CIVITAS
Conference on International Visions and Transformative Action for Society

Civitas Proceedings
Pancasila and Civic Education Program Study
Faculty of Education and Teacher Training
Universitas Nusantara PGRI Kediri, Indonesia
Volume 1 2026

A Clustering-Based Framework for Student Academic Performance Analysis and Learning Recommendations

Fardan Zamakhsyari * [1, a], Nadhira Fasya Salsabila [2, b], Rizka Suhana [1, c], Sean Ryan Al Dziqri [1, d]

[1] Universitas Cahaya Surya, Ngronggo, Kediri, 64129, Indonesia
[2] Universitas Negeri Yogyakarta, Sleman, Yogyakarta, 55281, Indonesia
[a] fardanzam@ucs.ac.id *; [b] nadhirafasya.2023@student.uny.ac.id; [c] rsuhana@ucs.ac.id; [d]seanryan1339@gmail.com
* Corresponding Author

Abstract: Student academic data contains valuable information that can be used to support improved learning and decision-making in the field of education. However, conventional academic evaluations often rely on aggregate indicators such as the Cumulative Grade Point Average (GPA), which may not adequately represent the student's learning characteristics and academic needs. This research proposes a clustering-based system to analyze students' academic achievement and generate personalized learning recommendations. The academic records of first-semester students are collected and processed first through feature engineering, by categorizing courses into numerical, practical, and theoretical domains. Three clustering algorithms, namely K-Means, Agglomerative Clustering, and Gaussian Mixture Model (GMM), were evaluated using the Silhouette Score, Davies-Bouldin Index, and Calinski-Harabasz Index. The results showed that K-Means achieved the best performance with a Silhouette Score of 0.4926, the Davies-Bouldin Index of 0.5085, and the Calinski-Harabasz Index of 48.6175. The selected model resulted in three meaningful clusters, representing high, medium, and low-performing students. Based on the results of the grouping and learning profiles, a recommendation framework consisting of nine personalized learning strategies was developed and validated through expert assessment. The proposed system demonstrates the potential of integrating academic performance clustering with learning recommendations to support personalized learning, early academic intervention, and data-driven efforts to improve educational quality.

Keywords: Educational Data Mining; K-Means Clustering; Academic Performance; Personalized Learning; Learning Recommendation

INTRODUCTION

The use of academic data is one of the important aspects in supporting the improvement of the quality of learning in higher education. Educational institutions regularly collect various student data, such as course scores, evaluation results, and learning activities (Rizki et al., 2024). The data can be used to identify learning patterns, monitor academic development, and support data-driven decision-making. However, academic evaluation still relies heavily on aggregate indicators such as the Cumulative Achievement Index (GPA) or final grades of courses that are often not able to comprehensively describe the characteristics, strengths, and learning weaknesses of each student (Dewi et al., 2022).

The development of Educational Data Mining (EDM) and Learning Analytics allows the use of machine learning techniques to gain a better understanding of students' academic performance (Azzahra & Sriani, 2025). As part of the broader application of Artificial Intelligence (AI) in education, these techniques provide opportunities to transform academic data into meaningful information that can support data-driven educational decision-making. One of the widely used methods is clustering, which is an unsupervised learning technique that groups students based on similar characteristics (Makawana & Mehta, 2023). In the context of education, clustering can help institutions identify groups of students with different patterns of academic achievement so that learning interventions can be carried out in a more targeted manner (Li & Li, 2025). Such data-driven and targeted interventions have the potential to support personalized learning and contribute to efforts to strengthen the quality of education.

Several studies have shown the effectiveness of clustering in the field of education. Azzahra and Sriani (2025) uses the K-Means algorithm to group students based on academic achievement and shows that the results of clustering can be used as the basis for learning evaluation. Kinjal et al. (2024) utilize K-Means to group courses based on difficulty level and generate course recommendations for students. In addition, Sharif et al. (2024) developed a personalized learning system that combines clustering and recommendation systems to identify student learning patterns and support more targeted learning interventions.

Nonetheless, some limitations were still found in previous studies. Most of the research focuses on the grouping process without linking the clustering results with learning recommendations that can be directly utilized by students and educational institutions (Kinjal et al., 2024). In addition, the implementation of the preprocessing and validation stages of clusters is often not comprehensively explained, even though these two aspects affect the quality of clustering results (Villegas-ch et al., 2024). On the other hand, research that integrates academic achievement clustering with a learning recommendation system based on student profiles is still relatively limited (Ouassif et al., 2025).

