Implementation of the K-Means Algorithm in Grouping Students' Learning Outcomes in Mathematics at MAN 1 Cirebon

Hani Tarkasih[1]
[1]Bimba Rainbow Kids, Jl. Widuri 4, Megu Cilik, Kec. Weru, Kabupaten Cirebon, Jawa Barat 45154

Article Info

Article history:

Received 05 18, 2025
Revised 06 07, 2025
Accepted 06 08, 2025

Keywords:

K-Means Algorithm
Learning evaluation
Learning outcomes
Data mining methods
Grouping learning outcomes

ABSTRACT

This study aims to apply the K-means Algorithm in grouping students' mathematics learning outcomes at MAN 1 Cirebon and measure its accuracy using the Davies-Bouldin Index. This type of research employs an experimental design with a quantitative approach. The research population consists of all students at MAN 1 Cirebon, with a sample of class X students majoring in Mathematics and Natural Sciences. The data used came from daily test scores and odd semester end-of-semester assessments. The results showed that the K-Means Algorithm successfully grouped student learning outcomes into three categories: Cluster 1 (high category), comprising 83 students; Cluster 2 (medium category), containing 64 students; and Cluster 3 (low category), comprising 27 students. The Davies-Bouldin Index value of 0.30 indicates that the grouping results are good quality. This research contributes to the application of data mining methods in supporting the objective and systematic evaluation of student learning outcomes.

Corresponding Author:

Hani Tarkasih
Bimba Rainbow Kids, Jl. Widuri 4, Megu Cilik, Kec. Weru, Kabupaten Cirebon, Jawa Barat 45154
Email: hanitarkasih24@gmail.com

1. INTRODUCTION

Mathematics from elementary school to college is one of the main subjects. Mathematics education has excellent power to advance superior human resources in the era of globalization. This power can be realized if mathematics education allows students to acquire mathematical concepts. Studying mathematics in primary school provides a solid foundation for entering higher education, and the role of mathematics continues in today's era of globalization. The development of science and technology cannot be separated from the role of mathematics, as mathematics significantly enhances students' ability to think logically, systematically, and orderly [1].

Mathematics education plays a vital role because mathematics is a comprehensive science applicable in all areas of life. Mathematics education aims to help students develop into individuals who can think logically, thoroughly, creatively, innovatively, critically, imaginatively, and work diligently, making mathematics a significant aspect of Indonesia's educational development [2].

Although the role of mathematics in daily life is no longer in doubt, the truth is that many students in school remain less interested and consider mathematics one of the more difficult subjects [3]. Internal and external factors can influence learning difficulties in students. Internal factors encompass physiological aspects (such as physical traits) and psychological aspects (including spiritual characteristics). Meanwhile, external factors affecting learning difficulties include social and physical environment factors [4]. These difficulties then affect students' ability to understand math material, ultimately leading to differences in learning outcomes between one student and another.

Learning outcomes refer to changes in cognitive, affective, and psychomotor abilities that students acquire after completing learning activities. They are typically demonstrated through evaluation results, as indicated by the grades teachers assign. Learning outcomes determine how much students understand the material taught [5]. Therefore, knowing learning outcomes is useful for teachers and students in managing education.

For teachers, learning outcomes can be used to assess the quality of learning that has been provided. By evaluating student learning outcomes, teachers can determine how much students have understood and mastered the material, allowing them to design more effective teaching strategies. This can ultimately help students get better grades [6]. For students themselves, knowing their learning outcomes can be a source of motivation to continue improving their learning abilities and skills. When students see progress in improved grades and a deeper understanding, it can boost enthusiasm and confidence to face the next learning challenge [7]. Additionally, students can evaluate the learning methods that have been used.

Knowing the precise and accurate grouping of student learning outcomes is critical to improving the quality of learning at MAN 1 Cirebon, especially in mathematics subjects. However, grouping student learning outcomes, especially with large amounts of data, can be complex and time-consuming. One approach that can be used to overcome this challenge is to implement algorithms.

An algorithm attempts to solve a problem using a logical and systematic sequence of operations to produce a specific output [8]. Algorithms have many types and uses in grouping learning outcomes; an algorithm is needed to group data according to its characteristics or attributes. One algorithm that can group data is the *K-means Algorithm* [9].

