Journal of Information Technology and Education
Volume 18, No. 2, September 2025
https://doi.org/10.24036/jtip.v18i2.967

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

Grouping of Student Achievement Based on Student Names in Class VII of SMPN 28 Sarolangun Using the K-Means Clustering Method

Nur Chintia Ningsih[1*✉], Dony Novaliendry[1], Yeka Hendriyani[1], Delvi Asmara[1]
[1] Department of Electronics Engineering, Universitas Negeri Padang, Padang, Indonesia
**Corresponding Author: nurchintia@student.unp.ac.id*

Article Information

Article history:

No. 967
Rec. May 28, 2025
Rev. July 10, 2025
Acc. July 15, 2025
Pub. July 18, 2025
Page. 864 – 882

Keywords:

- *K-Means*
- *Elbow Method*
- *Clustering*
- *Student Achievements*

ABSTRACT

Although the selection of outstanding students is important to provide awards and recognition for student achievement, the methods currently used by schools are not optimal. The process often takes a long time and requires a lot of manpower to collect and process student data, which can ultimately disrupt daily school operations. This study aims to identify outstanding students in class VII at SMP 28 Sarolangun a junior high school located under the supervision of the Ministry of Education and Culture. The school is situated on Lintas Sumatera Street, Simpang Nibung, in Simpang Nibung Village, Singkut District, Sarolangun Regency, Jambi Province. using the clustering method with the K-Means algorithm. This type of research is quantitative research. The method used in this study is K-Means Clustering , with the determination of the optimal number of clusters using the Elbow Method. The results of the study obtained a grouping of students into four clusters, including Cluster 1 with 10 students (15.2%), Cluster 2 with 16 students (24.2%), Cluster 3 with 25 students (37.9%) and Cluster 4 with 15 students (22.7%). From the resulting Elbow graph, the elbow point is seen at the value of K = 4, which indicates that four clusters are the most effective and efficient number to separate student data into meaningful groups.

How to Cite:

Ningsih, N. C., & et al. (2025). Grouping of Student Achievement Based on Student Names in Class VII of SMPN 28 Sarolangun Using the K-Means Clustering Method. Jurnal Teknologi Informasi Dan Pendidikan, 18(2), 864-882. https://doi.org/10.24036/jtip.v18i2.967

1. INTRODUCTION

Improving the quality of education is one of the main priorities in development policies in Indonesia that must continue to be prioritized. One important factor that influences this achievement is the role of teachers in providing effective and quality teaching. The progress of quality and student activity in learning are basic elements of the success of the education process, although not all students have the same speed in understanding the material taught [1]. This is certainly a special concern for schools in efforts to achieve educational success. Every student has the potential to develop hard skills and soft skills that can support them in achieving success in the future. Hard skills are generally related to technical expertise or skills that can be learned through formal education and training [2]. On the other hand, soft skills include social skills, personality, and interpersonal skills, such as the ability to communicate, cooperate, lead, and adapt [3]. In the era of big data, clustering algorithms like K-Means have become increasingly important for analyzing educational data, as they can help uncover patterns and insights that support more effective decision-making in teaching and learning processes [4].

Currently, the process of selecting outstanding students in schools still has several weaknesses. One of the main problems is that data processing takes a long time and is prone to human error in decision making [5]. Although selecting outstanding students is important to provide awards and recognition for student achievement, the methods currently used by schools are not optimal. The process often takes a long time and requires a lot of manpower to collect and process student data, which can ultimately disrupt daily school operations. One solution that can be applied is the use of technology to automate and improve the accuracy of data processing, so that the decision-making process becomes faster and more efficient [6]. In this context, data clustering methods such as K-Means have been increasingly adopted for student data analysis because of their ability to process large-scale data and identify meaningful patterns, supporting more effective and efficient decision-making [7].

Based on these problems, researchers conducted research at SMP 28 Sarolangun with the aim of identifying outstanding students in class VII using the clustering method with the K-Means algorithm. Clustering is the process of grouping data into several groups that have maximum similarity [8]. K-Means is one of the popular algorithms used for this purpose because it is able to group large numbers of objects based on similarity through an iterative partition procedure. This process is often used in data mining to identify natural groups of similar cases, which can then be analyzed further [9]. In addition, the use of K-Means for educational data segmentation has been shown to effectively support targeted decision-making and optimize educational services. Furthermore, the effectiveness of K-Means has been demonstrated in other fields, such as information management systems, where the algorithm can improve data organization and enhance data retrieval performance through cluster-based classification [10];[11];[12].

SMP Negeri 28 Sarolangun is one of the junior high schools located in Sarolangun Regency, with 66 students in grade VII, consisting of 22 students in grade 7A, 22 students in grade 7B, and 22 students in grade 7C. This school does not yet have a superior class program aimed at outstanding students as a form of appreciation and motivation to continue to improve their academic and non-academic abilities. This superior class is intended for grade VII students who show the best achievements during one school year and will be promoted to the superior class when they move up to grade VIII. However, so far the selection process for superior classes has been carried out manually by homeroom teachers and the school, which tends to take a long time and has the potential to contain elements of subjectivity. Thus, a data-based system is needed that is able to group students objectively and measurably.

To overcome these problems, the K-Means Clustering method was chosen as a technique for grouping data for class VII students at SMPN 28 Sarolangun. This method is used because it is able to group students into several clusters based on similarities in the characteristics of the data they have. In the context of this study, students will be grouped into three clusters, namely cluster 1 (C1) for students with high achievement, cluster 2 (C2) for students with moderate achievement, and cluster 3 (C3) for students with low achievement [13];[14]. The use of this method is relevant because K-Means is able to handle quantitative data such as academic grades, as well as qualitative data that has been quantified such as extracurricular activity and students' social attitudes. The attributes used in this clustering process include mid-term exam scores, final exam scores, homeroom teacher assessments regarding students' social attitudes, active participation in extracurricular activities, and attendance percentage. The selection of these attributes is supported by several previous studies, such as research conducted by Kurniawan & Ferdiansyah, who applied K-Means to group high-achieving students based on academic and non-academic data [15]. Another study conducted by Saputra and Nataliani, which emphasized the importance of multi-aspect based grouping in the selection of superior classes [16]. By considering these various aspects, the clustering results will provide a comprehensive picture of student achievement, both in terms of academics and character development and talent interests.