Based on these problems, this study proposes a clustering system for the analysis of students' academic achievement using K-Means, Agglomerative Clustering, and Gaussian Mixture Model algorithms. The system developed includes the stages of data preprocessing, determination of the optimal number of clusters using the Elbow Method, and evaluation of cluster quality. The clustering results are then used as a basis to produce learning recommendations that are in accordance with the characteristics of each student group.

The main contribution of this research is the development of a framework that integrates academic achievement analysis, clustering validation, and learning recommendations into a unified system. In contrast to previous studies that primarily focus on student clustering or recommendation generation, this study combines academic performance levels with learning profiles derived from Numeric, Practical, and Theoretical course categories. This integration produces nine personalized learning recommendation scenarios by combining three academic performance levels (High, Moderate, and Low) with three learning profiles (Numeric, Practical, and Theoretical). Thus, this study not only provides information about students' academic achievement groups but also generates more targeted recommendations that can be used as a basis for academic interventions and personalized learning in higher education.

METHODS

This study was conducted as an exploratory quantitative study using first-semester student academic records as the primary dataset. The study explores the application of clustering algorithms to identify patterns in students' academic performance and to develop learning recommendations based on the resulting clusters and student learning profiles. Given the relatively limited dataset, the proposed framework is intended as an initial exploration of the integration between academic performance clustering and personalized learning recommendations.

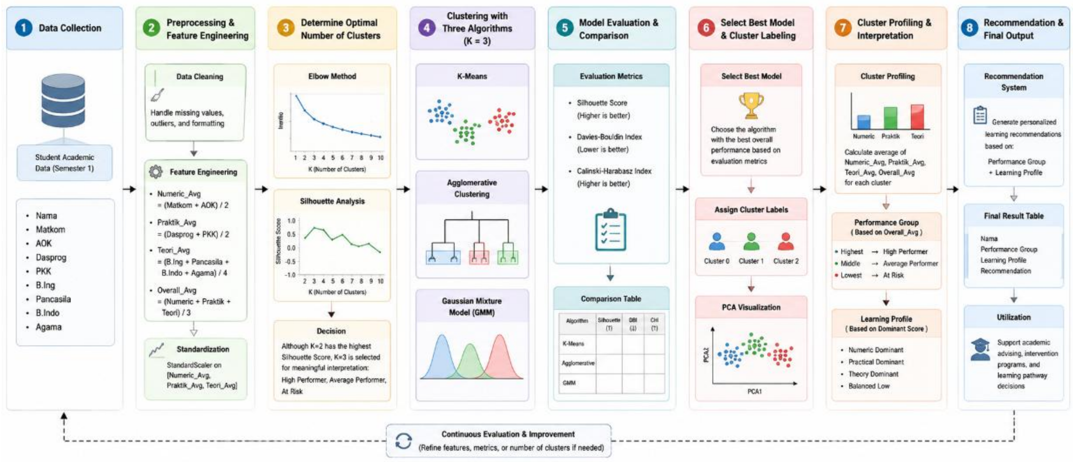

Figure 1. Research Flow Methodology

There are several stages that are passed in this study, based on Figure 1. This research begins with the data collection process, followed by pre-processing of data so that it can be processed optimally in the clustering process. Before entering the clustering stage, it is necessary to determine the optimal number of clusters, and only then process it at the clustering stage. The results of the clustering process will be evaluated to see the performance of each algorithm. And finally, create a recommendation system based on clusters and student profiling based on their grades.

Data Collecting

The quality and accuracy of the data used often determine the success of research. Therefore, the first important step is data collection (Latief et al., 2024). In this study, data was collected from the final exam results of the first semester of Cahaya Surya University students. The grades of all courses of each student are collected manually from the database.

Pre-processing Data

For the clustering and recommendation process to run properly, data needs to go through the data pre-processing stage. This stage aims to convert the raw data into a suitable format so that it can be properly processed by the deep learning model (Zamakhsyari & Fatwanto, 2025). In this process, the data will be categorized according to the type of course, which will be divided into three categories, namely Numeric, Theoretical, and Practic. By going through this process, the application of the clustering algorithm process to classify students based on their final grades will be more efficient (Suhana et al., 2022). Through data analysis, we can gain practical insights and suggestions to improve student engagement and learning processes (Sharif et al., 2024).

