Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 3, July 2022
DOI : https://doi.org/10.33395/sinkron.v7i3.11458
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Superior Class to Improve Student Achievement Using the K-Means Algorithm

Yopi Hendro Syahputra[1)], Juniar Hutagalung[2)*]
[1)2)]STMIK Triguna Dharma, Medan, Indonesia
[1)]yopihendro@gmail.com, [2)]juniarhutagalung991@gmail.com

Submitted : June 9, 2022 | **Accepted** : July 13, 2022 | **Published** : July 14, 2022

Abstract: The accumulation of new student data every year makes searching and processing data difficult, including selecting superior class students according to their talents and abilities. Therefore, the application of the K-Means Clustering data mining method is carried out to support decisions in grouping superior classes. The report card values for each class were used as parameters with a data sample of 80 students and 3 clusters were taken which then resulted in the selection and distribution of superior classes. The purpose of the study was to classify students in the superior class so that they could improve student achievement at SMK Raksana 2 Medan. Results Based on the calculation of the variable distance at the initial centroid with a sample of 80 students and the third iteration, the WCV value is 360.9745 and the BCV value is 7.3575 with a ratio value of 0.0203. Each cluster, namely: Cluster 1 has 43 students including the superior class category. Cluster 2 has 18 students and Cluster 3 has 19 students. Clusters 2 and 3 are included in the regular class category with a total of 37 students. The web-based K-Means application can provide information and solutions needed by schools to classify and determine superior classes so that they can improve student achievement in schools. These results can be used by the school to analyze student achievement and can assist teachers in forming superior classes so as to motivate students to study harder.

Keywords: Clustering; Iteration; K-Means; Student, Superior Class

INTRODUCTION

The role of education in a nation aims to form educated, competent, and characterized human resources (Uly Aldini & Wara Pramesti, 2020). The success of education lies in how to carry out the functions of the learning management program. Both from planning, organizing, implementing and evaluating the teaching and learning process. To achieve these objectives through the implementation of superior class programs effectively and efficiently (Hakim, 2021). Data science-based computing can be used to classify outstanding students so as to make it easier for schools to make objective decisions (Primanda et al., 2021).

SMK Raksana 2 Medan is a vocational high school that has a good learning system and implements a superior class student program. The accumulation of new student data every year makes searching and processing data difficult, including selecting superior class students according to their talents and abilities.

Therefore, the application of data mining using the K-Means method is carried out to support decisions in the superior class grouping by referring to the students' best scores. The report card values for each class are used as attributes and 3 clusters are taken and then result in the selection and division of superior classes.

The application of data mining techniques with the web-based K-Means method is carried out so that data processing is more optimal and becomes a strategic source of information. This method can help schools that have a lot of data to produce information that can support objective decision making that refers to student scores. Thus the school or teacher knows the achievement of student achievement both academically and through extracurricular activities, in order to improve the quality of education at SMK Raksana 2 Medan.

The system used to determine the value of student learning outcomes in the grouping of superior class students at SMK Raksana 2 Medan, was obtained by calculating the average value of knowledge, skills and expertise, extracurricular and discipline competencies based on student report cards.

Various data mining methods can be used for grouping, classification, regression, variable selection, association and analysis (Hutagalung & Sonata, 2021). Clustering is a method for grouping data that have similar characteristics and dividing the data into several groups (Hardianti. siti et al., 2018). K-Means is one of the

*Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 3, July 2022 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v7i3.11458 p-ISSN : 2541-044X

clustering algorithms used in unsupervised learning groups which is used to classify data into several classes with system partitions. This algorithm accepts data entry in the form of class labels (Hutagalung et al., 2021).