The selection process for the superior class itself is carried out at the end of each school year, specifically for grade VII students who will move up to grade VIII. With the results of the K-Means method grouping, schools will get recommendations for the best students from each class, who will then be considered for inclusion in the class VIII superior class. With this clustering-based system, the selection process can be carried out more transparently and measurably, and reduce the potential for subjectivity that often occurs if it only relies on manual assessments from homeroom teachers. Based on the solutions that have been presented, it is hoped that the development of this system will be able to produce groups (clusters) of students that are in accordance with their abilities and assessments.

2. RESEARCH METHODS

The type of research is quantitative research because it analyzes numerical and statistical data to identify student achievement patterns. The method used in this study is

K-Means Clustering, with the determination of the optimal number of clusters using the Elbow Method. This approach allows grouping students based on academic and non-academic criteria in depth and systematically, so that the results can be used to optimize coaching in superior classes [17][18]. The Elbow Method is a technique applied to K-Means Clustering to determine the optimal number of clusters to be formed. The purpose of the Elbow Method is to choose a small k value but still has a low within-cluster sum of squares (WSS) value, so that the number of clusters selected can minimize variation within clusters and maximize variation between clusters [19]. Similar methods have been successfully implemented in other fields, such as optimizing the grouping of patient medical records in hospitals, which also utilizes data mining and K-Means Clustering algorithm to improve the accuracy and efficiency of data grouping [20].

The study was conducted in class VII of SMP 28 Sarolangun. The location of the school is on Jln. Lintas Sumatera Simpang Nibung, Simpang Nibung Village, Singkut District, Sarolangun Regency, Jambi Province. The population in this study includes all academic and non-academic data of students in class VII of SMP 28 Sarolangun. A total of 66 student records were used, consisting of 22 students each from class 7A, 7B, and 7C. The data used are academic and non-academic scores from the 2023/2024 academic year. The data source used is secondary data. The secondary data in question is data collected from previously existing data. This data is available in the form of documents, reports, or databases that are relevant to the research needs. In this study, secondary data includes student assessment documents, attendance recapitulations, and extracurricular activity records.

The research procedure was carried out through the following steps:

- Step 1 – Data Collection: Data were collected from class VII students at SMPN 28 Sarolangun, including academic scores (UTS and UAS), class teacher's assessments of behavior, extracurricular participation (Pramuka and PMR), and attendance records. All data were manually entered into a table based on a predefined template.
- Step 2 – Data Input and System-Based Conversion: The prepared data were input into the system interface. Academic scores were already in numeric form, while behavior and extracurricular participation were in letter format (e.g., A, B, C). These were automatically converted to numeric values by the system before clustering.
- Step 3 – Determination of the Optimal Cluster (Elbow Method): The Elbow Method was used to determine the optimal value of K by analyzing the WCSS graph. The elbow point was found at K = 4.
- Step 4 – K-Means Clustering Process: The system ran the K-Means Clustering algorithm to group students based on similarities in their academic and non-academic attributes.
- Step 5 – Evaluation and Interpretation of Results : The clustering results were analyzed and interpreted based on patterns in academic performance, attendance, behavior, and extracurricular activity.

The research stages are visualized through a series of flowcharts that highlight the core processes relevant to the clustering method, including dataset import, cluster determination using the Elbow Method, and execution of the K-Means algorithm. Supporting system flows such as login or user authentication are not discussed, as they are not directly related to the research objectives.:

2.1. Flowchart dataset

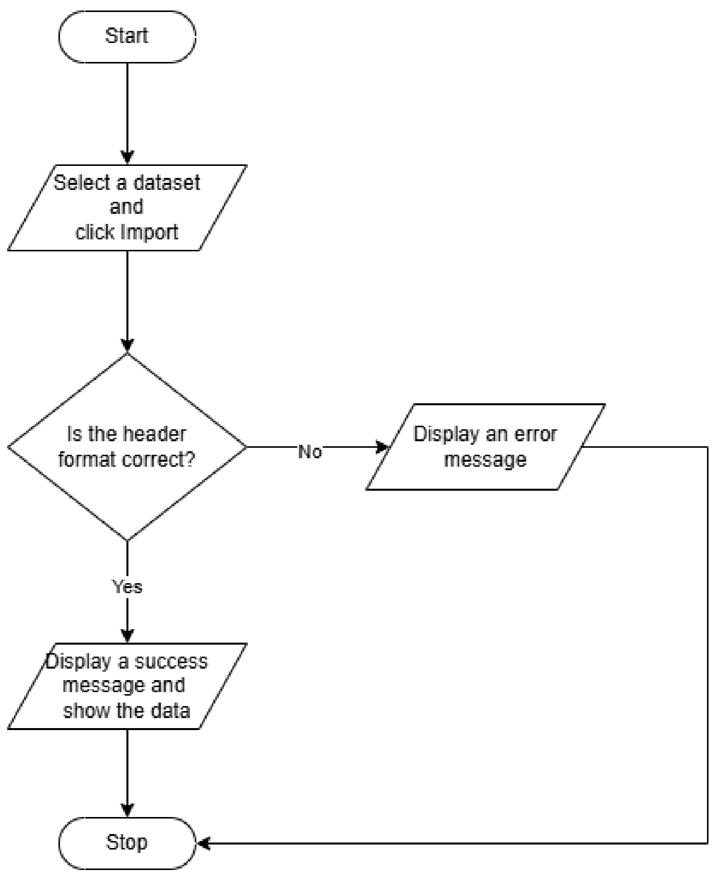


Figure 1. Flowchart dataset

Figure 1 illustrates the process flow of importing a dataset into the system. 1) Start: The process begins. 2) Select dataset and click import: The user selects the dataset file to be imported and then presses the import button. 3) Is the header format correct?: The system checks whether the header format in the dataset is as expected. 4) If No: The system displays an error message, then the process returns to the dataset selection stage. 5) If Yes: The system displays a message that the import was successful and displays the data. 6) Stop: The process is complete

