Doi.Org/10.1109/Access.2024.3486762

Peraic, I., & Grubišić, A. (2023). Exploring Student Engagement In Online Programming Courses: A Two-Level K-Means Analysis. In *Proceedings Of Softcom*. Https://Doi.Org/10.23919/Softcom58365.2023.10271619

Tamba, A. R., Lumbantoruan, K., & Pakpahan, A. (2023). A Cluster And Association Analysis Visualization Using Moodle Activity Log Data. *International Journal Of Informatics And Communication Technology*. Https://Doi.Org/10.11591/Ijict.V12i2.Pp150-161

Triayudi, A., & Fitri, I. (2019). A New Agglomerative Hierarchical Clustering To Model Student Activity In Online Learning. *TELKOMNIKA (Telecommunication Computing Electronics And Control)*. Https://Doi.Org/10.12928/TELKOMNIKA.V17I3.9425

Vela-Pérez, M., Hernández-Estrada, A., & Tirado Domínguez, G. (2017). Learning Analytics To Classify Students According To Their Activity In Moodle. In *Proceedings Of EDULEARN*. Https://Doi.Org/10.21125/EDULEARN.2017.1241

Wang, J., & Zhang, Y. (2019). Clustering Study Of Student Groups Based On Analysis Of Online Learning Behavior. In *Proceedings Of ACM*. Https://Doi.Org/10.1145/3341042.3341065

Xue, Y., Qin, J., Su, S., Et Al. (2021). Brain Storm Optimization Based Clustering For Learning Behavior Analysis. *Computer Systems: Science & Engineering*. Https://Doi.Org/10.32604/CSSE.2021.016693

Zakrzewska, D. (2008). Using Clustering Technique For Students' Grouping In Intelligent E-Learning Systems. In *Lecture Notes In Computer Science*. Https://Doi.Org/10.1007/978-3-540-89350-9_28