JSCE (Journal of System and Computer Engineering)
Vol.6, No.3, July 2025, pp. 271-280
ISSN (online): 2723-1240
DOI : https://doi.org/10.61628/jsce.v6i3.2100

Implementation of the K-Means Algorithm for Clustering Students' Web Programming Course Grades Using Silhouette Score

Josua Josen A. Limbong [1,a,*]; Fervin Mayos Likumahwa [2,a]
[1] *Universitas Papua, Jl.Gunung Salju Amban, Manokwari 98314, Indonesia*
[2] *Universitas Papua, Jl.Gunung Salju Amban, Manokwari 98314, Indonesia*
[a] *jja.limbong@unipa.ac.id;* [b] *f.likumahwa@unipa.ac.id;*
* *Corresponding author*

Abstract

The development of information technology requires students majoring in informatics engineering to master web programming as one of the core competencies of the study program. Variations in students' ability to understand the material are reflected in significant differences in grades, so an objective analysis approach is needed to determine the ability of students. This study aims to group students based on academic grades in Web Programming courses using the K-Means algorithm. The data analyzed includes 1-3 assignment grades, attendance, UTS, and UAS from 32 students in the Department of Informatics Engineering, University of Papua. The research stages include preprocessing, data normalization, and clustering process using Orange Data Mining tools. Determination of the optimal number of clusters is done using the Silhouette Score method, and the best results are obtained at K = 4 with a Silhouette Score value of 0.513 which indicates a good cluster structure. The clustering results show that Cluster 1 has the highest score with a final score ranging from 0.93-1 with an Excellent score category consisting of 8 students, Cluster 2 with a Poor score category consists of 10 students with a final score range of 0.23-0.61, then Cluster 3 with a Good score category consists of 10 students with a Final score of 0.78-0.87 and Cluster 4 with a Fair score category consists of 4 students with a score range of 0.64-0.75. The results of this study provide information about the distribution of student abilities and can be the basis for improving learning strategies in the future.

Keywords *— K-Means, Silhoutte Score, Clustering, Web Programming, Orange Data Mining.*

1. Introduction

The rapid development of information technology requires students majoring in informatics engineering to master competencies in web programming as one of the key skills in the digital industry. Web programming courses are an important foundation in the curriculum because they are applicable and become the basis for developing web-based systems in the world of work. However, students' understanding of the material is a big challenge for lecturers in adjusting teaching methods. (Maulana et al., 2024) This phenomenon is often seen from the results of learning evaluations which show significant variations in scores between students, ranging from those who really master the material to those who still have difficulty in understanding basic concepts. This condition indicates the need for a more comprehensive evaluation of learning methods, where analysis of student grades can be an objective basis for assessing the effectiveness of the teaching strategies applied. This wide variation in academic achievement indicates that there is a need for evaluation in learning methods so that in the future the applied learning methods can reach all levels of student abilities.

To address this learning evaluation challenge, the K-Means algorithm was chosen as the solution in this study due to its ability to cluster data efficiently. This clustering capability is very relevant to objectively analyze the distribution of student grades so that it can be the basis for assessing the effectiveness of teaching strategies. As one of the popular unsupervised learning algorithms, K-Means works by partitioning data into clusters based on the similarity of student academic grade characteristics. The algorithm follows an iterative process consisting of randomly determining the initial centroid, calculating the distance of each data point to the centroid using Euclidean distance and updating the centroid position until convergence. The main advantage of K-Means lies in its simplicity of implementation and computational speed for small to medium sized datasets such as the student grade data in this study.

Several studies have shown the effectiveness of clustering algorithms in analyzing academic data. For example, (Mohd Talib et al., 2023) successfully applied K-Means clustering with Silhouette Score validation to identify behavioral patterns among students in higher education, which enabled targeted academic interventions and improved student support systems. Moreover, other studies have also confirmed the capability of K-Means clustering in academic contexts. The study by (Reza Pahlevi Kurniawan & Ferdiansyah, 2023) successfully grouped students based on academic grades into 3 categories namely high, medium and low grades. The clustering results in this study resulted in 16 students with the lowest scores, 92 students with medium scores, and 141 students with the highest scores. The clustering results were evaluated using the Davies Bouldin Index method which produced a value of 1.12388. Another study by (Yudhistira & Andika, 2023) grouped student grades using the attributes of student grades, discipline grades and attitude grades, then found the results of cluster 0 totaling 59 students, cluster 1 totaling 94 students, and cluster 2 totaling 1 student. Meanwhile, (Basri et al., 2023) conducted research to test the optimization of the number of clusters in the K-Means algorithm using the Elbow method based on the calculation of the Sum of Square Error (SSE). The dataset used contains GPA parameters and the number of credits to group student graduation times. Further research was conducted by (Safitri Juanita, 2024) to compare the Elbow and Silhouette methods. The results of this study prove that the Silhouette Score provides more stable results in determining the best number of clusters. Finally, (Arslan & Özdener Dönmez, n.d.) used K-Means to cluster students based on learning styles, providing a new view on how student segmentation can improve the overall learning experience.