2.2. Flowchart elbow

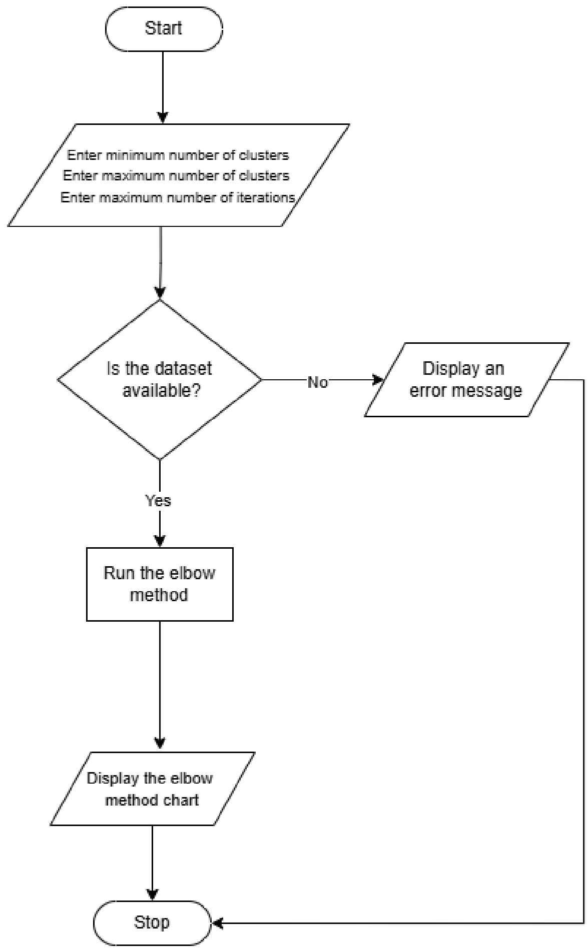


Figure 2. Elbow flowchart

Figure 2 explains the stages in the process of determining the optimal number of clusters using the elbow method. 1) Start: The process begins. 2) Input minimum cluster, maximum cluster, maximum iteration: The user enters the minimum and maximum number of clusters and the maximum number of iterations. 3) Does the dataset exist?: The system checks whether the dataset is available. 4) If No: The system displays an error message, then the process returns to the parameter input stage. 5) If Yes: The system continues to the elbow process. 6) Elbow method process: The system performs the elbow method calculation to determine the optimal number of clusters. 7) Display elbow method graph: The calculation results are visualized in the form of an elbow graph. 8) Stop: The process is complete.

Journal of Information Technology and Education
Volume 18, No. 2, September 2025
https://doi.org/10.24036/jtip.v18i2.967

2.3. Clustering setup

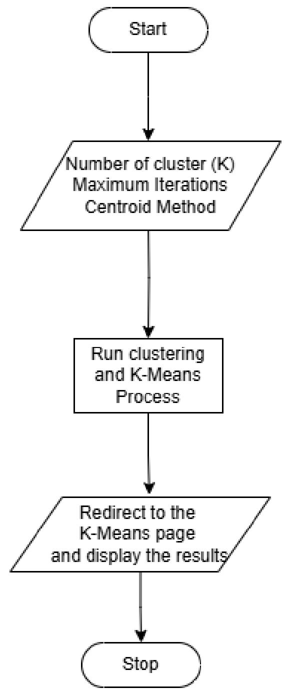


Figure 3. Clustering setup flowchart

Figure 3 explains the process of clustering setup and execution using the k-means method. 1) Start: The process begins. 2) Input the number of clusters (K), maximum iterations, centroid method: The user enters the parameters for the number of clusters, maximum iterations, and centroid method used. 3) Clustering and k-means process: The system runs the clustering process using the k-means algorithm based on the parameters entered. 4) Switch to the k-means page and display results: After the process is complete, the system displays the clustering results on a special k-means page. 5) Stop: The process is complete

Independent Variable: K-Means Clustering Method, this technique is used to group data based on attributes such as academic grades, attendance, social attitudes, and extracurricular activities, which are used to analyze the similarities in student characteristics in class VII of SMP N 28 Sarolangun. Dependent Variable: Patterns and clusters of student grouping results in class VII of SMPN 28 Sarolangun, which reflect the category of high-achieving students based on high grades, medium grades and low grades analyzed from the specified attributes.

The research instruments used in the creation of this system include various software, frameworks, programming languages, and design and mock-up tools that support efficient and structured development. The back-end framework used is Laravel, which is

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

known to be lightweight and easy to learn, and provides features such as routing, database management, and data validation that facilitate the development of structured web applications. For the development of the user interface (front-end), a combination of HTML, CSS, and JavaScript is used, with the support of the Bootstrap framework to create a responsive and user-friendly design. The database used is MySQL, which provides efficient and reliable data management capabilities.