Clustering: Methods and Behavioral Analysis

Before starting the clustering process, a model configuration is required to determine the appropriate value for the clustering model. In this study, the Elbow Method is used to determine the appropriate number of clusters by evaluating the Sum of Squared Errors (SSE) across different values of K based on the students' academic features derived from the Numeric, Practical, and Theoretical course categories. The optimal number of clusters is identified by observing the point at which the decrease in SSE begins to diminish. In addition, the Silhouette Score is used to evaluate cluster cohesion and separation and to support the selection of an appropriate number of clusters. The number of squared errors (SSE) is displayed along with a reasonable number of cluster values (Syakur et al., 2018). The value chosen at the apex of the graph curve is called the K value. In addition, the silhouette score will be used to measure the quality of the cluster by assessing how similar an object is to its own cluster compared to other clusters (Danu et al., 2025).

Clustering is a grouping model based on its similarity value (Kinjal et al., 2024). The clustering models used are K-Means Clustering, Agglomerative Clustering, and Gaussian Mixture Model (GMM). K-Means Clustering is used because based on several previous studies, this model has high accuracy and efficiency and is suitable for small to medium data (Utami et al., 2024). K-Means Clustering forms groups whose members have a high degree of similarity, which are measured based on proximity to the same central point (Ouassif et al., 2025). The model that will be used as a comparison is Agglomerative Clustering. This model works with the bottom-up principle, where each data point is initially considered a single cluster. This algorithm will iteratively combine the two most similar clusters into one larger cluster until one single cluster remains (Hakim et al., 2024). In addition, this study also uses the Gaussian Mixture Model (GMM) as a comparison. This model is based on the assumption that there is a set of Gaussian distributions, each of which represents a group of observations (Ouassif et al., 2025).

Evaluation Clustering Model

After the clustering process uses three different models, the next process is to evaluate the results of the clustering to see how each model performs (Maguate et al., 2024). In this study, the evaluation process will use three models, namely the Silhouette Score, the Davies-Bouldin Index, and the Calinski-Harabasz Index. Silhouette Score is used to measure the quality of a cluster by assessing how similar an object is to its own cluster compared to other clusters. Then, the Davies-Bouldin Index is a matrix used to measure the quality and validation of clusters. This method works by evaluating the level of cohesion (the distance between points in the same cluster) and the degree of separation (the distance between different clusters) (Dinata et al., 2020). And finally, the Calinski-Harabasz Index is an evaluation matrix that compares how similar an item is to its cluster (cohesion) with other objects in other clusters (separation) (Ullah et al., 2023).

Recommendation System Model

Using the clustering results, the researcher created a recommendation system that aims to provide suggestions on the learning process according to the specific needs of each cluster. Recommendation systems analyze users' preferences for an item and provide proactive suggestions on items that match their interests (Ouassif et al., 2025). The creation of this recommendation system will be based on the performance of each student and the profile of their course group, so that there is a combination of performance and student course profile. The recommendations given are based on independent learning theories and are intended for students to improve learning experiences and outcomes. To ensure objectivity, validity, and accuracy, the recommendations provided have also been validated by two educational experts evaluated the recommendation framework based on relevance, feasibility, and usefulness.

RESULTS AND DISCUSSION

Data Collecting and Pre-processing Data

Data collection is carried out manually by taking the student's final semester grade. The data that was successfully collected was 31 data from 1st semester students from a total of 8 courses. Each of these data will later be processed based on the course category that has been determined based on the dominant learning outcomes of each course and validated by two domain experts from the Computer Science department, which can be seen in Table 1.

Table 1. Course Categories

Category	Courses
Numeric	Computational Mathematics
	Computer Organization Architecture
Practice	Programming Basics
	Introduction to Computer Science
Teoritic	English
	Indonesian
	Civic Education
	Religious Study

In the pre-processing stage, there is a feature engineering process that adjusts the course to its category. After adjusting, the data in the value column will be converted into numeric values so that it can be processed by the system. Finally, the data will be normalized according to the values of each category.

Clustering: Methods and Behavioral Analysis

Before starting the clustering process, it is necessary to perform an analysis to determine the optimal number of clusters. The determination of this cluster is carried out through the Elbow Method approach, which involves calculating the SUM Square Error (SSE) value. For the elbow chart, a significant decrease is the optimal number of K. In addition, the determination of the number of clusters also uses the Silhouette Score approach to provide validation based on the Elbow Method analysis.