This study applies the K-Means algorithm to classify student learning outcomes efficiently and systematically, thereby helping schools evaluate and plan learning more effectively. This algorithm was chosen because it offers advantages in terms of speed, process simplicity, and effectiveness in handling extensive numerical data, such as student learning outcomes [10].

Several previous studies also support the selection of this algorithm. For example, research by Oladipupo and Olugbara [11] shows that K-Means has the highest potential in accurately mining information from student engagement data. In addition, according to Santosa, Lukito, and Chrismanto [12], this algorithm yields clustering results that are pretty accurate in predicting student GPA, making it highly relevant for application in educational data. Research by Goh [13] also demonstrates that the K-Means algorithm can be utilized to profile students' levels of academic achievement and segment them based on their learning performance.

The application of the K-Means Algorithm has been widely used in various studies, such as in research by Arofah and Marisa [14] to determine students' learning interests, Aditya, Jovian, and Sari [15] in grouping the results of the students' national exams, as well as the research of Jamaludin, Martanto, and Bahtiar [16] which grouped student learning outcomes to support education quality assurance during the Covid-19 pandemic.

Based on this background, this study aims to apply the K-Means Algorithm to group student learning outcomes in Mathematics subjects at MAN 1 Cirebon. Thus, this research is expected to provide practical solutions for grouping learning outcomes, thereby improving the quality of learning and academic evaluation.

2. METHODS

The research stage of data grouping is part of the KDD (*Knowledge Discovery in Databases*) process [17]. The initial stage of this research is data collection. This stage involves collecting data on mathematics learning outcomes in odd semesters, sourced from mathematics teachers at MAN 1 Cirebon. The initial dataset was then selected based on attributes such as student name, daily test scores, and end-of-semester exam scores. The population in this study consists of MAN 1 CIREBON students. Meanwhile, the sample consists of 174 students from class X MIPA.

The second stage involves data cleaning, which is performed through the data normalization process [18] using the following normalization equation.

Equation 1. Data Normalization

$$Normalization = \frac{(Initial\ score - Minimum\ score)}{(Maximum\ score - Minimum\ score)}$$

The third stage involves implementing the K-Means Algorithm, and the final stage entails evaluating the clusters using the Davies-Bouldin Index (DBI).

The K-Means algorithm is a *clustering method* that uses the *K* parameter to determine the desired number of *clusters*. The word "*means"* in K-Means refers to the average value of a cluster data group. So, *K-Means clustering* is a data analysis or data mining method used for unsupervised modeling, and is one of the partition methods that performs data grouping [19].

The steps of the K-Means algorithm can be explained as follows [20]

1. Specify the number of clusters (k) in the dataset.
2. Determine the value of the center (*centroid*). Determination of the center value (*centroid*). The determination of *the centroid* value is done randomly, but in the repetition phase, the following equation is used

Equation* 2. *Determination of Centroid Values

$$V_{ij} = \frac{1}{N_i}\sum_{k=0}^{N_i} X_{kj}$$

where

V_{ij} = Average for *cluster* i variable j

N_i = Number of cluster members i

i, k = Cluster index,

j = Variable indices

X_{kj} = The data value for *the cluster* of k variables of j

3. Calculate the shortest distance to *each dataset's centroid using the* Euclidean distance formula.

***Equation* 3. Euclidean Distance**

$$D(i,j) = \sqrt{(x_{1i} - x_{1j})^2 + \cdots + (x_{1i} - x_{1j})^n}$$

where

x_{1i} = The i-th data in the 1st data attribute.

x_{1j} = The j-th centroid in the 1st data attribute.

4. Group objects based on distance to the nearest focus.
5. Repeat steps 3 and 4 until *the centroid* is optimally valued.

After completing the grouping process, a cluster evaluation is conducted using *the Davies-Bouldin Index* (DBI). The DBI calculates the average ratio between clusters, which describes the degree of similarity within the cluster and the differences between clusters. The best number of *clusters* is indicated by a lower DBI value [21]. The DBI calculation begins by calculating the value of SSW (*Sum of Squares Within*) with the following equation [22]:

Equation* 4. *Nilai Sum of Squares Within

$$SSW_i = \frac{1}{m_i}\sum_{j-i}^{m_i} d(x_j, c_i)$$

where

m_i = The number of data points in the i-th cluster

x_j = The data points in the j-th cluster.

c_i = The centroid of the i-th cluster.

$d(x_j, c_i)$ = The distance of each data point to the i-th centroid.