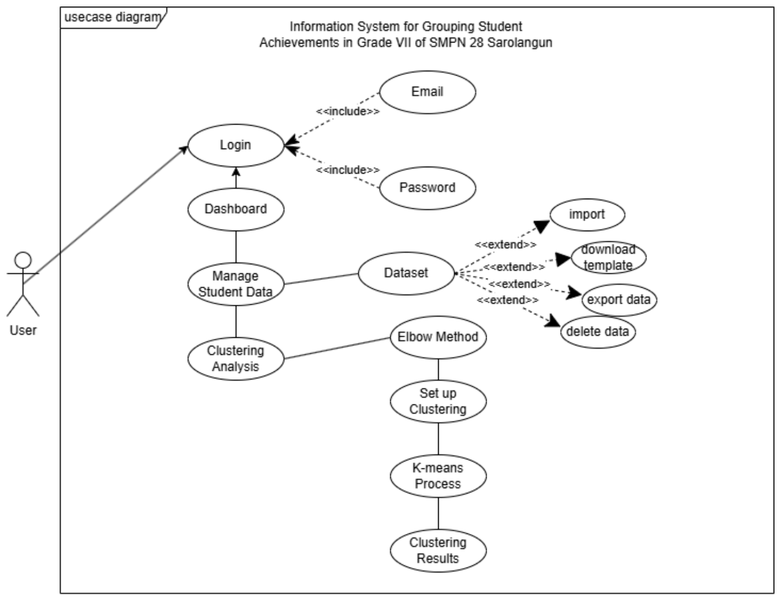


Figure 4. Use Case diagram of the information system for grouping the achievements of class VII students at SMPN 28 Sarolangun.

Based on figure 4, the Use Case Diagram illustrates the interaction between a user and a clustering system based on the K-Means method. In the diagram, there is one actor, namely the user who can perform several activities. The user starts by logging in using an email and password, then the account is verified. After logging in, the user can manage student data, such as importing data, deleting all data, exporting data, or downloading data templates. After the data is ready, the user can determine the optimal number of clusters using Elbow Method. After the number of clusters is determined, the user can proceed to the cluster setup stage, at this stage the system runs the k-means process to group student data. The final result of this process is the clustering result that displays groups of students based on their achievements.

Journal of Information Technology and Education
Volume 18, No. 2, September 2025
https://doi.org/10.24036/jtip.v18i2.967

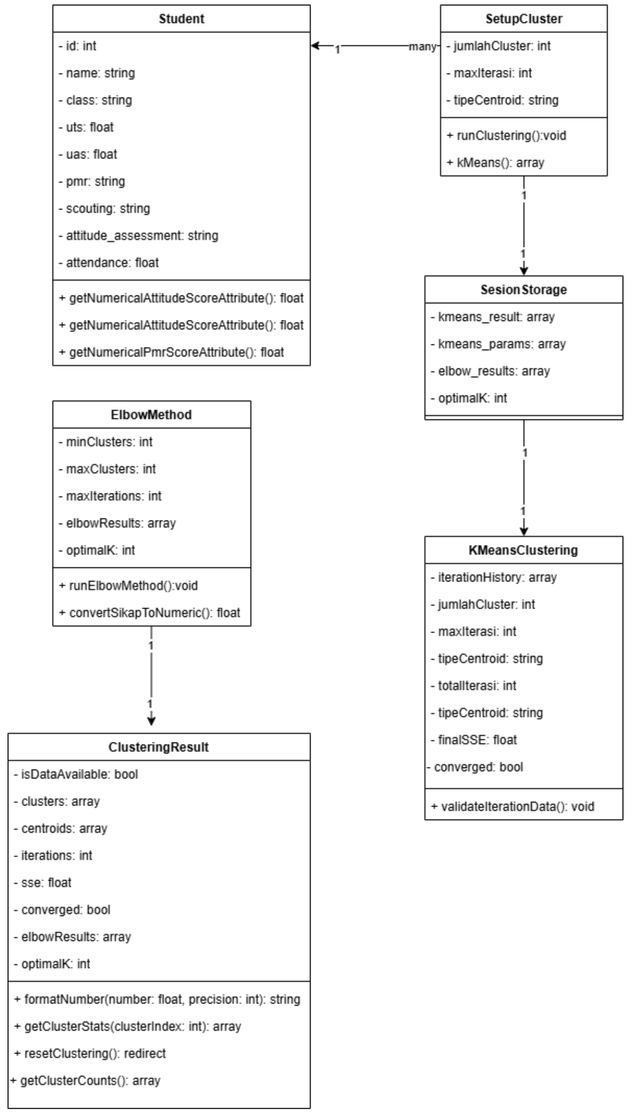


Figure 5. Class diagram of the information system for grouping the achievements of class VII students at SMPN 28 Sarolangun.

In figure 5, the Class Diagram illustrates the data structure used in the system, including the relationships between classes. This process starts from the student class which acts as a provider of raw data in the form of student information. These data are the basis for the clustering process that will be carried out. The results of this analysis will be a reference for the next clustering process. This process is carried out based on the parameters that have been set by the Cluster Setup and the number of clusters obtained from the Elbow Method. K-Means Clustering will group student data into several clusters through a series of iterations until stable results are achieved.

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

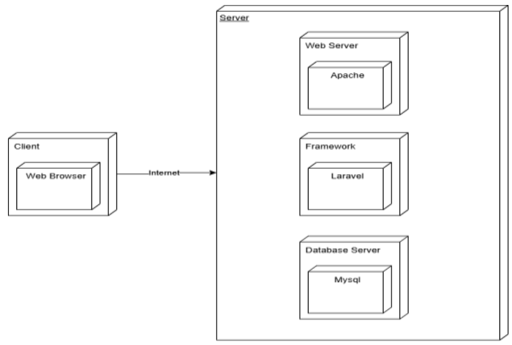


Figure 6. Deployment diagram of the information system for grouping the achievements of class VII students at SMPN 28 Sarolangun.

In the diagram of figure 6, this communication flow illustrates how web applications run in real life, starting from user requests in the browser, processing on the server, to data management in the database. Each software component is placed on the appropriate node, so that the system can run effectively and integrated.

3. RESULT AND DISCUSSION

3.1. Data collection

The data obtained comes from various sources that are grouped into separate files, namely the report card value dataset which includes Attitude values, UTS and UAS values. Extracurricular dataset which includes student activity assessments in PMR and Scout activities. Student attendance dataset which includes the percentage of student attendance during odd and even semesters.