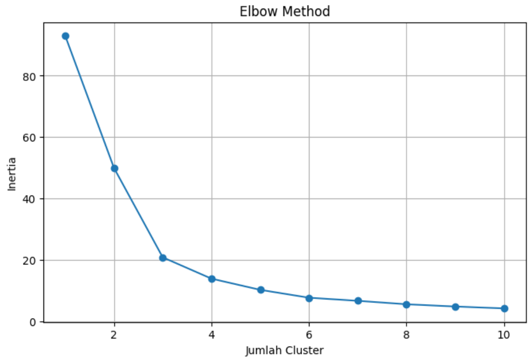

Figure 2. Elbow Method

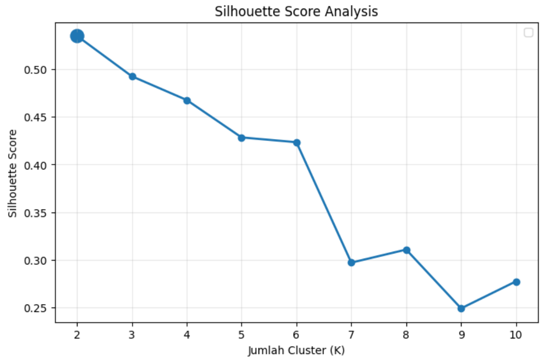

Figure 3. Silhouette Score Analysis

Figure 2 is the result of the analysis of the optimal number of clusters using the Elbow Method, and Figure 3 is the result of the Silhouette Score Analysis to validate the determination of the optimal number of clusters from the Elbow Method. The results of the Silhouette Score analysis showed that K=2 had the highest score with a score of 0.5350, as well as in the Elbow Model approach, K=2 got an inertia value of 49.9. The Silhouette Score analysis showed that K=2 achieved the highest score of 0.5350, while K=3 obtained a slightly lower score of 0.4926. Although K=2 provided the highest internal clustering validity based on the Silhouette Score, K=3 was selected by considering both cluster quality and the educational interpretability required by the objective of this study. Specifically, the use of three clusters enables the academic performance patterns to be represented as High, Moderate, and Low performance groups, which provides a more actionable basis for developing differentiated learning recommendations. Therefore, K=3 was not selected solely on the basis of the highest Silhouette Score, but as a balance between acceptable clustering quality and the practical interpretability of the clusters for personalized learning recommendations.

After the number of clusters is determined, the next process is clustering on the processed data. In this process, a comparison of three clustering algorithm models was carried out: K-Means, Agglomerative, and Gaussian Mixture Model. Each model has a different way of defining clusters. K-means determine cluster values based on similarities between data, while agglomeration is more hierarchical by considering each data to be a cluster. The Gaussian Mixture Model determines the value of clusters based on probability.

Table 2. Cluster Data Distribution of Each Model

Cluster	K-Means	Agglomerative	GMM
0	12	17	17
1	18	1	2
2	1	13	12

Based on the results of clustering, the results were obtained as shown in Table 2. Each model shows different data distribution values, according to the method used. These results will be continued in the evaluation process to determine the algorithm model that has the best performance.

Evaluation Clustering Model

The clustering results from each model has different values and data distribution, so it is necessary to evaluate each result of the clustering algorithm model to determine which model has the best performance. In this case, the evaluation process will use three different approaches, namely the Silhouette Score, the Davies-Bouldin Index, and the Calinski-Harabasz Index. This process is aimed at obtaining accurate results by using several evaluation models.

Table 3. Clustering Evaluation Results of each Algorithm model

Algoritma	Silhouette	Davies_Bouldin	Calinski_Harabasz
K-Means	**0.492609**	**0.508491**	**48.617529**
GMM	0.487592	0.669595	41.106316
Agglomerative	0.474108	0.531671	46.764983

Based on the results of the clustering model evaluation shown in Table 3, the K-means model excelled in all evaluation models. K-Means model shows a value of 0.492609 in the Silhouette Score and 48.617529 in the Calinski-Harabasz. Meanwhile, in the evaluation of the Davies-Bouldin model, the lowest evaluation value was taken, and K-Means got a value of 0.508491. From this data, it can be concluded that the K-Means model has the best performance in this study.

Recommendation System Model

In line with the results of clustering and evaluation, the processed data will be reprocessed as a guide to create a recommendation system. This system will assess student performance based on their grades, and also their learning profiles in accordance with numerical, theoretical, and practical categories.

Table 4. Results of Clustering of Student Performance by Grades

Cluster	Performa	Numeric_Avg	Praktik_Avg	Teori_Avg	Overall_Avg
0	High	83.89	84.17	82.44	83.50
1	Moderate	75.89	79.64	77.65	77.73
2	Low	53.50	65.50	65.75	61.58

Table 4 shows the results of clustering which are used as a reference in determining student performance, which is based on the average score of the course group and the overall average score.

The High Performance cluster shows that academically students excel and give good results, while Moderate Performance is a group of students with an overall average score, and Low Moderate is a group of students who are lacking in academics.

However, this research does not stop at this cluster alone. The researcher also profiled each student based on the grades obtained. This profiling will help in the classification of recommendations that are appropriate for students, not only based on academic achievement, but also based on students' tendencies to models and learning concentrations.