Next, calculate the value of SSB (*Sum of Squares Between Clusters*) using the equations below.

***Equation* 5. Sum of Square Between Cluster value**

$$SSBij = d(c_i, c_j)$$

After that, the ratio value between the two clusters is calculated by the following formula:

***Equation* 6. Ratio**

$$R_{ij} = \frac{SSWi + SSWj}{SSBij}$$

Finally, the value DBI (*Davies Bouldien Index*) can be calculated using the following equation:

***Equation* 7. Davies Bouldien Index**

$$DBI = \frac{1}{K}\sum\nolimits_{i}^{k} max_I \neq j\ (R_{ij})$$

K = *Existing clusters*

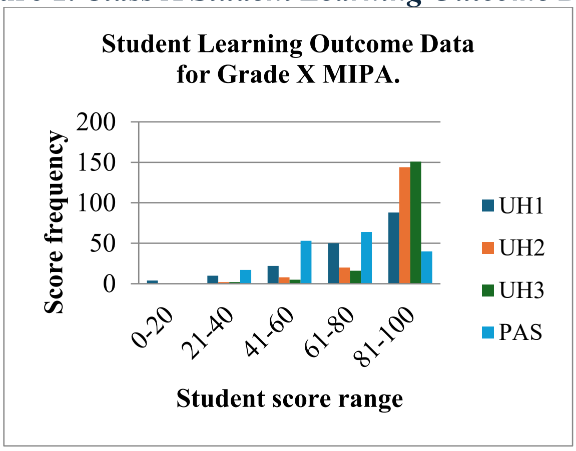
R_{ij} = rasio antara cluster i dan j

A lower DBI value indicates better grouping results, as it indicates the optimal level of compactness and separation between clusters [21].

3. RESULTS AND DISCUSSION

Applying the K-Means Algorithm in this study began with collecting data from the archives of class X mathematics teachers at MAN 1 CIREBON in the 2022/2023 odd-semester academic year. The data comprised daily tests and end-of-semester assessments (PAS) for 174 mathematics students. The attributes used in this study consisted of student names, daily test 1 (UH1), daily test 2 (UH2), daily test 3 (UH3), and end-of-semester assessment (PAS).

Figure 1. *Class X Student Learning Outcome Data*



The next step is to perform data cleaning, which involves refining data that does not conform to its attributes. No missing values or data redundancy are found at this stage, so the process can be immediately continued to the normalization stage to prevent *outliers*. Data normalization is carried out using ***Equation 1***. Here is a table of maximum and minimum values for each attribute:

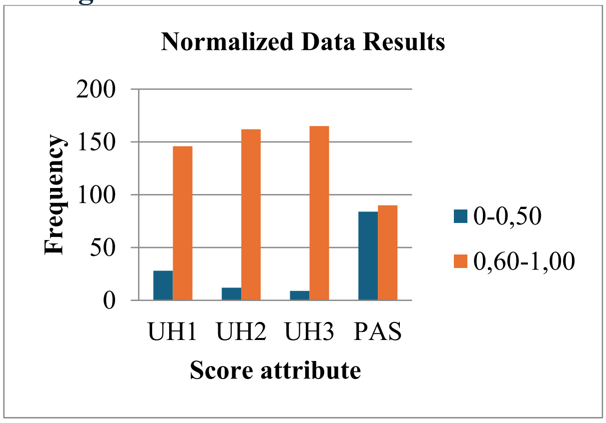
Table 1. Maximum and Minimum Values

Attribute	*Highest Score (Xmaks)*	*Lowest Score (Xmin)*
Daily Repetition 1 (X1)	100	15

Daily Repetition 9 (X2)	100	33
Daily Repetition 3 (X3)	100	30
PAS Values (X4)	100	30

The full normalization results for the following attributes are presented in the figure below.

Figure 2. Normalization Results Data



Once the data is normalized, the K-Means Algorithm is implemented using the Microsoft Excel 2016 application. The stages carried out in the K-Means Algorithm are as follows.

1) Specifies the number of clusters. By paying attention to the dataset and needs, learning outcome data can be grouped into 3 clusters.
2) Determine the initial centroid. In this initial iteration process, the centroid value is determined randomly, provided that it remains within the distance of each attribute.