The application of the K-Means method on other references, obtained the results of grouping the selection of superior class students, where the center point no longer changes and there is no data moving between clusters (Satria & Anggrawan, 2021). This method can solve problems that often occur in processing student data, teachers, employees and student grades in determining students who are entitled to superior class (Kusuma & Aryati, 2019). Analyzing student achievement by grouping low, medium, and high achievement categories, so that students who excel and form an ideal class to improve achievement and motivate students (Saputra & Nataliani, 2021). Data processing performed using K-Means can produce groups of outstanding students based on grades, so that the school curriculum is easier to select students to enter superior classes (Azaria Bella Bernissa, 2020). The application of K-Means resulted in the clarification of the superior class formed in 2 clusters for each class from a total of 192 data tested (Sulistiyawati & Supriyanto, 2020).

The grouping of students who are entitled to enter the superior class by applying the K-Means method is quite effective. So for this study, using the K-Means method for the cluster process where pre-processing will be carried out first to process the superior class student data which aims to produce better accuracy in the superior class grouping so as to improve student learning achievement.

This study aims to classify superior class students to improve student achievement at SMK Raksana 2 Medan, so that an improvement in the learning process can be created according to the talent value of students' abilities to achieve optimal results. The K-Means method is applied to process data to make it more useful and as a tool to facilitate decision making for the school, it can also produce superior class divisions according to students' abilities so as to positively improve the quality of education.

LITERATURE REVIEW

Data mining is one of the broad fields of computer science and the availability of large amounts of data, as well as the urgent need to turn data into useful information and knowledge (Hussain et al., 2018). The purpose of data mining is to gain knowledge that is still hidden in chunks of data (Nasyuha et al., 2021). Application of data mining techniques in digging knowledge can be seen as formative evaluation technique (Desiani et al., 2020). In many cases, this can described as the preprocessing stage of pattern recognition (Riana et al., 2022). Data mining is the best option used by researchers in data processing to fully recognize problems and solve them by adopting data mining techniques (Triayudi & Fitri, 2021).

Grouping is one of the method of grouping data based on similarity data properties. If the data do not have the same properties, then it will be grouped in another group (Wibowo et al., 2021). In this clustering process, the processed data is data pre-processing results (Faesal et al., 2020). The development of data grouping has a simple and close nature to humans way of thinking every time big data is presented tends to summarize the amount of data (Salbinda & Handayani, 2022). Data cluster evaluation is an important task in knowledge discovery and data mining through the process of creating data clusters based on the similarity of data from large data sets (Hossain et al., 2019). Clustering is also known as data segmentation, grouping has widely used in various fields (Monalisa & Kurnia, 2019).

METHOD

K-means is one part of the clustering algorithm, including on unsupervised learning (Tempola et al., 2020). K-Means is an algorithm that groups objects with the same characteristics into a cluster, which is determined repeatedly by the value of k (Amalia et al., 2021). Data mining process using K-Means is aimed at grouping data into each cluster which corresponds to the center point (centroid) from each cluster used to identify patterns (Puspitasari et al., 2020).

Distance matrix has a very important role in processing grouping on the K-means algorithm (Sugriyono & Siregar, 2020). In K-means the quality of the grouping results can run consistently in all the distance functions tested. The distance with the best evaluation result is the Euclidean distance (Eliyanto & Surono, 2021). Use of K-Means algorithm is very sensitive to initialize the cluster center because it is already done randomly and using the mean value as the center of the cluster (Sembiring Brahmana et al., 2020).

The flowchart of the K-means algorithm is illustrated in Figure 1 (Seta & Hartomo, 2020).

*Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 3, July 2022
DOI : https://doi.org/10.33395/sinkron.v7i3.11458
e-ISSN : 2541-2019
p-ISSN : 2541-044X