Based on existing literature studies, there is (GAP) research that can be further analyzed, previous research in grouping student grades or students has not grouped grades based on a more complete composition of academic grades. So that in this study will use a combination of more detailed assessment variables in order to get more in-depth analysis results.

Then in this research conducted using a dataset taken from student final grade data in the odd semester of 2024 which consists of the grades of 32 students of the Informatics Engineering Department of the University of Papua, with assessment variables including, the value of assignment 1, assignment 2, assignment 3, attendance value, as well as the value of the Midterm Exam and Final Semester Exam. Then the data will be processed using the K-Means algorithm implemented through Orange Data Mining tools and determining the number of clusters using the silhoutte score method to segment students into four category groups based on grades, namely Excellent, Good, Fair, and Poor. The determination of four clusters (C=4) is based on the consideration of category groups to obtain a more detailed gradation of ability. One of the main advantages of Orange Data Mining in this research is its ability to present interactive data visualization which significantly facilitates the process of analysis and interpretation of clustering results.

The results of this study are expected to help lecturers map students' abilities objectively, develop more targeted learning strategies, and provide clear feedback for students about their academic position.

2. Method

This research begins with the data collection stage as an initial step to obtain information that is relevant and aligned with the research objectives. The collected data then undergoes a preprocessing process to clean, filter, and prepare it for further analysis. Subsequently, the number of clusters is determined based on the characteristics of the data and the intended purpose of the grouping. The K-Means algorithm is then implemented using Orange Data Mining version 3.36 on Windows 11, which facilitates automated clustering based on data similarity. The resulting clusters are analyzed to identify patterns or meaningful insights, and the research concludes with a summary of findings that support the overall objectives of the study.

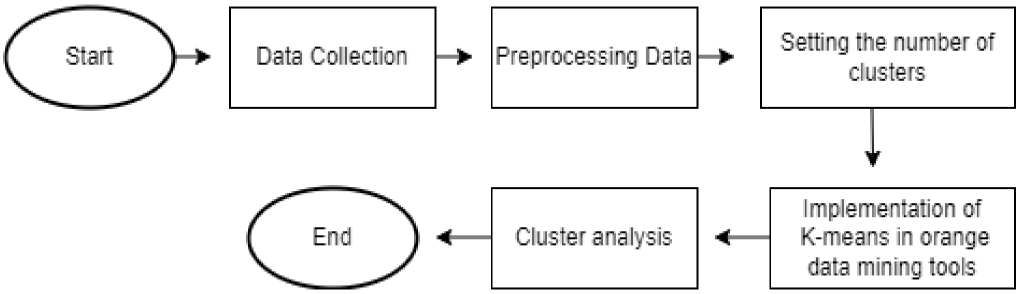


Figure 1. Research Method

2.1 Data Collection

The first stage is to collect data on student grades in web programming courses. The data used includes the value of assignments (1-3), attendance, Midterm Examination (UTS) and Final Semester Examination (UAS) from 32 students majoring in informatics engineering.

1	NIM	Tugas 1	Tugas 2	Tugas 3	UTS	UAS	Kehadiran	Nilai Akhir	Kategori Nilai
2	2024xxx37	80	75	0	0	60	75	41	Kurang
3	2021xxx67	70	80	0	30	60	75	50	Kurang
4	2022xxx17	75	80	0	0	0	75	20	Kurang
5	2022xxx34	80	80	90	80	75	100	81	Baik Sekali
6	2022xxx38	80	80	85	60	75	100	75	Baik
7	2022xxx65	80	80	0	80	75	93	73	Baik
8	2023xxx01	80	80	80	80	85	100	85	Baik Sekali
9	2023xxx03	80	80	0	80	68	93	70	Baik
10	2023xxx04	80	80	80	80	58	100	75	Baik
11	2023xxx05	80	80	90	80	58	100	75	Baik
12	2023xxx07	85	80	70	80	68	100	79	Baik
13	2023xxx08	90	80	95	80	76	100	84	Baik Sekali
14	2023xxx10	80	80	90	80	76	100	81	Baik Sekali
15	2023xxx12	80	80	0	65	72	93	67	Cukup
16	2023xxx14	100	80	90	80	75	100	83	Baik Sekali
17	2023xxx16	80	80	0	80	66	93	70	Baik
18	2023xxx18	80	80	90	75	76	100	80	Baik Sekali
19	2023xxx20	80	60	0	0	0	75	19	Kurang
20	2023xxx24	80	80	90	65	50	100	68	Cukup
21	2023xxx26	80	80	0	50	60	93	59	Kurang
22	2023xxx32	80	80	90	80	58	100	75	Baik
23	2023xxx37	80	80	0	75	75	93	71	Baik
24	2023xxx41	80	30	0	55	0	75	33	Kurang
25	2023xxx43	85	80	80	65	75	100	76	Baik
26	2023xxx47	80	80	90	86	68	100	80	Baik Sekali
27	2023xxx49	80	80	0	70	66	93	67	Cukup
28	2023xxx52	75	80	0	80	0	87	46	Kurang
29	2023xxx56	75	80	0	80	0	93	46	Kurang
30	2023xxx62	80	80	0	80	40	93	61	Cukup
31	2023xxx66	80	80	90	80	75	100	81	Baik Sekali
32	2023xxx69	80	50	0	0	0	62	17	Kurang
33	2023xxx73	80	40	0	70	0	75	39	Kurang