Table 2. Early Centroid Values

Central Point	*Data number-*	*H1*	*H2*	*H3*	*PAS*
C1	*8*	*1*	*1*	*1*	*1*
C2	*35*	*0,8*	*0,63*	*1*	*0,66*
C3	*50*	*0,18*	*0,10*	*0,71*	*0,17*

3) Calculate data distance. Once the initial central point, the centroid, is determined, the next stage calculates the data distance for each cluster using **Equation 3**. An example of the calculation is as follows.

$$d(1,1) = \sqrt{\begin{matrix}(0,76-1)^2 + (0,91-1)^2 \\ +(1-1)^2 + (0,57-1)^2\end{matrix}} = 0,43$$

$$d(1,2) = \sqrt{\begin{matrix}(0,76-0,8)^2 + (0,91-0,63)^2 \\ +(1-1)^2 + (0,57-0,66)^2\end{matrix}} = 0,12$$

$$d(1{,}3) = \sqrt{\begin{matrix}(0{,}76 - 0{,}18)^2 + (0{,}91 - 0{,}10)^2 \\ +(1 - 0{,}71)^2 + (0{,}57 - 0{,}17)^2\end{matrix}} = 1{,}49$$

4) Group data into the same group. The next stage is to group the data into the same cluster. The cluster is determined by comparing the distance of the calculation results with the Euclidean distance, and the closest distance between the data and the cluster's center is chosen [23]. The following table provides an example of the results of grouping data by cluster.

Table 3. Iteration 1 Results

Data	*C1*	*C2*	*C3*	*Closest distance*	*Cluster*
1	*0,43*	*0,12*	*1,49*	*0,12*	*2*
2	*0,01*	*0,39*	*2,21*	*0,01*	*1*
3	*0,04*	*0,35*	*2,15*	*0,04*	*1*
4	*0,12*	*0,31*	*1,96*	*0,12*	*1*
5	*0,62*	*0,33*	*1,13*	*0,33*	*2*
…	…	…	…	….	…
172	*0,48*	*0,29*	*1,65*	*0,29*	*2*
173	*0,30*	*0,30*	*1,76*	*0,30*	*1*
174	*0,81*	*0,36*	*1,46*	*0,36*	*2*

From the first iteration of data grouping, it was found that each cluster had 39 members, Cluster 2 had 121 members, and Cluster 3 had 14 members. The clustering process has not been completed in the first iteration; it must be repeated so that the members' results in each cluster are more optimal.

5) Determination of the New Centroid of the Last Iteration

The calculation of the new centroid is carried out using the following equation, which is by summing all the data in the cluster and dividing it by the number of cluster members.

Equation 8. Centroid Formula

$$c(i,j) = \frac{The\ sum\ of\ all\ data\ in\ the\ cluster}{the\ number\ of\ cluster\ members}$$

In this study, the centroid stopped at the 7th iteration, with the value of the centroid remaining unchanged, as shown in the following table.

Table 4. Final Iteration Centroid Data

Central Point	*H1*	*H2*	*H3*	*PAS*
centroid 1	*0,96*	*0,92*	*0,97*	*0,58*
centroid 2	*0,67*	*0,77*	*0,92*	*0,52*
centroid 3	*0,30*	*0,80*	*0,72*	*0,27*

6) Cluster Results

To determine the category of each cluster, the sum of the averages of each final centroid is obtained. Then, from the average number of results to the category of each cluster [24]. Determining the category for each cluster is carried out by calculating the average of the final centroid value of each attribute, namely, Daily Exam 1 (UH1), UH2, UH3, and Final Semester Assessment (PAS). Based on the calculation results displayed in ***Table 5***, three final centroids are obtained. Cluster 1 has the highest average score of 0.86, indicating that students in this group have achieved excellent learning outcomes. Cluster 2 has an average of 0.72, which is categorized as good learning outcomes, while Cluster 3 has the lowest average of 0.52 and is classified as having poor learning outcomes.