Table 5. Recommendations based on cluster results

Performa	Learning Profile	Recommendation	Key Activities	Theoritical Framework
High	Practical	Project based enrichment & peer mentoring	Coding challenge, independent project, peer tutoring	Experiential Learning (Kolb, 1984); Peer Learning
	Theoretical	Critical Thinking & Knowledge	Paper analysis & review, mind mapping, peer tutoring	HOTS — Bloom Taxonomy L4–6 (Anderson & Krathwohl, 2001)
	Numeric	Advanced Problem Solving	Competitive math, mathematical projects, concept mapping	Mastery Learning — Enrichment (Bloom, 1968); SRL (Zimmerman, 2002)
Moderate	Practical	Deliberate Practice & Reflective Coding	Repractice, journal debugging, core coding training.	Kolb (1984) ELT; Deliberate Practice (Ericsson, 1993)

	Theoretical	Active Reading & Concept Consolidation	SQ3R, digital flashcards, small group discussion.	Cognitive Load Theory (Sweller, 1988); Active Recall
	Numeric	Structured Practice & Formative Self-Assessment	Gradual training, mistake identification, tutorial video.	Mastery Learning — Corrective (Bloom, 1968; Guskey, 2007)
Low	Practical	Step-by-Step Guided Coding & Peer Support	Worked examples, retyping code, buddy system, small projects	Scaffolding (Vygotsky, 1978); Worked Example Effect (Sweller, 1988)
	Theoretical	Structured Reading Support & Oral Comprehension	Summary sheet, short videos, collective reading, daily glossary	Scaffolding (Vygotsky, 1978); UDL (CAST, 2018)
	Numeric	Foundational Scaffolding & Guided Remediation	Understanding core concept, animation video, mini daily target, peer/tutor	ZPD & Scaffolding (Vygotsky, 1978); Mastery Remediation (Bloom, 1968)

Using the clustering results, the researcher created a recommendation system shown in Table 5. This recommendation aims to provide advice for students to improve their independent learning experience according to the specific needs of each cluster. Recommendation systems analyze users' preferences for an item and provide proactive suggestions on items that match their interests (Ouassif et al., 2025). The main basis of the recommendations given is the theory of Self-Regulated Learning, which states that the success of independent learning is determined by the ability of students to manage three cognitive phases cyclically: Forethought (planning), Performance (implementation and monitoring), and Self-Reflection (self-evaluation) (Schunk & Zimmerman, 2022). Recent research verifies that structured SRL process patterns are significantly correlated with student achievement (Cheng et al., 2025). Based on this theory, High-Performance students are directed to carry out strengthening activities that require high-level thinking operations, Moderate Performance students, is recommended to focus on consolidation, reinforcement with structured exercises, and active reflection, while Low-Performance students need intensive strengthening of basic concepts with simpler, realistic, and structured learning targets.

For the High-Performance students in Numeric profile, recommendations focus on advanced computational mathematics and real-world problem-based mathematical modelling, which underpins the enrichment pathway in the theory of Mastery Learning (Betts et al., 2021; Stanny, 2016). In courses that emphasize practice, students can enhance the learning experience by developing self-paced programming projects and acting as peer mentors, this walks in line with the Active Experimentation stage in Kolb's Experiential Learning Theory (ELT) (Devi & Thendral, 2023; Kolb & Kolb, 2022). Furthermore, in the theoretical profile, students are encouraged to analyze scientific articles, write analytical essays, and create concept maps. All of these activities involve high-level thinking skills according to the latest model of Bloom's Taxonomy (Vrbanović Lisac et al., 2025; Wilson, 2012).

The recommendation for students with moderate performance in the numeric profile is to learn using a corrective instruction approach based on Mastery Learning, by conducting gradual practice questions and identifying and correcting weaknesses by making log errors (Betts et al., 2021). In the practice profile, the recommendations are based on the theory of reflective observation which is in line with the concept of deliberate practice, which includes the repetition of the practice and the writing of debugging journals as a form of focused practice (Devi & Thendral, 2023; Kolb & Kolb, 2022). Meanwhile, in lectures that emphasize theory, students can apply S3QR reading techniques, utilize digital flashcards, and participate in small group discussions. This recommendation is considered to increase the effectiveness of learning for students in the transition phase to full mastery because it combines active cognitive strategies and learning technologies (Meng et al., 2025).