Figure 2. Student grade data

2.2 Preprocessing Data

After the data is collected, preprocessing is done by determining the features used to normalize the data so that it can be used in further analysis (Wongoutong, 2024). The normalization process is done to ensure that all variables have the same scale which is important for the K-Means algorithm (Ahmad Harmain et al., 2022).

2.3 Clustering

The next step is to determine the optimal number of clusters. In this research, the method used is Silhouette Score to evaluate the number of clusters that best fit the dataset. The selection of the number of clusters aims to produce meaningful and representative groupings.

2.3.1 Silhoutte Score

Silhouette Score is a clustering evaluation metric that measures how well an object is placed in its cluster compared to other clusters. Its value ranges from -1 to 1 (Shutaywi & Kachouie, 2021), where:

1. A value close to 1 indicates that the object fits perfectly into its cluster and is far away from other clusters.
2. A value of 0 means that the object is on the border between two clusters.
3. A negative value indicates that the object may be misclustered.

2.3.2 Algoritma K-Means

K-Means is a partition-based clustering algorithm that groups data into clusters by minimizing intra-cluster variance. Thus, in accordance with research conducted by (Lnc. Prakash et al., 2023), the K-means algorithm can be tested in this study to group students based on grades with excellent, good, fair and poor grade categories. The following are the steps of the K-Means Algorithm

1. Determine the value of k as the number of clusters to be formed.
2. Determine a random or random value for the initial cluster center centroid of k, to calculate the distance of each input data to each centroid using the Euclidean Distance formula, namely:
   $$d(xi,\mu j) = \sqrt{\sum(xi - \mu j)^2}$$
3. Group each data based on its proximity to the centroid or find the smallest distance.
4. Update the new centroid value, the new centroid value is obtained from the average of the cluster concerned using the formula, namely:
   $$\mu j(t+1) = \frac{1}{Nsj}\sum_{j \in sj} xj$$
5. If the data of each cluster has not stopped, repeat from steps 2 to 5, until the members of each cluster have not changed. (Nurdiyansyah & Akbar, 2021)

2.4 Cluster Analysis

The clustering results are visualized using scatter plot, silhouette plot, and box plot widgets to provide a clear picture of the distribution of data in each cluster. Analysis is carried out to understand the characteristics of each cluster such as the level of student ability to understand Web Programming material based on the student score component. The following data visualization design using orange data mining can be seen in Figure 3.

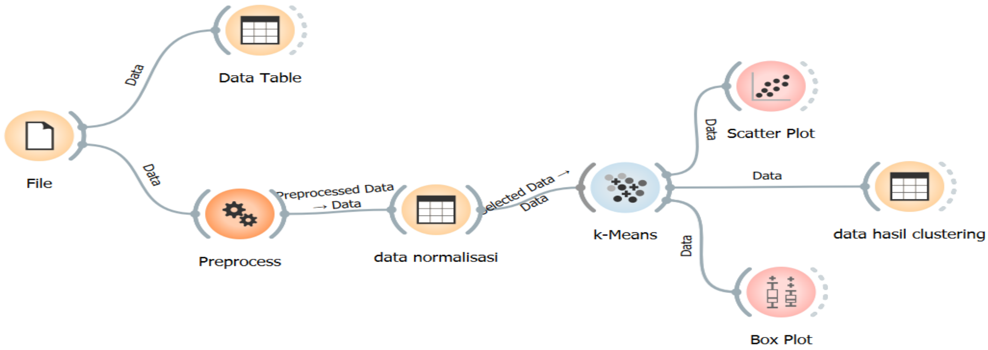


Figure 3. Data visualization design

3. Results And Discussion

Figure 4 displays student performance data from the Web Programming course, which has been imported into Orange Data Mining using the File widget and viewed through the Data Table widget. The dataset contains multiple assessment-related attributes, while the student ID (NIM) serves as the label.