Table 5. Cluster Category Determination

attribute	Cluster 1	Cluster 2	Cluster 3
UH1	0,96	0,67	0,30
UH2	0,92	0,77	0,80
UH3	0,97	0,92	0,72
PAS	0,58	0,52	0,27
Average	0,86	0,72	0,52

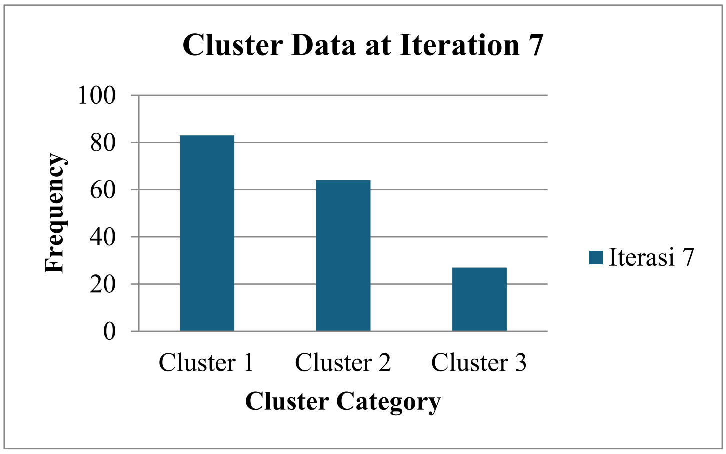
The last iteration of the K-Means algorithm produced the following distribution of the number of students in each cluster: Cluster 1 (high learning outcomes) had 83 students, Cluster 2 (medium learning outcomes) had 64 students, and Cluster 3 (low learning outcomes) had 27 students. This indicates that most students achieve high learning outcomes, while others still require additional attention in the mathematics learning process.

Furthermore, when viewed from the centroid values, students in Cluster 1 showed high consistency in scores on all daily tests and performed exceptionally well on the PAS scores. This signifies that they are active in regular learning and have a fairly in-depth understanding when facing final exams. This cluster can indicate the success of teachers' learning strategies in the classroom.

Meanwhile, students in Cluster 2 exhibited a relatively stable performance, albeit slightly lower than those in Cluster 1, particularly in UH1 and PAS scores. This can concern improving independent learning strategies or additional guidance ahead of PAS. Meanwhile, students in Cluster 3, who have the lowest average scores, especially in UH1 and PAS, require special interventions, such as remedial instruction, reinforcement of basic concepts, or a more individualized approach to learning.

These results suggest that the K-Means algorithm can aid in the educational decision-making process, for example, in designing learning strategies that cater to the needs of each student group. By understanding the characteristics of each cluster, teachers can provide more effective and targeted interventions to enhance the overall quality of learning. The results of data grouping in each cluster in iteration 7 are as follows.

Figure 3. Cluster Results



The evaluation of cluster quality is carried out by calculating the SSW (*Sum of Squares Within Cluster*) value. In calculating the SSW (*Sum of Squares Within Cluster*) value, the value to be calculated uses the last iteration with the *unchanged centroid* quality. The next step is to calculate the SSW value using **Equation 4**. Here is an example of an SSW calculation.

$$SSW1 = \frac{0{,}32+0{,}29+0{,}14+0{,}44+\cdots+0{,}21+0{,}11}{83} = 0{,}27$$
$$SSW2 = \frac{0{,}19+0{,}31+0{,}33+0{,}18+\cdots+0{,}28+0{,}37}{64} = 0{,}32$$
$$SSW3 = \frac{0{,}79+0{,}68+0{,}71+0{,}19+\cdots+0{,}33+0{,}37}{27} = 0{,}43$$

The next step is calculating the SSB (*Sum of Squares Between Clusters*). After determining the SSW value, the SSB (Sum of Squares Between Clusters*) value is calculated*. To calculate the value of SSB, a centroid is required in the last iteration. Calculate the SSB value (*Sum of Squares Between Clusters*) with **Equation 5**. Examples of SSB calculations are as follows

$$SSB_{12} = \sqrt{\begin{matrix}(0{,}96-0{,}67)^2+(0{,}92-0{,}77)^2\\ +(0{,}97-0{,}92)^2+(0{,}58-0{,}52)^2\end{matrix}} = 0{,}33$$

$$SSB_{13} = \sqrt{\begin{matrix}(0{,}96-0{,}30)^2+(0{,}92-0{,}80)^2\\ +(0{,}97-0{,}72)^2+(0{,}58-0{,}27)^2\end{matrix}} = 0{,}78$$

$$SSB_{23} = \sqrt{\begin{matrix}(0{,}67-0{,}30)^2+(0{,}77-0{,}80)^2\\ +(0{,}92-0{,}72)^2+(0{,}52-0{,}27)^2\end{matrix}} = 0{,}49$$

The results of further calculation of SSB are as follows.