3.2. Data Preprocessing

In the data preprocessing stage , the previously obtained datasets are merged and cleaned. This process produces cleaner data that avoids missing value data . The process of converting letter values to numbers is carried out automatically by the system during the analysis process.

Table 1. Dataset resulting from data reprocessing

Name	Class	Uts	Uas	Attitude_Assessment	Scout	Pmr	Perncetage%
Aditya	7A	73	79	B	C	**A**	93
Arum Sugita	7A	90	89	SB	C	A	90
Auliya Ramadani	7A	40	70	B	B	B	82
Bagas Pratama Putra	7A	48	75	C	B	B	95
Cika Wulandarai	7A	50	70	C	B	A	88

Based on table 1, there are a total of 5 student data that will be used in the analysis process. All of these data have gone through the preprocessing stage and the merging of various related documents, such as academic grades, extracurricular activities, attitude assessments, and student attendance. This dataset will then be processed using the K-Means Clustering algorithm.

3.3. Implementation of K-Means

3.3.1 Create an Account

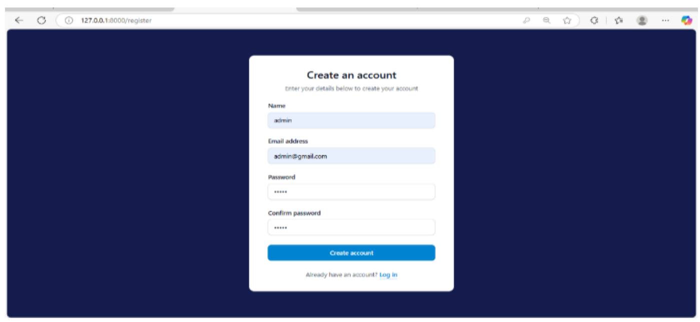


Figure 7. Create an account menu display

Based on figure 7, the first step that must be taken by the user is to create an account on the system. The creation of this account serves as initial authorization so that only registered users can access important features in the system.

3.3.2 Login to the system

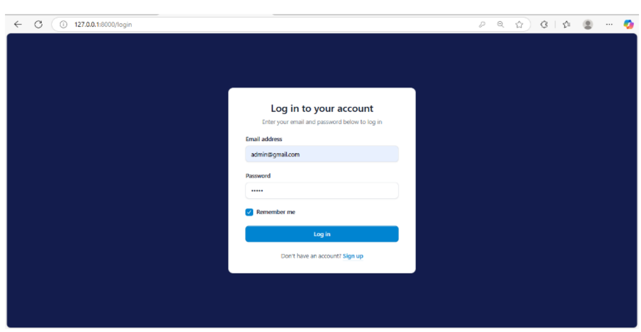


Figure 8. Login to your account menu display

From figure 8, once the account is successfully created, the user can log in to the system. The login feature provides security to the system so that data can only be accessed by users who have valid credentials.

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

3.3.3. Import Dataset to System

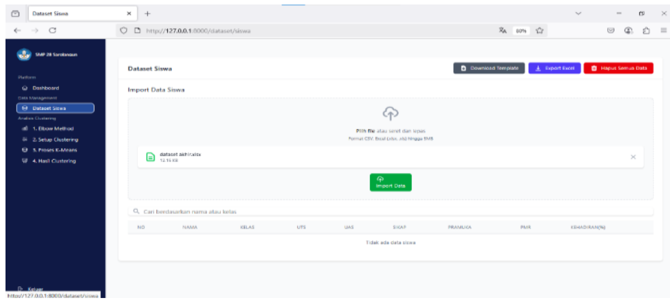


Figure 9. Student dataset import view

The previously obtained dataset is then uploaded in file form and will automatically be displayed in the system based on figure 9.

3.3.4. Determining the Optimal Number of Clusters Using the Elbow Method

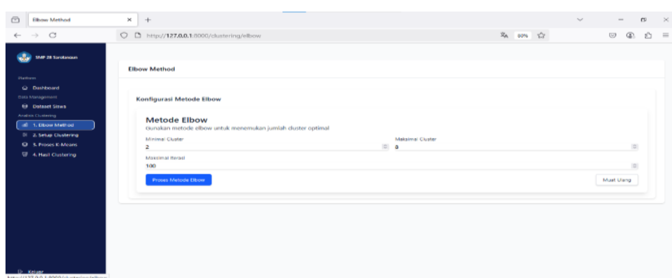


Figure 10. Elbow method settings display

After the dataset is successfully uploaded, the next step is to determine the optimal number of clusters based on figure 10. This method calculates the Within-Cluster Sum of Squares (WCSS) value for various values of K (number of clusters), then the results are plotted in a graph. In the graph, the elbow point indicates the ideal number of clusters.

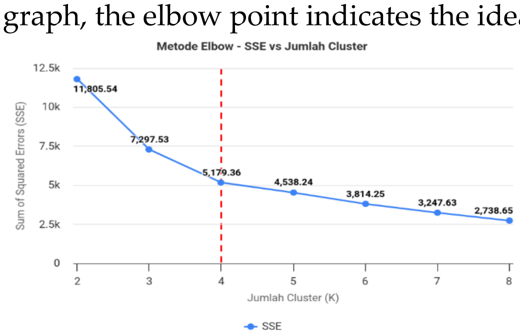


Figure 11. Graph of elbow method results

Journal of Information Technology and Education
Volume 18, No. 2, September 2025
https://doi.org/10.24036/jtip.v18i2.967

Based on the results of the Elbow graph on the system in figure 11, it was found that the elbow point occurred at K = 4, so the optimal number of clusters for this data is 4 groups.