Data Table - Orange

	NIM	Tugas 1	Tugas 2	Tugas 3	UTS	UAS	Kehadiran	Nilai Akhir	Kategori Nilai
1	2024xxx37	80	75	0	0	60	75	41.4167	Kurang
2	2021xxx67	70	80	0	30	60	75	50	Kurang
3	2022xxx17	75	80	0	0	0	75	20.4167	Kurang
4	2022xxx34	80	80	90	80	75	100	81.0833	Baik Sekali
5	2022xxx38	80	80	85	60	75	100	75.0833	Baik
6	2022xxx65	80	80	0	80	75	93	72.8833	Baik
7	2023xxx01	80	80	80	80	85	100	84.5833	Baik Sekali
8	2023xxx03	80	80	0	80	68	93	70.4333	Baik
9	2023xxx04	80	80	80	80	58	100	75.1333	Baik
10	2023xxx05	80	80	90	80	58	100	75.1333	Baik
11	2023xxx07	85	80	70	80	68	100	79.05	Baik
12	2023xxx08	90	80	95	80	76	100	83.5167	Baik Sekali
13	2023xxx10	80	80	90	80	76	100	81.4333	Baik Sekali
14	2023xxx12	80	80	0	65	72	93	67.3333	Cukup
15	2023xxx14	100	80	90	80	75	100	82.75	Baik Sekali
16	2023xxx16	80	80	0	80	66	93	69.7333	Baik
17	2023xxx18	80	80	90	75	76	100	79.9333	Baik Sekali
18	2023xxx20	80	60	0	0	0	75	19.1667	Kurang
19	2023xxx24	80	80	90	65	50	100	67.8333	Cukup
20	2023xxx26	80	80	0	50	60	93	58.6333	Kurang
21	2023xxx32	80	80	90	80	58	100	75.1333	Baik
22	2023xxx37	80	80	0	75	75	93	71.3833	Baik
23	2023xxx41	80	30	0	55	0	75	33.1667	Kurang
24	2023xxx43	85	80	80	65	75	100	76.1667	Baik
25	2023xxx47	80	80	90	86	68	100	80.4333	Baik Sekali
26	2023xxx49	80	80	0	70	66	93	66.7333	Cukup
27	2023xxx52	75	80	0	80	0	87	45.6167	Kurang
28	2023xxx56	75	80	0	80	0	93	46.2167	Kurang
29	2023xxx62	80	80	0	80	40	93	60.6333	Cukup
30	2023xxx66	80	80	90	80	75	100	81.0833	Baik Sekali
31	2023xxx69	80	50	0	0	0	62	17.0333	Kurang
32	2023xxx73	80	40	0	70	0	75	38.5	Kurang

32 32 | 32

Figure 4. Dataset

3.1 Data Preprocessing

The data preprocessing process begins with normalization using the Normalize to interval [0, 1] feature to equalize the scale of numerical attributes such as assignment grades, UTS, UAS, and attendance. This is important so that variables with different ranges (for example, 0-100 for attendance and 0-20 for UTS) have comparable weights in the analysis.

Next, the One feature per value option is applied to categorical columns to convert them into a binary format using one-hot encoding. This transformation ensures that each category is represented numerically without implying any ordinal relationship, allowing machine learning algorithms such as K-Means to process the data accurately. These preprocessing steps improve the consistency and comparability of the data prior to clustering. The selected features for normalization can be seen in Figure 5 below.

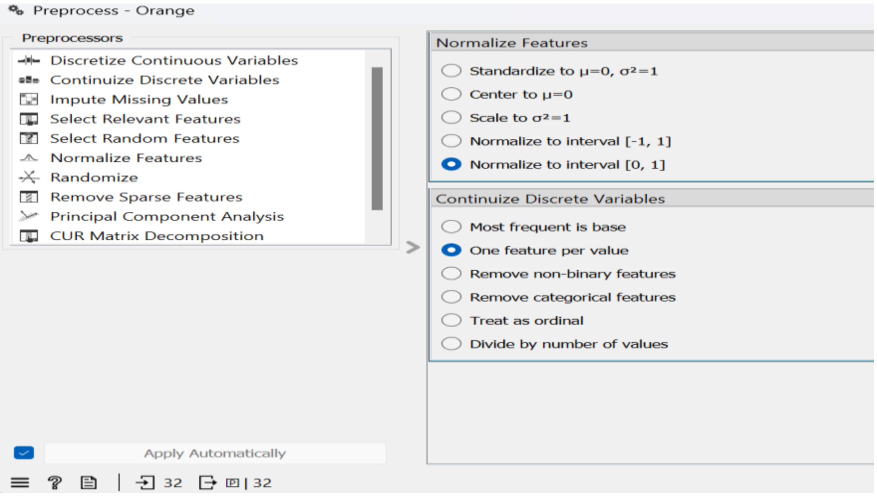
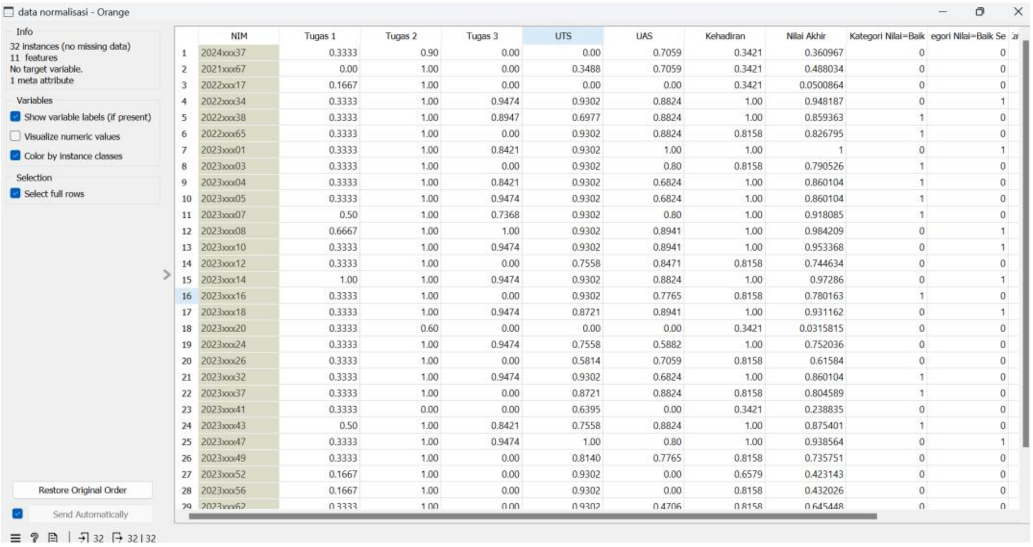