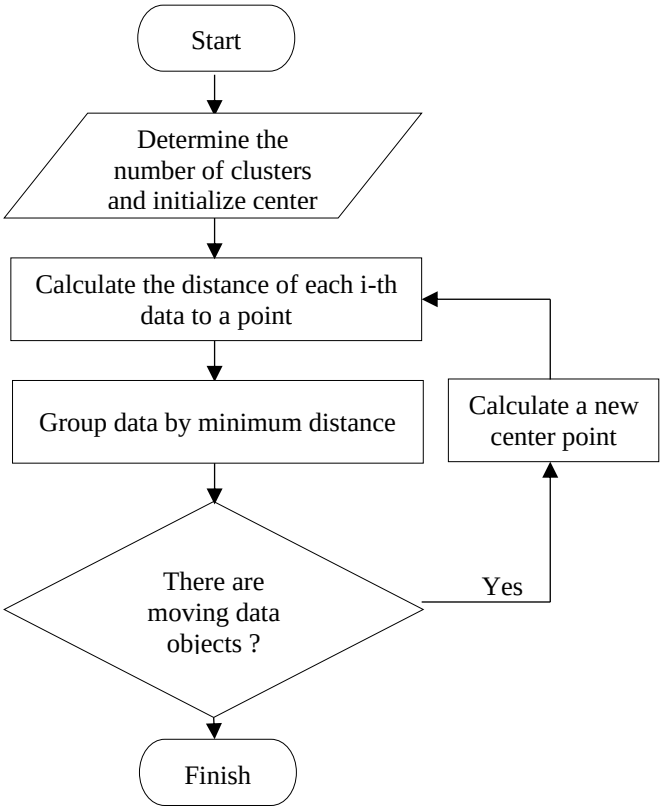

Figure 1. K-means Algorithm Flowchart

The stages of the algorithm in the K-Means clustering are as follows (Sari et al., 2019):
1. Select k centroid point randomly.
2. Group the data to form k clusters with the centroid point of each cluster is the centroid point that has been selected previously.
3. Update the centroid point value.
4. Repeat steps 2 and 3 until the value from the centroid point no longer changes.

In this study, the method used is k-means. Figure 2 below is a research framework.

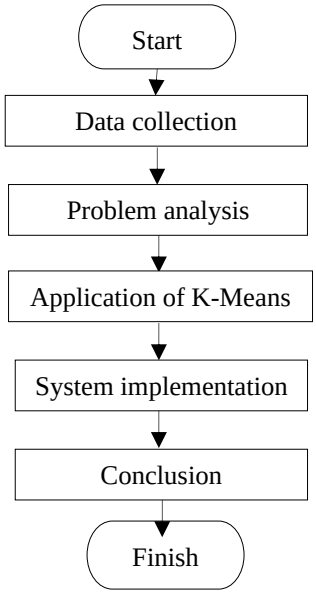

Figure 1. Research Framework

*Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 3, July 2022
DOI : https://doi.org/10.33395/sinkron.v7i3.11458
e-ISSN : 2541-2019
p-ISSN : 2541-044X

The following is an explanation of Figure 2 above:

For the beginning of the study, data was collected through observation by directly observing SMK Raksana 2 Medan. Conducting direct question and answer with the school, Ibu Fatimah, S. Kom as a teacher for the RPL class and studying literature by collecting references from several journals and books. Then analyze the problem in classifying superior classes to improve student achievement and find solutions to these problems by applying the K-Means algorithm. Furthermore, it is implemented in a web-based application. The final step is to provide conclusions from the results of the application of the K-Means algorithm.

RESULT

The data in this study went through the data pre-processing stage before being used to carry out the clustering process. Grouping of data related to superior class to improve student achievement using K-Means, with a sample of 80 students and 3 clusters, namely "High, Medium and Low". In determining this superior class, several parameters are used as determinants to get the superior class, namely the value of knowledge, skills and expertise, extracurricular and discipline competencies based on student report cards.

Initialization of knowledge assessment data, competency skills and expertise, extracurricular and discipline, according to table 1 below.

Table 1. Initialization of Parameter Assessment

Value	Description
≤ 50	1
51-60	2
61-70	3
71-80	4
81-100	5

Student assessment data in accordance with predetermined parameters can be seen in table 2 below. Data source from SMK Raksana 2.