The most intensive recommendation is given to Low-Performance cluster students because it is considered that they still need gradual adaptive scaffolding as the main foundation (Allagui, 2024; Chen et al., 2024). In the numeric profile, students can be accompanied by peer tutors and learn basic mathematical concepts through animated videos. This strategy can be combined with a self-regulation strategy, namely a daily progress tracker to increase student confidence (Schunk & Zimmerman, 2022). For the practice profile, students are advised to do a worked example, for example by imitating and understanding the existing program, and then making modifications. This strategy can also be combined with a buddy system to develop student's independence gradually (Ouwehand et al., 2025). The theoretical profile will be greatly helped when the learning materials are available in various formats adapting the Universal Design Learning (UDL) framework. This strategy has been proven to increase student learning outcomes, engagement, and a sense of belonging in an inclusive learning environment (Almeqdad et al., 2023; Moriña et al., 2025).

Discussion

The results of this study show that the K-Means method is superior to Agglomerative Clustering and Gaussian Mixture Model (GMM) in grouping students' academic achievements. Based on the Silhouette Score, the Davies-Bouldin Index, and the Calinski-Harabasz Index, K-Means produces groups with better cohesion and separation. Although the highest Silhouette Score was obtained when the number of clusters was set to two, three clusters were chosen because they provided a more meaningful interpretation of education by representing high, medium, and low-achieving students.

The results of the grouping were then combined with learning profiles based on Numeric, Practical, and Theoretical course categories. This profiling approach allows the system to identify not only students' overall academic performance, but also their dominant learning characteristics. As a result, students with similar performance levels can receive different recommendations according to their learning profiles. These findings support previous studies that highlighted the potential of clustering techniques for personalized education and learning analysis.

In contrast to previous research that primarily focused on student grouping, this study integrates clustering and learning recommendations into a single framework. The proposed system generates personalized recommendations by combining academic achievement levels and learning profiles, resulting in nine recommendation scenarios. These recommendations can assist educators in providing more targeted academic interventions and support personalized learning strategies. However, this study should be interpreted as an exploratory study due to the relatively small dataset size and the use of academic grades as the main data source. Therefore, the findings are intended to demonstrate the initial feasibility of the proposed framework rather than to provide broad generalization across different student populations or educational contexts. Further research should involve larger and more diverse datasets and additional learning behavior indicators to improve the robustness and applicability of the proposed framework.

The validity of the proposed recommendations is supported by their alignment with established educational theories and by expert assessment of their relevance, feasibility, and usefulness. The recommendation framework was designed to translate the identified academic performance levels and learning profiles into practical learning activities, such as peer mentoring, structured practice, reflective coding, and guided remediation. However, expert validation primarily establishes the conceptual and practical appropriateness of the recommendations rather than their empirical effectiveness in improving student learning outcomes. Therefore, implementation in real learning contexts is necessary to evaluate how students respond to the recommended strategies and whether the recommendations lead to measurable improvements in academic performance and learning engagement. Future implementation could involve integrating the recommendation framework into academic advising or Learning Management System (LMS) environments, followed by longitudinal evaluation of student progress and learning outcomes.

This study proposes a clustering system to analyze students' academic achievement and produce learning recommendations that are in accordance with student profiles. This system integrates data pre-processing, clustering analysis, and recommendation generation into one unified system. Academic achievement data is categorized into numerical, practical, and theoretical learning profiles to provide a more comprehensive picture of student learning characteristics. The results showed that K-Means outperformed the Agglomerative Clustering and Gaussian Mixture Model based on the Silhouette Score, Davies-Bouldin Index, and Calinski-Harabasz Index. The clustering model chosen succeeded in identifying three groups of student performance, namely students with High performance, Moderate performance, and Low performance. These groups were then leveraged to develop a recommendation system that combined academic performance with learning profiles, resulting in nine personalized learning recommendation scenarios.

Based on the results of this study, it can be concluded that clustering can not only serve as an analytical tool for student segmentation, but also as a foundation for producing actionable learning recommendations. By integrating academic achievement analysis with educational theories such as Self-Regulated Learning, Mastery Learning, and Experiential Learning, the proposed system provides practical support for personalized learning and early intervention in the academic field. However, as an exploratory study, this research is limited by its relatively small sample size and its focus on a single group of first-semester students. Therefore, the findings should be interpreted as preliminary evidence of the feasibility of the proposed framework. Further research should use a larger and more diverse sample, include additional learning behavior indicators from the Learning Management System (LMS), and evaluate the effectiveness of the proposed recommendations in real educational settings.