Figure 5. Preprocessing data

Then after the feature selection is done the data will be normalized so that it is ready to be used in the clustering process, the results of data normalization can be seen in Figure 6.



	NIM	Tugas 1	Tugas 2	Tugas 3	UTS	UAS	Kehadiran	Nilai Akhir	Kategori Nilai=Baik	egori Nilai=Baik Se
1	2024xxx37	0.3333	0.90	0.00	0.00	0.7059	0.3421	0.360967	0	0
2	2021xxx67	0.00	1.00	0.00	0.3488	0.7059	0.3421	0.488034	0	0
3	2022xxx17	0.1667	1.00	0.00	0.00	0.00	0.3421	0.0500864	0	0
4	2022xxx34	0.3333	1.00	0.9474	0.9302	0.8824	1.00	0.948187	0	1
5	2022xxx38	0.3333	1.00	0.8947	0.6977	0.8824	1.00	0.859363	1	0
6	2022xxx65	0.3333	1.00	0.00	0.9302	0.8824	0.8158	0.826795	1	0
7	2023xxx01	0.3333	1.00	0.8421	0.9302	1.00	1.00	1	0	1
8	2023xxx03	0.3333	1.00	0.00	0.9302	0.80	0.8158	0.790526	1	0
9	2023xxx04	0.3333	1.00	0.8421	0.9302	0.6824	1.00	0.860104	1	0
10	2023xxx05	0.3333	1.00	0.9474	0.9302	0.6824	1.00	0.860104	1	0
11	2023xxx07	0.50	1.00	0.7368	0.9302	0.80	1.00	0.918085	1	0
12	2023xxx08	0.6667	1.00	1.00	0.9302	0.8941	1.00	0.984209	0	1
13	2023xxx10	0.3333	1.00	0.9474	0.9302	0.8941	1.00	0.953368	0	1
14	2023xxx12	0.3333	1.00	0.00	0.7558	0.8471	0.8158	0.744634	0	0
15	2023xxx14	1.00	1.00	0.9474	0.9302	0.8824	1.00	0.97286	0	1
16	2023xxx16	0.3333	1.00	0.00	0.9302	0.7765	0.8158	0.780163	1	0
17	2023xxx18	0.3333	1.00	0.9474	0.8721	0.8941	1.00	0.931162	0	1
18	2023xxx20	0.3333	0.60	0.00	0.00	0.00	0.3421	0.0315815	0	0
19	2023xxx24	0.3333	1.00	0.9474	0.7558	0.5882	1.00	0.752036	0	0
20	2023xxx26	0.3333	1.00	0.00	0.5814	0.7059	0.8158	0.61584	0	0
21	2023xxx32	0.3333	1.00	0.9474	0.9302	0.6824	1.00	0.860104	1	0
22	2023xxx37	0.3333	1.00	0.00	0.8721	0.8824	0.8158	0.804589	1	0
23	2023xxx41	0.3333	0.00	0.00	0.6395	0.00	0.3421	0.238835	0	0
24	2023xxx43	0.50	1.00	0.8421	0.7558	0.8824	1.00	0.875401	1	0
25	2023xxx47	0.3333	1.00	0.9474	1.00	0.80	1.00	0.938564	0	1
26	2023xxx49	0.3333	1.00	0.00	0.8140	0.7765	0.8158	0.735751	0	0
27	2023xxx52	0.1667	1.00	0.00	0.9302	0.00	0.6579	0.423143	0	0
28	2023xxx56	0.1667	1.00	0.00	0.9302	0.00	0.8158	0.432026	0	0
29	2023xxx62	0.3333	1.00	0.00	0.9302	0.4706	0.8158	0.645448	0	0

Figure 6. Data normalization results

3.2 Clustering Data

In this clustering process begins with determining the number of clusters to be used, then the author determines 4 clusters in accordance with the predetermined value categories, namely Excellent, Good, Fair, and Poor. Student grade data that has gone through the preprocessing stage including normalization to equalize the numerical feature scale is then processed using the K-Means algorithm. This algorithm works iteratively to cluster students based on the similarity of grade characteristics by minimizing the intra-cluster distance. The clustering results are evaluated and visualized using scatter plot and box plot widgets to analyze the cluster quality and grade distribution in each group. The determination of the number of clusters can be seen in Figure 7.