Table 2. Student Assessment Data

Number	Student's name	Knowledge	Skills	Skill Competence	Extracurricular	Discipline
1	Sakinah	72	75	70	82	81
2	Shakira Safitri	75	73	78	72	70
3	Adiba Khairiyah	83	71	85	66	80
4	Arumi Herliana	69	82	73	63	68
5	Balkis Kiranah	81	77	80	81	73
6	Intan Natasha	74	63	78	84	72
7	Fikri Anggar	83	79	83	92	81
8	Muhammad Fadhil	50	63	57	73	61
9	Nazra Saskia	82	92	88	74	72
10	Gali Prawira	70	77	61	52	64
11	Muhammad Ikhsan	64	73	53	61	63
12	Naila Aliza	71	82	72	63	72
13	Nur Laila	83	92	76	70	78
14	Rizky Afrianda	75	81	69	80	74
15	Andini Syahputri	63	72	53	81	78
…	…	…	…	…	…	…
80	Wizan Restu	69	63	76	68	79

The data in data table 2 above can be processed using the web-based K-Means algorithm, then normalized or cleaned of unused data and initialize the data.

The following table 3 is student assessment data normalized based on predetermined values using the web-based K-Means algorithm.

*Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 3, July 2022
DOI : https://doi.org/10.33395/sinkron.v7i3.11458
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Table 3. Student Assessment Normalization Data

Number	Student's name	Knowledge	Skills	Skill Competence	Extracurricular	Discipline
1	Sakinah	4	4	3	5	5
2	Shakira Safitri	4	4	4	4	3
3	Adiba Khairiyah	5	4	5	3	4
4	Arumi Herliana	3	5	4	3	3
5	Balkis Kiranah	5	4	4	5	4
6	Intan Natasha	4	3	4	5	4
7	Fikri Anggar	5	4	5	5	5
8	Muhammad Fadhil	1	3	2	4	3
9	Nazra Saskia	5	5	5	4	5
10	Gali Prawira	3	4	3	2	3
11	Muhammad Ikhsan	3	4	2	3	3
12	Naila Aliza	4	5	4	3	4
13	Nur Laila	5	5	4	3	4
14	Rizky Afrianda	4	5	3	4	4
15	Andini Syahputri	3	4	2	5	4
…	…	…	…	…	…	…
80	Wizan Restu	3	3	4	3	4

Clustering alternative data using web-based K-Means clustering, here are the steps:

1. Determine the Number of Clusters = 3 and Maximum Iteration = 100
2. Determine the Cluster Center

After determining the number of clusters formed as much as 3 (three), determine the center (cluster). The initial centroid was chosen randomly based on the existing data. The following centroids are selected in table 4.

Table 4. Initial Cluster Center View

Cluster	Knowledge	Skills	Skill Competence	Extracurricular	Discipline
C1	3	3	3	3	3
C2	1	2	1	3	2
C3	4	2	3	3	1

3. Determining the Distance to the Cluster Center

Calculate the data distance on each centroid with Euclidean Distance equation according to equation 2, its implementation can be seen in table 5.

$$d\left(x_j, c_j\right) = \sqrt{\sum_{j=1}^{n}(x_j, c_j)^2} \quad (1)$$

Table 5. Display of Distance to Cluster Center

		Cluster		
Number	Student's name	C1	C2	C3
1	Sakinah	3, 1622	5.477	4,8989
2	Shakira Safitri	2	4,8989	3,1622
3	Adiba Khairiyah	3, 1622	6,3245	4.2426

*Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 3, July 2022
DOI : https://doi.org/10.33395/sinkron.v7i3.11458
e-ISSN : 2541-2019
p-ISSN : 2541-044X

4	Arumi Herliana	2, 2360	4,7958	3,8729
5	Balkis Kiranah	3,3166	6,0827	4,3588
6	Intan Natasha	2.6457	5.1961	3.8729
7	Fikri Anggar	4.1231	7	5.3851
8	Muhammad Fadhil	2,4494	2	4
9	Nazra Saskia	3.7416	6,7823	4,8989
10	Gali Prawira	1,4142	3,7416	3,1622
11	Muhammad Ikhsan	1,4142	3,1622	3,1622
12	Naila Aliza	2,6457	5.5677	4,3588
13	Nur Laila	3,1622	6,1644	4,4721
14	Rizky Afrianda	2,6457	5,1961	4,3588
15	Andini Syahputri	2,6457	4,1231	4,3588
…	…	…	…	…
80	Wizan Restu	1,7320	2,6457	2,2360

4. Determining Cluster Membership and Minimum Distance.
   The following determines the cluster and drinking distance can be seen in table 6 below.