3.3.5. K-Means Setup and Iteration Process

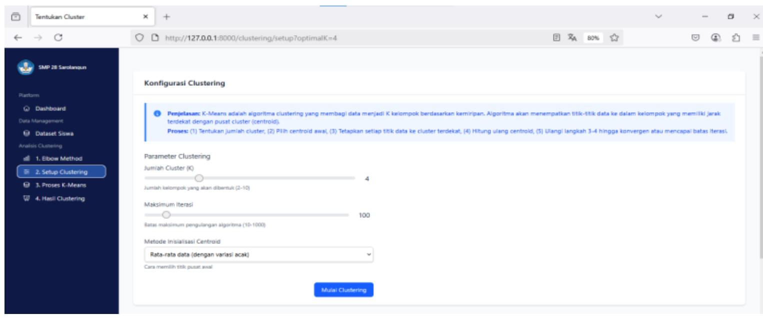


Figure 12. Clustering setup view

Following the determination of the optimal number of clusters as shown in Figure 12, the system advances to the implementation phase of the K-Means Clustering algorithm, starting with the random initialization of centroids for the four clusters.

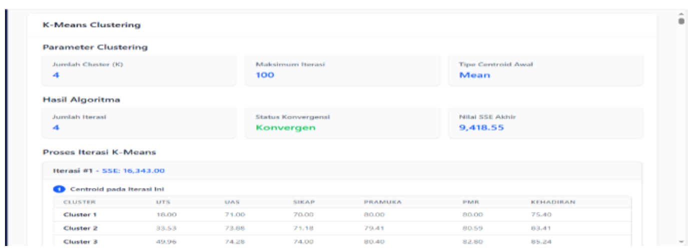
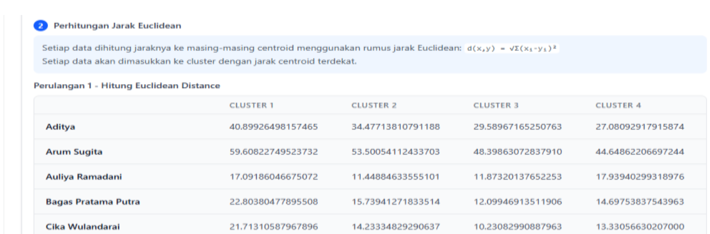


Figure 13. K-means iteration process

From figure 13, iteration is carried out 4 times until the grouping results are stable (convergent).

3.3.6. Euclidean Distance Calculation



Perhitungan Jarak Euclidean

Setiap data dihitung jaraknya ke masing-masing centroid menggunakan rumus jarak Euclidean: $d(x,y) = \sqrt{\Sigma(x_i - y_i)^2}$
Setiap data akan dimasukkan ke cluster dengan jarak centroid terdekat.

Perulangan 1 - Hitung Euclidean Distance

	CLUSTER 1	CLUSTER 2	CLUSTER 3	CLUSTER 4
Aditya	40.89926498157465	34.47713810791188	29.58967165250763	27.08092917915874
Arum Sugita	59.60822749523732	53.50054112433703	48.39863072837910	44.64862206697244
Auliya Ramadani	17.09186046675072	11.44884633555101	11.87320137652253	17.93940299318976
Bagas Pratama Putra	22.80380477895508	15.73941271833514	12.09946913511906	14.69753837543963
Cika Wulandarai	21.71310587967896	14.23334829290637	10.23082990887963	13.33056630207000

Figure 14. Euclidean distance calculation display

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

From figure 14, the distance between each student data and each centroid is calculated using the Euclidean Distance formula. The purpose of this calculation is to find out how close the student data is to each centroid.

3.3.7. Temporary Grouping

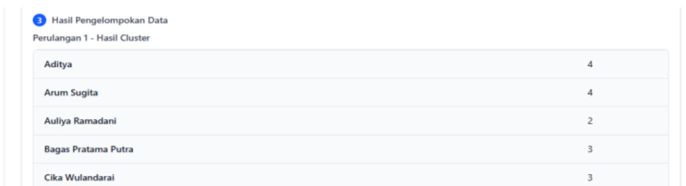


Figure 15. Display of data clustering results (loop 1-cluster results)

Based from figure 15, after the distance is calculated, each student data will be placed into the cluster that has the closest distance to the centroid.

3.3.8. Centroid Update

The average of the data included in each cluster is recalculated to form a new centroid position. This process is repeated continuously. In the 4th iteration, the system detects that there are no more significant changes in the data grouping. This indicates that the process has reached convergence, and the clustering results are considered final.

3.3.9. Final Results and Visualization of K-Means

The final results of the K-Means Clustering process show that students were successfully grouped into four main clusters. The number of students in each cluster is as follows

3.3.9.1. K-Means Clustering Information Results

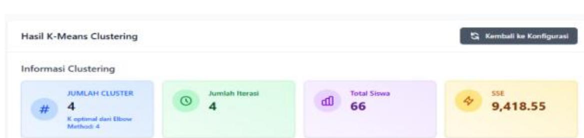


Figure 16. Cluster information display

From figure 16, the data of class VII students of SMP 28 Sarolangun are grouped using the K-Means Clustering method. The data used includes five attributes, namely mid-term exam scores, final exam scores, attendance, homeroom teacher assessments regarding social

Journal of Information Technology and Education
Volume 18, No. 2, September 2025
https://doi.org/10.24036/jtip.v18i2.967

attitudes, and activity in extracurricular activities. The number of clusters used in grouping is K = 4 based on the optimal value of the Elbow method previously conducted.