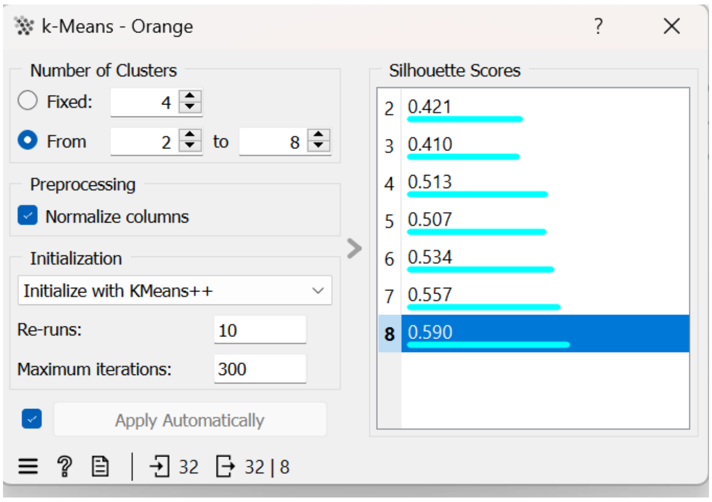


Figure 7. Silhoutte score

Figure 7 shows the evaluation results using the Silhouette Score with varying numbers of clusters from 2 to 8. Based on these results, the number of clusters of 4 has a Silhouette Score value of 0.513, which is above the threshold of 0.5 and indicates that the cluster structure is sufficiently good and clearly defined. For comparison, the Silhouette Scores for 2, 3, and 5 clusters are 0.421, 0.410, and 0.507, respectively. Therefore, 4 clusters were selected because they offer sufficient clustering quality without dividing the data into too many groups. Although the highest score was obtained at the number of clusters = 8 (0.590), the selection of 4 clusters is considered most appropriate for the purpose of segmentation based on academic categories and facilitates the interpretation of results.

Then after determining the number of clusters as many as 4 clusters in accordance with the needs of the analysis, the next stage is the clustering process with the k-means algorithm whose results are visualized using scatter plots and box plots.

3.2.1 Scatter Plot

The results of clustering analysis using scatter plots of student grade data show the existence of four groups or clusters, each of which represents the following categories of student final grades:

1. Cluster 1: This group consists of students with final grades that fall into the excellent category. This cluster is dominated by high scores on all assessment components, namely Assignments 1-3, UTS, UAS, and attendance. Students in this group show the best academic achievement among all clusters.
2. Cluster 3: Students in this cluster have final grades in the good category. In general, their assessment component scores are above average.
3. Cluster 4: This cluster represents the fair category. Students in this cluster tend to score in the middle/average range, with some assessment components that may require improvement.
4. Cluster 2: Students in this cluster have final grades in the poor category. Overall, the scores on each assessment component in this cluster tend to be low, indicating a need for more attention in the learning process.

This scatter plot visualization shows the distribution of students based on their final grades on the X-axis and the clusters on the Y-axis. The different colors for each cluster help in identifying the grade categories more clearly. The colored areas on the graph show the regional divisions of each cluster, giving a clear picture of the grouping of students' final grades.