Table 6. Display of Cluster Membership and Minimum Distance

Number	Student's name	Cluster	Minimum Distance	Minimum Square of Distance
1	Sakinah	C1	3, 1622	10
2	Shakira Safitri	C1	2	4
3	Adiba Khairiyah	C1	3, 1622	10
4	Arumi Herliana	C1	2, 2360	5
5	Balkis Kiranah	C1	3,3166	11
6	Intan Natasha	C1	2.6457	7
7	Fikri Anggar	C1	4.1231	17
8	Muhammad Fadhil	C2	2	4
9	Nazra Saskia	C1	3.7416	14
10	Gali Prawira	C1	1,4142	2
11	Muhammad Ikhsan	C1	1,4142	2
12	Naila Aliza	C1	2,6457	7
13	Nur Laila	C1	3,1622	10
14	Rizky Afrianda	C1	2,6457	7
15	Andini Syahputri	C1	2,6457	7
…	…		…	…
80	Wizan Restu	C1	1,7320	3

5. Calculate the value of WCV (Within Cluster Variation) with Raise the nearest cluster distance and add up each.
   WCV value = 360.9745

6. Calculate the value of BCV (Between Cluster Variation) by add up the result of the distance between each centroid.
   BCV value = 7.3575

Calculate the value of the ratio by comparing the value of BCV and WCV, the results can be seen on the display following web-based implementation in table 7 below.

Table 7. Display of BCV Results and Ratio

Cluster		Distance Between Clusters
C1	C2	2,6168
C1	C3	1,8927
C2	C3	2,8480
BCV		7,3575

*Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 3, July 2022
DOI : https://doi.org/10.33395/sinkron.v7i3.11458
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Ratio	0,0203

Continue iteration until the position of the cluster member no longer exists which changed again. Proceed to the second iteration with steps the same as the first iteration. The final result of the calculation using a -based system. The web can be seen in table 8 below.

Table 8. Display of Final Cluster Membership Results

Number	Student's name	Cluster
1	Sakinah	C1
2	Shakira Safitri	C1
3	Adiba Khairiyah	C1
4	Arumi Herliana	C1
5	Balkis Kiranah	C1
6	Intan Natasha	C1
7	Fikri Anggar	C1
8	Muhammad Fadhil	C2
9	Nazra Saskia	C1
10	Gali Prawira	C1
11	Muhammad Ikhsan	C1
12	Naila Aliza	C1
13	Nur Laila	C1
14	Rizky Afrianda	C1
15	Andini Syahputri	C1
…	…	
80	Wizan Restu	C1

Based on the calculation of the variable distance on the centroid then each cluster has the following members:

C1 has 43 members
C2 has 18 members
C3 has 19 members

DISCUSSIONS

Based on the calculation of the variable distance at the initial centroid with a data sample of 80 students until iteration 3, the WCV (Within Cluster Variation) value is 360.9745 and the BCV (Between Cluster Variation) = 7.3575 with a ratio value = 0.0203. Each cluster has the following members: Cluster 1 has 43 students in the superior class category. Cluster 2 has 18 students. Cluster 3 has 19 students. Clusters 2 and 3 are included in the category for the regular class having a total of 37 students. With the web-based K-Means application, it can provide information and solutions needed by schools to classify and determine superior class students so that they can improve student achievement at school. These results can be used by schools to analyze student achievement and can help teachers and homeroom teachers in forming an ideal class in the sense of motivating students to study harder.