3.3.9.2. K-Means Clustering Visualization Results

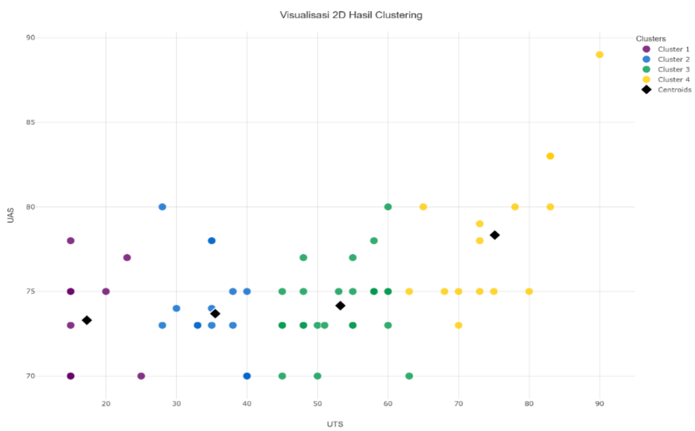


Figure 17. 2D K-Means clustering visualization view

In the graph of figure 17, there are student data divided into four different clusters, marked with purple, blue, green, and yellow. Each point represents a student, while the black rhombus-shaped point shows the center point (centroid) of each cluster. The analysis of each cluster is as follows.

Cluster 1 (purple) is located on the bottom left which shows students with very low mid-term exam scores around 15-30 and moderate final exam scores around 70-76. Students in this cluster represent students with poor academic performance, both at the beginning and end of the semester, so they need special attention, for example, they can do additional tutoring or a more personal learning approach.

Cluster 2 (blue) is located in the middle left of the graph indicating students with low mid-term exam scores between 25-40 and medium final exam scores between 72-80, with a tendency for lower mid-term exam scores. Students in this cluster have low mid-term academic ability but better performance during final exams, they need more attention, such as additional tutoring or creating appropriate learning strategies.

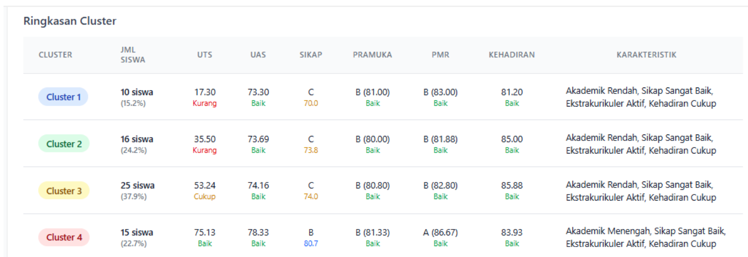
Cluster 3 (green) is located in the middle of the graph which reflects students with moderate academic performance in mid-term exam scores ranging from 45-60 and final exam scores ranging from 70-80. Students in this cluster group show fairly good academic performance stability, but still need to have room for improvement so that groups in this cluster have the potential to rise to a higher level of performance by providing the right motivation and learning strategies.

Cluster 4 (yellow) is in the upper right area of the graph, with a high range of mid-term and final exam scores of around 65 to 90. Students in this cluster group describe a

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

group of students with excellent and consistent academic performance throughout the semester. Students in this cluster can be role models or can also be a mentor for their classmates, and have the potential to continue to be developed through achievement coaching programs.

3.3.9.3 Cluster Summary Results


Ringkasan Cluster

CLUSTER	JML SISWA	UTS	UAS	SIKAP	PRAMUKA	PMR	KEHADIRAN	KARAKTERISTIK
Cluster 1	10 siswa (15.2%)	17.30 Kurang	73.30 Baik	C 70.0	B (81.00) Baik	B (83.00) Baik	81.20 Baik	Akademik Rendah, Sikap Sangat Baik, Ekstrakurikuler Aktif, Kehadiran Cukup
Cluster 2	16 siswa (24.2%)	35.50 Kurang	73.69 Baik	C 73.8	B (80.00) Baik	B (81.88) Baik	85.00 Baik	Akademik Rendah, Sikap Sangat Baik, Ekstrakurikuler Aktif, Kehadiran Cukup
Cluster 3	25 siswa (37.9%)	53.24 Cukup	74.16 Baik	C 74.0	B (80.80) Baik	B (82.80) Baik	85.88 Baik	Akademik Rendah, Sikap Sangat Baik, Ekstrakurikuler Aktif, Kehadiran Cukup
Cluster 4	15 siswa (22.7%)	75.13 Baik	78.33 Baik	B 80.7	B (81.33) Baik	A (86.67) Baik	83.93 Baik	Akademik Menengah, Sikap Sangat Baik, Ekstrakurikuler Aktif, Kehadiran Cukup

Figure 18. Cluster summary view

Based on figure 18, you can see the results of the cluster summary display and details of the members of each cluster group.

3.3.9.4. Student Distribution Results for each Cluster

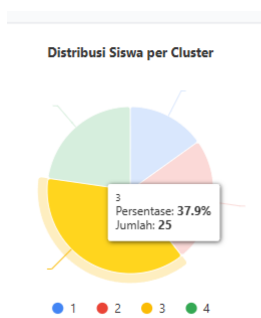


Figure 19. Display of student distribution per cluster

Figure 19 display of student distribution per cluster in the form of a pie chart, which shows the number and percentage of students in each cluster quantitatively. Based on the results displayed, Cluster 1 contains 10 students (15.2%), Cluster 2 contains 16 students (24.2%), Cluster 3 contains 25 students (37.9%) and Cluster 4 contains 15 students (22.7%).

4. CONCLUSION

Based on the research that has been conducted, students were successfully grouped into four clusters: Cluster 1 with 10 students (15.2%), Cluster 2 with 16 students (24.2%), Cluster 3 with 25 students (37.9%), and Cluster 4 with 15 students (22.7%). Each cluster represents students with different characteristics, such as high academic achievement, active participation in extracurricular activities, consistent attendance, or positive social attitudes. This approach allows schools to identify high-achieving students more objectively and efficiently.

The evaluation of the K-Means Clustering algorithm was carried out using the Elbow method, which indicated that K = 4 is the optimal number of clusters based on the Within-Cluster Sum of Squares (WCSS) value. The resulting clustering produces balanced and meaningful groupings of students. This confirms that the K-Means algorithm can be used effectively to support data-driven decision-making in selecting outstanding students, as intended in the purpose of this study.