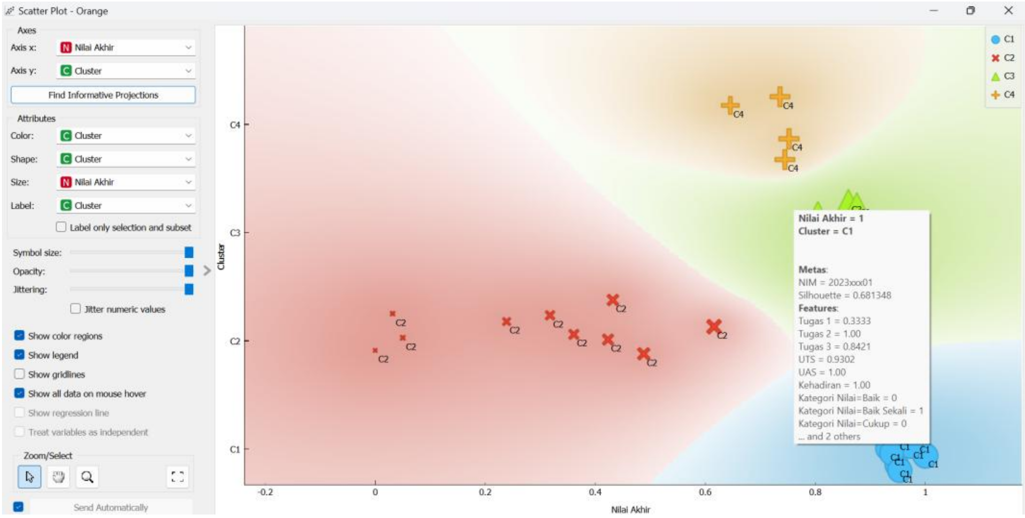


Figure 8. Scatter plot result

3.2.2 Box Plot

The box plot visualization results in Figure 9 show the distribution of student grades based on the clusters formed. Cluster 1 as many as 8 students show the best academic performance with excellent value categories in all components, while Cluster 3 as many as 10 students reflect the good value category. Cluster 4 with 4 students has an fair score category, while Cluster 2 with 10 students is in the poor category with the lowest score. The chi-square test ($\chi^2 = 96.00$, $p = 0.000$) confirmed that the difference in distribution between the clusters was statistically significant, reflecting relevant groupings based on the students' score patterns.

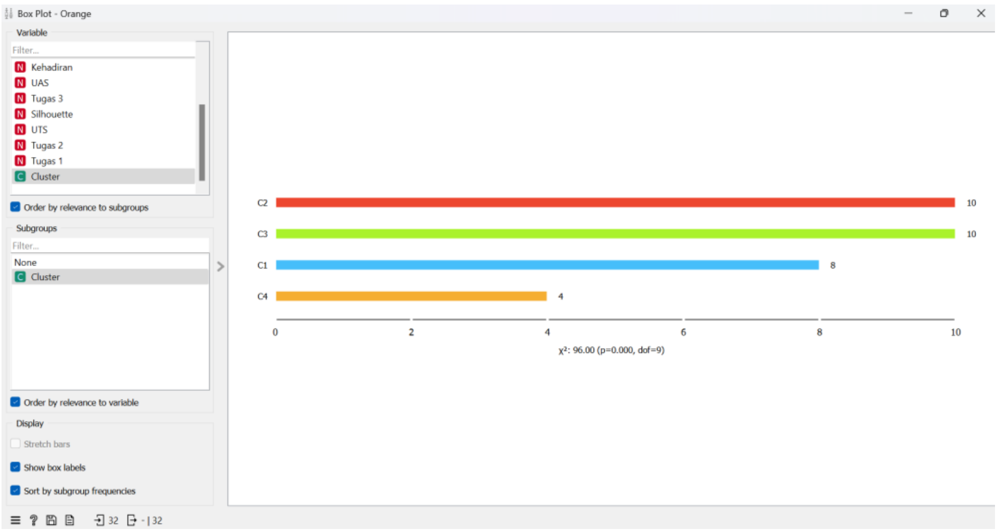


Figure 9. Box Plot

3.3 Cluster Analysis

Table 1 below presents a summary of the clustering results, showing the number of students, final score range, and Silhouette Score interval for each group. Cluster 1 (Excellent) includes students with outstanding academic performance, characterized by final scores above 0.93 and Silhouette Scores between 0.58–0.69, indicating a cohesive and well-defined cluster. These students demonstrate strong consistency across all assessment components and are well-suited for advanced learning opportunities, such as project-based assignments and independent exploration to further develop their potential. Cluster 3 (Good) consists of students with final

DOI : https://doi.org/10.61628/jsce.v6i3.2100

scores ranging from 0.78 to 0.87 and Silhouette Scores of 0.65–0.67, indicating clear separation. This group performs well and can be encouraged to engage more deeply through structured tasks and collaborative discussions to improve their achievements. Group 4 (Fair) includes students with moderate performance, characterized by final scores of 0.64–0.75 and Silhouette Scores between 0.63–0.69. Although the grouping is fairly clear, these students may require additional support such as academic guidance, reinforcement of basic concepts, and interactive learning approaches. Group 2 (Poor) consists of students with the lowest academic results (0.23–0.61) and inconsistent attendance ranging from 0.34–0.81. The Silhouette scores in this group (0.49–0.62) which still indicates an acceptable clustering structure. Students in this group need special attention through individual guidance, scaffolded learning strategies, and regular monitoring to improve participation and academic achievement.

Table 1. Clustering Result

Cluster	Category	Number of Students	Final Grade Range	Silhoutte Score Range
C1	Excellent	8	0.93-1.00	0.58–0.69
C3	Good	10	0.78-0.87	0.65–0.67
C4	Fair	4	0.64-0.75	0.63–0.69
C2	Poor	10	0.23-0.61	0.49–0.62

4. Conclusions

This study successfully applied the K-Means algorithm using Orange Data Mining to cluster 32 students into four performance-based groups in a Web Programming course. The resulting clusters demonstrated distinct academic profiles with an overall Silhouette Score of 0.513, indicating a well-structured grouping suitable for further pedagogical analysis.