Table 6. SSB Calculation Results

SSB	*1*	*2*	*3*
C1	*0*	*0,33*	*0,78*

C2	*0,33*	*0*	*0,49*
C3	*0,78*	*0,49*	*0*

After determining the value of SSB, a calculation is carried out to find the value of the ratio using **Equation 6**. An example of the Ratio calculation is as follows.

$$R_{12} = \frac{0{,}27+0{,}32}{0{,}33} = 1{,}79$$
$$R_{13} = \frac{0{,}27+0{,}43}{0{,}78} = 0{,}90$$
$$R_{23} = \frac{0{,}32+0{,}43}{0{,}49} = 1{,}52$$

The final step is to calculate the Davies-Bouldin Index (DBI) Value. After determining the Ratio value, calculations are carried out to find the value of DBI using **Equation 7**. An example of the DBI (*Davies Bouldien Index*) calculation is as follows.

$$DBI_{12} = \frac{1{,}79}{3} = 0{,}60$$
$$DBI_{13} = \frac{0{,}90}{3} = 0{,}30$$
$$DBI_{23} = \frac{1{,}52}{3} = 0{,}51$$

From the DBI calculation above, the result of DBI is 0.30 because the value is closest to 0; the smaller the DBI value, the better the cluster value.

4. CONCLUSION

Based on the research results and discussion, it can be concluded that the K-Means algorithm is effectively used to classify student learning outcomes. Three clusters with different characteristics were obtained from the clustering process. Cluster 1 consisted of 83 students with an average centroid score of 0.86, indicating that this group had a high level of learning outcomes and could be categorized as students with excellent academic achievements. Cluster 2 consisted of 64 students with an average centroid score of 0.72, indicating moderate or fairly good learning outcomes. Meanwhile, Cluster 3 consisted of 27 students with an average centroid score of 0.52, indicating low learning outcomes, which suggests that students in this group require special attention during the learning process.

The evaluation of clustering accuracy using the Davies-Bouldin Index (DBI) value yielded a value of 0.30, which falls within the good category. This value indicates that the clusters formed exhibit clear separation and a compact internal structure, making the resulting grouping valid enough to serve as a basis for educational decision-making. These findings suggest that data-mining approaches, such as K-Means, can be used to assist teachers in understanding the distribution of students' abilities and developing more targeted learning strategies based on learning group profiles.

REFERENCES

[1] M. Surjana, N. Riastini dan Y. Pustika, “Penerapan Pendekatan Kontekstual Berbantuan Media Konkret Untuk Meningkatkan Aktivitas Dan Hasil Belajar,” *International Journal Of Elementary Education,* 2017.

[2] E. K. Hasibuan, “Analisis kesulitan belajar matematika siswa pada pokok bahasan bangun ruang sisi datar di smp negeri 12 bandung,” *Jurnal Pendidikan dan Matematika,* 2018.

[3] N. Ainin, "Analisis Kesalahan Siswa Dalam Menyelesaikan Soal Matriks dan Kaitannya dengan Motivasi Belajar Matematika Pada Kelas XI," *Euclid,* 2020.

[4] M. Syah, Psikologi Pendidikan, Bandung: PT Remaja Rosdakarya, 2010.

[5] P. Evaluasi Hasil Belajar, Yogyakarta: Pustaka Belajar, 2010.

[6] A. R. Z. Wati e S. Trihantoyo, “Strategi Pengelolaan Kelas Unggulan Dalam Meningkatkan Prestasi Belajar Siswa,” *Jurnal Dinamika Manajemen Pendidikan,* vol. 5, 2020.

[7] L. Lomu e S. A. Widodo, “Pengaruh Motivasi Belajar dan Disiplin Belajar terhadap Prestasi Belajar Matematika Siswa,” em *Prosiding Seminar Nasional Pendidikan Matematika Etnomatnesia*, 2018.

[8] K. Modul pengantar algoritma dan pemrograman, Tangerang Selatan: Universitas Terbuka, 2020.

[9] Z. Mustofa e I. S. Suasana, "Algoritma clustering K-medoids Pada e-government Bidang information and communication technology Dalam Penentuan status Edgi," *Jurnal Teknologi Informasi dan Komunikasi,* 2018.

[10] W. e R. S. Wahono , “Algoritma Cluster Dinamik untuk Optimasi Cluster pada Algoritma K-Means dalam Pemetaan Nasabah Potensial,” *Journal of Intelligent Systems,* p. 33, 2015.