REFERENCES

Allagui, B. (2024). A scaffolding intervention to improve self-efficacy in source-based argumentative writing. *Frontiers in Psychology*, *15*. https://doi.org/10.3389/fpsyg.2024.1454104

Almeqdad, Q. I., Alodat, A. M., Alquraan, M. F., Mohaidat, M. A., & Al-Makhzoomy, A. K. (2023). The effectiveness of universal design for learning: A systematic review of the literature and meta-analysis. *Cogent Education*, *10*(1). https://doi.org/10.1080/2331186X.2023.2218191

Azzahra, C., & Sriani. (2025). Clustering of High School Students Academic Scores Using the K-Means Algorithm. *Journal of Information Systems and Informatics*, *7*(1), 572–586. https://doi.org/10.51519/journalisi.v7i1.1029

Betts, A., Thai, K.-P., & Gunderia, S. (2021). *Personalized Mastery Learning Ecosystems: Using Bloom's Four Objects of Change to Drive Learning in Adaptive Instructional Systems* (pp. 29–52). https://doi.org/10.1007/978-3-030-77857-6_3

Chen, T., Zhao, Y.-J., Huang, F.-Q., Liu, Q., Li, Y., Alolga, R. N., Zhang, L., & Ma, G. (2024). The effect of problem-based learning on improving problem-solving, self-directed learning, and critical thinking ability for the pharmacy students: A randomized controlled trial and meta-analysis. *PLOS ONE*, *19*(12), e0314017. https://doi.org/10.1371/journal.pone.0314017

Cheng, Y., Guan, R., Li, T., Raković, M., Li, X., Fan, Y., Jin, F., Tsai, Y.-S., Gašević, D., & Swiecki, Z. (2025). Self-regulated Learning Processes in Secondary Education: A Network Analysis of Trace-based Measures. *Proceedings of the 15th International Learning Analytics and Knowledge Conference*, 260–271. https://doi.org/10.1145/3706468.3706502

Danu, R., Putra, K., Palupi, K. S., Wakhidah, N., Informasi, J. S., & Akademik, N. (2025). Pengelompokkan Data Nilai Mahasiswa Menggunakan Metode K-Means. *Jurnal Algoritme*, *6*(1), 88–99. https://doi.org/10.35957/algoritme.v6i1.11313

Devi, M. K. K., & Thendral, M. S. (2023). Using Kolb's Experiential Learning Theory to Improve Student Learning in Theory Course. *Journal of Engineering Education Transformations*, *37*(1), 70–81. https://doi.org/10.16920/jeet/2023/v37i1/23133

Dewi, F. P., Aryni, P. S., & Umaidah, Y. (2022). Implementasi Algoritma K-Means Clustering Seleksi Siswa Berprestasi Berdasarkan Keaktifan dalam Proses Pembelajaran. *JISKA (Jurnal Informatika Sunan Kalijaga)*, *7*(2), 111–121.

Dinata, R. K., Novriando, H., Hasdyna, N., & Retno, S. (2020). Reduksi Atribut Menggunakan Information Gain untuk Optimasi Cluster Algoritma K-Means. *JEPIN ((Jurnal Edukasi Dan Penelitian Informatika))*, *6*(1), 48–53.

Hakim, H. T., Usino, W., Informasi, S., Informasi, F. T., Luhur, U. B., Selatan, J., & Clustering, A. (2024). Penerapan Algoritma Agglomerative Clustering Untuk Mengelompokkan Provinsi Di Indonesia

Berdasarkan Indikator Pendidikan Application Of Agglomerative Clustering Algorithm To Group Provinces In Indonesia Based On. *Bit (Fakultas Teknologi Informasi Universitas Budi Luhur)*, *21*(2), 153–160.

Kinjal, Parida, S. M., Suthar, J., & Pande, S. D. (2024). Predicting Academic Success : A Comparative Study of Machine Learning and Clustering-Based Subject Recommendation Models. *EAI Endorsed Transactions on Internet on Things Research*, *10*, 1–7. https://doi.org/10.4108/eetiot.5378

Kolb, A. Y., & Kolb, D. A. (2022). Experiential Learning Theory as a Guide for Experiential Educators in Higher Education. *Experiential Learning and Teaching in Higher Education*, *1*(1), 38. https://doi.org/10.46787/elthe.v1i1.3362

Latief, M. A., Nabila, L. R., Miftakhurrahman, W., Ma'rufatullah, S., & Tantyoko, H. (2024). Handling Imbalance Data using Hybrid Sampling SMOTE-ENN in Lung Cancer Classification. *International Journal of Engineering and Computer Science Applications (IJECSA)*, *3*(1), 11–18. https://doi.org/10.30812/ijecsa.v3i1.3758

Li, S., & Li, D. (2025). RESEARCH ON PERSONALIZED LEARNING RECOMMENDATION SYSTEM BASED ON MACHINE LEARNING ALGORITHM. *Scalable Computing: Practice and Experience*, *26*, 432–440. https://doi.org/10.12694/scpe.v26i1.3844

Maguate, G. S., Odango, J. S., Cruz, J. N. D. Dela, Cornel, M. A., M., A. A., & Uy, F. T. (2024). Analyzing Student Performance : A Clustering Approach for Academic Intervention. *INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY, AND EXCELLENCE (IMJRISE)*, (2021), 941–946.