CONCLUSION

From the results of the research conducted, then the following conclusions are drawn: System results by applying web-based k-means for the grouping of superior class students using 3 clusters with a data sample of 80 students, Superior Class Cluster (C1), with a high score of 43 students. Clusters (C2) and (C3) with medium and low grades are included in the regular class category with 37 students. With the web-based K-Means application, you can provide information and solutions needed by schools to classifying and determining superior class students so as to improve student achievement in school. These results can be used by schools to analyze the achievements of students and can help teachers and homeroom teachers in forming an ideal class in the sense of motivating students to study harder. The featured class here is not to make a regular class a stepchild. However, as a form of the school's seriousness in increasing competitiveness in the academic and non-academic fields in a special way. In terms of recruitment itself, the superior class has to go through a different series of selections from the regular class. Regular classes are still provided with services that have been programmed by the school.

*Corresponding author

ACKNOWLEDGMENT

Thanks to the Kemenristekdikti (Ministry of Research, Technology and Higher Education) and DRPM (Directorate of Research and Community Service) who have provided assistance funding for this scientific publication. Based on contract Beginner Lecturer Research Research (PDP) Fiscal Year 2022.

REFERENCES

Amalia, N. L. R., Supianto, A. A., Setiawan, N. Y., Zilvan, V., Yuliani, A. R., & Ramdan, A. (2021). Student Academic Mark Clustering Analysis and Usability Scoring on Dashboard Development Using K-Means Algorithm and System Usability Scale. *Jurnal Ilmu Komputer Dan Informasi*, *14*(2), 137–143. https://doi.org/10.21609/jiki.v14i2.980.

Azaria Bella Bernissa, F. (2020). Implementasi Algoritma K-Means Untuk Menentukan Kelas Unggulan Pada Smpn 1 Bojong. *EProsiding Sistem Informasi (POTENSI)*, *1*(1), 128–135.

Desiani, A., Yahdin, S., & Rodiah, D. (2020). Prediksi Tingkat Indeks Prestasi Kumulatif Akademik Mahasiswa dengan Menggunakan Teknik Data Mining. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *7*(6), 1237. https://doi.org/10.25126/jtiik.2020722493.

Eliyanto, J., & Surono, S. (2021). Distance Functions Study in Fuzzy C-Means Core and Reduct Clustering. *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika*, *7*(1), 118. https://doi.org/10.26555/jiteki.v7i1.20516.

Faesal, A., Muslim, A., Ruger, A. H., & Kusrini, K. (2020). Sentimen Analisis Terhadap Komentar Konsumen Terhadap Produk Penjualan Toko Online Menggunakan Metode K-Means. *MATRIK : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, *19*(2), 207–213. https://doi.org/10.30812/matrik.v19i2.640.

Hakim, L. (2021). Manajemen Program Kelas Unggulan untuk Meningkatkan Citra Madrasah Tsanawiyah 1 Kabupaten Madiun. *Southeast Asian Journal of Islamic Education Management*, *2*(1), 1–14. https://doi.org/10.21154/sajiem.v2i1.37.

Hardianti. siti, Sinawati, & Praseptian M, D. (2018). Implementasi Clustering dengan Metode Minimum Spanning Tree Untuk Pengelompokkan siswa berdasarkan nilai hasil studi. *Journal of Big Data Analytic and Artificial Intelligence*, *4*(1), 23–28.

Hossain, M. Z., Akhtar, M. N., Ahmad, R. B., & Rahman, M. (2019). A dynamic K-means clustering for data mining. *Indonesian Journal of Electrical Engineering and Computer Science*, *13*(2), 521–526. https://doi.org/10.11591/ijeecs.v13.i2.pp521-526.

Hussain, S., Dahan, N. A., Ba-Alwib, F. M., & Ribata, N. (2018). Educational data mining and analysis of students' academic performance using WEKA. *Indonesian Journal of Electrical Engineering and Computer Science*, *9*(2), 447–459. https://doi.org/10.11591/ijeecs.v9.i2.pp447-459.

Hutagalung, J., & Sonata, F. (2021). Penerapan Metode K-Means Untuk Menganalisis Minat Nasabah Asuransi. *Jurnal Media Informatika Budidarma*, *5*(3), 1187–1194. https://doi.org/10.30865/mib.v5i3.3113.