[11] O. O. Oladipupo e O. O. Olugbara, "Evaluation of data analytics based clustering algorithms for knowledge mining in a student engagement data,” *Intelligent Data Analysis,* vol. 23, nº 5, p. 1055, 2019.

[12] R. G. Santosa, Y. Lukito e . A. R. Chrismanto, “Classification and Prediction of Students’ GPA Using KMeans Clustering Algorithm to Assist Student Admission Process,” *Journal of Information Systems Engineering and Business Intelligence,* vol. 7, 2021.

[13] Y. L. Goh, Y. H. Goh, C. Y. Yip, C. H. Ting, R. L. L. Bin e K. P. Chen, “Prediction of students' academic performance by K-means clustering,” *Science Proceedings Series,* 2020.

[14] S. N. Arofah e F. Marisa, “Penerapan Data Mining untuk Mengetahui Minat Siswa pada Pelajaran Matematika menggunakan Metode K-Means Clustering,” *OINTECS (Journal Inf. Technol. Comput),* 2018.

[15] A. Aditya, . I. Jovian e . B. N. Sari , “Implementasi K-Means Clustering Ujian Nasional Sekolah Menengah Pertama di Indonesia Tahun 2018/2019,” *Jurnal Media Informatika Budidarma,* 2020.

[16] M. Jamaludin, M. Martanto e A. Bahtiar, “Pengelompokan Hasil Belajar Siswa pada Masa Covid-19 dengan Algoritma K-Means untuk Menjamin Mutu Pendidikan di SMK Bina Cendekia,” *JURSIMA,* 2022.

[17] C. Nas, “Data Mining Prediksi Minat Calon Mahasiswa Memilih Perguruan Tinggi Menggunakan Algoritma C4.5,” *Jurnal Manajemen Informatika (JAMIKA),* 2021.

[18] N. Puspitasari dan H. , “Penerapan metode k-means dalam pengelompokkan curah hujan di kalimantan timur,” *Proceedings seminar nasional riset ilmu komputer (SNRIK) Universitas Muslim Indonesia,* p. 22, 2016.

[19] F. Yunita, “Penerapan data mining menggunakan algoritma k-means clustering pada penerimaan mahasiswa baru (studi kasus : universitas islam indragiri),” *Jurnal SISTEMASI,* p. 240, 2018.

[20] A. Wanto, M. N. H. Siregar, A. P. Windarto, D. Hartama, D. Napitupulu, E. S. Negara dan M. R. Lubis, Data mining : algoritma dan implementasi, Medan: Yayasan Kita Menulis, 2020.

[21] A. F. Khairati, A. A. Adlina, G. F. Hertono dan B. D. Handari, “Kajian indeks validitas pada algoritma k-means enhanced dan k-means MMCA,” *Prosiding Seminar Nasional Matematika,* 2019.

[22] R. J. Hablum, A. Khairan e R. Rosihan, “Clustering Hasil Tangkap Ikan Di Pelabuhan Perikanan Nusantara (Ppn) Ternate Menggunakan Algoritma K-Means,” *JIKO (Jurnal Informatika dan Komputer,* 2019.

[23] A. Sulistiyawati e E. Supriyanto, “Implementasi Algoritma K-means Clustring dalam Penetuan Siswa Kelas Unggulan,” *Jurnal TEKNO KOMPAK,* 2021.

[24] S. Anwar, T. Suprapti, G. Dwilestari e A. Ali, “Pengelompokkan Hasil Belajar Siswa dengan Metode Clustering K-Means,” *Jurnal Riset Sistem Informasi dan Teknologi Informasi (JURSISTEKNI),* 2022.

[25] M. A. Misri, Teori Grup, Cirebon: CV, 2018.

[26] M. A. Misri, K. H. Qadr e M. A. Rahmatullah, "SEIR Mathematical Model with the Use of Hand Sanitizers to Prevent the Spread of Covid-19 Disease," *CAUCHY: Jurnal Matematika Murni dan Aplikasi,* vol. 9, nº 1, pp. 138-154, 2024.

[27] R. Hariyani, M. A. Misri e H. Handoko, "Implementasi Pewarnaan Graf Menggunakan Algoritma Welsh-Powell pada Peta Indonesia," *JURNAL SILOGISME: Kajian Ilmu Matematika dan Pembelajarannya,* vol. 9, nº 2, pp. 94-103, 2024.