Makawana, M., & Mehta, R. G. (2023). Discovering Search Space Using M-distance Clustering of Semantic Relatedness Based Weighted Network for the Content-based Recommender System. *Journal of Scientometric Research*, *12*(2), 243–253.

Meng, N., Mat Deli, M., & Abdul Rauf, U. A. (2025). Educational Technology and AI: Bridging Cognitive Load and Learner Engagement for Effective Learning. *Sage Open*, *15*(4). https://doi.org/10.1177/21582440251395930

Moriña, A., Carballo, R., & Doménech, A. (2025). Transforming higher education: a systematic review of faculty training in UDL and its benefits. *Teaching in Higher Education*, *30*(7), 1722–1739. https://doi.org/10.1080/13562517.2025.2465994

Ouassif, K., Ziani, B., & Herrera-tapia, J. (2025). Empowering Education : Leveraging Clustering and Recommendations for Enhanced Student Insights. *Education Sciences*, 1–14.

Ouwehand, K., Lespiau, F., Tricot, A., & Paas, F. (2025). Cognitive Load Theory: Emerging Trends and Innovations. *Education Sciences*, *15*(4), 458. https://doi.org/10.3390/educsci15040458

Rizki, I. A., Sunarti, T., & Salsabila, N. F. (2024). A case study of how digital competencies affect Pre-Service Physics Teachers ' Professional growth : Realizing SDG 4. *E3S Web of Conferences*, *04007*, 1–10.

Schunk, D. H., & Zimmerman, B. J. (2022). *Self-regulation of Learning and Performance*. Routledge. https://doi.org/10.4324/9780203763353

Sharif, M. S., Tamang, M. T., Mani, D., & Elmedany, W. (2024). Application of Clustering Algorithms to enhance Personalized Learning through Recommendation Model. *International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)*. https://doi.org/10.1109/3ict64318.2024.10824599

Stanny, C. J. (2016). Reevaluating bloom's taxonomy: What measurable verbs can and cannot say about student learning. *Education Sciences*, *6*(4). https://doi.org/10.3390/educsci6040037

Suhana, R., Mahmudy, W. F., & Budi, A. S. (2022). Fish Image Classification using Transfer Learning Method withAdaptive Learning Rate Fish Image Classification using Transfer Learning Method with Adaptive Learning Rate. *Knowledge Engineering and Data Science*, *5*(1).

Syakur, M. A., Khotimah, B. K., Rochman, E. M. S., & Satoto, B. D. (2018). Integration k-means clustering method and elbow method for identification of the best customer profile cluster. *IOP Conference Series: Materials Science and Engineering*, *336*(1), 12017.

Ullah, A., Mohmand, M. I., Hussain, H., Johar, S., & Khan, I. (2023). Customer Analysis Using Machine Learning-Based. *Journal Sensors*. https://doi.org/https://doi.org/10.3390/s23063180

Utami, I. Q., Hwang, W., & Ningrum, R. A. (2024). Student ' s Behavior Clustering based on Ubiquitous Learning Log Data using Unsupervised Machine Learning. *Journal of Advanced Technology and Multidiscipline (JATM)*, *03*(01), 13–20.

Villegas-ch, W., García-ortiz, J., & Sánchez-viteri, S. (2024). *Personalization of Learning : Machine Learning Models for Adapting Educational Content to Individual Learning Styles*. *12*(August). https://doi.org/10.1109/ACCESS.2024.3452592

Vrbanović Lisac, D., Radanović, I., & Šimić Šašić, S. (2025). Enhancing higher order thinking skills to improve understanding of biological content: A quasi-experimental study. *Natural Sciences Education*, *54*(2). https://doi.org/10.1002/nse2.70023

Wilson, L. O. (2012). *Anderson and Krathwohl Bloom's Taxonomy Revised Understanding the New Version of Bloom's Taxonomy*.

Zamakhsyari, F., & Fatwanto, A. (2025). Language model optimization for mental health question answering application. *International Journal of Electrical and Computer Engineering (IJECE)*, *15*(5), 4829–4836. https://doi.org/10.11591/ijece.v15i5.pp4829-4836