The analysis results showed that Group 2 had the lowest academic performance, significantly influenced by low attendance rates ranging from 0.00 to 0.81. This finding emphasizes that attendance plays a significant role in determining students' academic outcomes across other assessment components. These results provide meaningful insights for lecturers and academic stakeholders. The cluster groupings can be utilized by instructors to design differentiated learning strategies based on students' ability levels, thereby enhancing learning effectiveness and academic support.

However, this study is limited by the small dataset, covering only one course in a single semester. Future research is recommended to expand the dataset and apply alternative clustering algorithms such as DBSCAN or hierarchical clustering. Additionally, it is suggested to evaluate cluster validity using other internal indices like the Davies-Bouldin Index for a more comprehensive comparison.

References

Ahmad Harmain, Paiman, P., Kurniawan, H., Kusrini, K., & Dina Maulina. (2022). Normalisasi Data Untuk Efisiensi K-Means Pada Pengelompokan Wilayah Berpotensi Kebakaran Hutan Dan Lahan Berdasarkan Sebaran Titik Panas. *TEKNIMEDIA: Teknologi Informasi Dan Multimedia*, *2*(2), 83–89. https://doi.org/10.46764/teknimedia.v2i2.49

Arslan, E., & Özdener Dönmez, N. (n.d.). *Clustering Students Based on Gamification User Types and Learning Styles*.

Basri, A., Mubarak, A., Siradjuddin, H. K., Abdullah, S. Do, Studi, P., Informatika, T., Teknik, F., Khairun, U., Metro, J. J., & Selatan, K. T. (2023). Penentuan Jumlah Klaster Terbaik pada K-Means dalam Melihat Pola Pengelompokan Data Mahasiswa yang Telah Lulus. *JATI (Jurnal Jaringan Dan Teknologi Informasi)* , *3*(1), 80–86. https://doi.org/00.0000/jati

Lnc. Prakash, K., Suryanarayana, G., Swapna, N., Bhaskar, T., & Kiran, A. (2023). Optimizing K-Means Clustering using the Artificial Firefly Algorithm. *International Journal of Intelligent Systems and Applications in Engineering*, *11*(9), 461–468.

Maulana, I., Adnandi, M. A., & Andriani, F. (2024). Evaluasi Penggunaan Metode Pembelajaran Berbasis Proyek Dalam Mata Kuliah Pengenalan Pemrograman: Studi Kasus. *JoiTechs (Journal Of Information Technology and Computer Science)*, *1*(1), 21–25.

Mohd Talib, N. I., Abd Majid, N. A., & Sahran, S. (2023). Identification of Student Behavioral Patterns in Higher Education Using K-Means Clustering and Support Vector Machine. *Applied Sciences (Switzerland)*, *13*(5). https://doi.org/10.3390/app13053267

Muhammad Zulfadhilah, Mambang, & Septyan Eka Prastya. (2022). Implementasi Metode K-Means Clustering untuk Meningkatkan Penjaringan Mahasiswa. *Tematik*, *9*(2), 152–160. https://doi.org/10.38204/tematik.v9i2.1053

Nurdiyansyah, F., & Akbar, I. (2021). Implementasi Algoritma K-Means untuk Menentukan Persediaan Barang pada Poultry Shop. *Jurnal Teknologi Dan Manajemen Informatika*, *7*(2), 86–94. https://doi.org/10.26905/jtmi.v7i2.6377

Reza Pahlevi Kurniawan, & Ferdiansyah. (2023). Penerapan Algoritme K-Means Clustering Untuk Mengelompokkan Siswa Berdasarkan Nilai Akademik Di Smp Negeri 207 Ssn. *Prosiding Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)*, *2*(2), 530–538.

Safitri Juanita, R. D. C. (2024). K-Means Clustering with Comparison of Elbow and Silhouette Methods for Medicines Clustering Based on User Reviews. *Jurnal Teknik Informatika (JUTIF)*, *5*(1). https://doi.org/https://doi.org/10.52436/1.jutif.2024.5.1.1349

Shutaywi, M., & Kachouie, N. N. (2021). Silhouette analysis for performance evaluation in machine learning with applications to clustering. *Entropy*, *23*(6), 1–17. https://doi.org/10.3390/e23060759

Wongoutong, C. (2024). The impact of neglecting feature scaling in k-means clustering. *PLoS ONE*, *19*(12), 1–19. https://doi.org/10.1371/journal.pone.0310839

Yudhistira, A., & Andika, R. (2023). Pengelompokan Data Nilai Siswa Menggunakan Metode K-Means Clustering. *Journal of Artificial Intelligence and Technology Information (JAITI)*, *1*(1), 20–28. https://doi.org/10.58602/jaiti.v1i1.22