This research contributes to the field of educational data mining by demonstrating how clustering techniques can be practically applied in school environments to support transparent and measurable evaluation processes. In addition, the system brings social benefits by minimizing subjectivity, reducing the risk of bias, and promoting fairness, so that every student has the opportunity to be assessed based on their true potential and performance.

REFERENCE

[1] M. Triandini, S. Defit, and GW Nurcahyo, "Data Mining in Measuring the Level of Student Activeness in Participating in the Learning Process at SMP IT Andalas Cendikia," 2021.

[2] I. Budiningsih and TS Dinarjo, "Hard Skill Versus Soft Skill in Achieving Employee Performance in Jakarta Mass Rapid Transit (MRT) Infrastructure Project," Akademika: Jurnal Teknologi Pendidikan, vol. 9, no. 02, 2020.

[3] O. Ubaydillah, "Teacher's Efforts in Inculcating Soft Skills and Hard Skills of Students in Learning Moral Beliefs in Madrasah Aliyah Negeri Malang. Thesis Islam Negeri University," 2019.

[4] Ikotun, Abiodun M., Absalom E. Ezugwu, Laith Abualigah, Belal Abuhajja, and Jia Heming. "K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data." Information Sciences 622 (2023): 178-210.

[5] Z. Fahmi and D. Faradika, "Decision Support System for Selecting Outstanding Students: Profile Matching Method," Journal of Technology and Business Information, vol. 1, no. 1, pp. 30-37, 2019.

[6] MR Seran, "Information Technology Forming Multiplier Effect in Corporate Business," Of Management, vol. 3, no. 2, pp. 263-274, 2016.

[7] Prastybudi, Wahyu Andy, Amalia Nur Alifah, and Arliyanti Nurdin. "Segmenting the higher education market: An analysis of admissions data using k-means clustering." Procedia Computer Science 234 (2024): 96-105.

[8] R. Nainggolan and E. Purba, "The Cluster Analysis of Online Shop Product Reviews Using K-Means Clustering," Data Science: Journal of Computing and Applied Informatics, vol. 4, no. 2, pp. 111-121, 2020.

[9] SN Arofah and F. Marisa, "Application of Data Mining to Find Out Students' Interest in Mathematics Lessons Using the K-Means Clustering Method," JOINTECS Journal of Information Technology and Computer Science, vol. 3, no. 2, p. 229, 2018.

[10] Wang, Pengxu. "Archive Information Management System Based on K-Means Clustering with Elbow Method." In 2025 3rd International Conference on Integrated Circuits and Communication Systems (ICICACS), pp.1-6. IEEE, 2025

[11] Novaliendry, D., Permana, A., Dwiyani, N., Ardi, N., Yang, C.H., Saragih, F.M. (2024). Development of a semantic text classification mobile application using TensorFlow Lite and Firebase ML Kit. Journal Européen des Systèmes Automatisés, Vol. 57, No. 6, pp. 1603-1611. https://doi.org/10.18280/jesa.570607

[12] Dony Novaliendry, Ahyanuardi, Irma Yulia Basri, Marthen Dangu Elu Beily, Fadhillah Majid Saragih, and Insan Matin Hanif, "Implementation of the Junior Mobile Programmer Professional Competency Certification Assessment Model", *Int. J. Interact. Mob. Technol.*, vol. 18, no. 21, pp. pp. 74–87, Nov. 2024.

[13] S. . Subbarayan and H. G. Gunaseelan, "A Review of Data and Document Clustering pertaining to various Distance Measures", *Salud, Ciencia y Tecnología*, vol. 2, p. 194, Dec. 2022, doi: 10.56294/saludcyt2022194.

[14] R. Radha, N. Radha, R. Swathika, N. Poongavanam, and S. Mishmala, "Time Series Analysis of Clinical Dataset Using ImageNet Classifier", *Salud, Ciencia y Tecnología*, vol. 4, p. 837, Apr. 2024, doi: 10.56294/saludcyt2024837.

[15] RP Kurniawan and Ferdiansyah, "Implementation of K-Means Clustering Algorithm to Group Students Based on Academic Scores at SMPN 207 SSN," in Proceedings of the National Seminar of Students of the Faculty of Information Technology (SENAFTI) , 2023.

[16] E. Saputra and Y. Natali, "Analysis of Student Grade Data Clustering to Determine High Achieving Students Using the K-Means Clustering Method," Journal of Information Systems and Informatics, vol. 3, no. 3, pp. 424-439, 2021.

[17] J. Albert Mayan, S.V. Manikanthan, Azham Hussain, S. Nithyaselvakumari, and A. Vinnarasi, "Clustering Technique for Mobile Edge Computing To Detect Clumps in Transportation-Related Problems", *Int. J. Interact. Mob. Technol.*, vol. 17, no. 04, pp. pp. 47–63, Feb. 2023.

[18] F. AlZyoud, M. Tarawneh, A. Almaghthawi, and A. Altalidi, "A New Approach for Cluster Head Selection in Wireless Sensor Networks", *Int. J. Onl. Eng.*, vol. 20, no. 09, pp. pp. 39–50, Jun. 2024.

[19] N.A. Maori and E. Evita, "Elbow Method in Optimizing the Number of Clusters in K-Means Clustering," Journal of Mechanical, Electrical, and Computer Science Engineering, vol. 14, no. 2, pp. 277-288, 2023.

[20] Dony Novaliendry, Tegar Wibowo, Noper Ardi, Tiolina Evi, and Dwi Admojo, "Optimizing Patient Medical Records Grouping through Data Mining and K-Means Clustering Algorithm: A

Case Study at RSUD Mohammad Natsir Solok", *Int. J. Onl. Eng.*, vol. 19, no. 12, pp. pp. 144–155, Aug. 2023.