Hutagalung, J., Ginantra, N. L. W. S. R., Bhawika, G. W., Parwita, W. G. S., Wanto, A., & Panjaitan, P. D. (2021). COVID-19 Cases and Deaths in Southeast Asia Clustering using K-Means Algorithm. *Journal of Physics: Conference Series*, *1783*(1). https://doi.org/10.1088/1742-6596/1783/1/012027

Kusuma, A. S., & Aryati, K. S. (2019). Sistem Informasi Akademik Serta Penentuan Kelas Unggulan Dengan Metode Clusttering Dengan Algoritama K-Means Di Smp Negeri 3 Ubud. *Jurnal Sistem Informasi Dan Komputer Terapan Indonesia (JSIKTI)*, *1*(3), 143–152. https://doi.org/10.33173/jsikti.29.

Monalisa, S., & Kurnia, F. (2019). Analysis of DBSCAN and K-means algorithm for evaluating outlier on RFM model of customer behaviour. *Telkomnika (Telecommunication Computing Electronics and Control)*, *17*(1), 110–117. https://doi.org/10.12928/TELKOMNIKA.v17i1.9394.

Nasyuha, A. H., Jama, J., Abdullah, R., Syahra, Y., Azhar, Z., Hutagalung, J., & Hasugian, B. S. (2021). Frequent pattern growth algorithm for maximizing display items. *Telkomnika (Telecommunication Computing Electronics and Control)*, *19*(2), 390–396. https://doi.org/10.12928/TELKOMNIKA.v19i2.16192.

Primanda, R. P., Alwi, A., & Mustikasari, D. (2021). Data Mining Seleksi Siswa Berprestasi Untuk Menentukan Kelas Unggulan Menggunakan Metode K-Means Clustering (Studi Kasus di MTS Darul Fikri). *Komputek*, *5*(1), 88. https://doi.org/10.24269/jkt.v5i1.686.

Puspitasari, N., Widians, J. A., & Setiawan, N. B. (2020). Customer segmentation using bisecting k-means algorithm based on recency, frequency, and monetary (RFM) model. *Jurnal Teknologi Dan Sistem Komputer*, *8*(2), 78–83. https://doi.org/10.14710/jtsiskom.8.2.2020.78-83.

Riana, D., Rahayu, S., Hadianti, S., Frieyadie, F., Hasan, M., Karimah, I. N., & Pratama, R. (2022). Identifikasi Citra Pap Smear RepoMedUNM dengan Menggunakan K-Means Clustering dan GLCM. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *6*(1), 1–8. https://doi.org/10.29207/resti.v6i1.3495.

*Corresponding author

Salbinda, V., & Handayani, R. I. (2022). CLUSTERING OF CLOTHING SALES DATA USING K-MEANS METHOD. *Sinkron : Jurnal Dan Penelitian Teknik Informatika*, *7*(2), 314–320. https://doi.org/https://doi.org/10.33395/sinkron.v7i2.11244 e-ISSN.

Saputra, E. A., & Nataliani, Y. (2021). Analisis Pengelompokan Data Nilai Siswa untuk Menentukan Siswa Berprestasi Menggunakan Metode Clustering K-Means. *Journal of Information Systems and Informatics*, *3*(3), 424–439. https://doi.org/10.51519/journalisi.v3i3.164.

Sari, D. P., Rosadi, D., Effendie, A. R., & Danardono. (2019). K-means and bayesian networks to determine building damage levels. *Telkomnika (Telecommunication Computing Electronics and Control)*, *17*(2), 719–727. https://doi.org/10.12928/TELKOMNIKA.V17I2.11756.

Satria, C., & Anggrawan, A. (2021). Aplikasi K-Means berbasis Web untuk Klasifikasi Kelas Unggulan. *MATRIK : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, *21*(1), 111–124. https://doi.org/10.30812/matrik.v21i1.1473.

Sembiring Brahmana, R. W., Mohammed, F. A., & Chairuang, K. (2020). Customer Segmentation Based on RFM Model Using K-Means, K-Medoids, and DBSCAN Methods. *Lontar Komputer : Jurnal Ilmiah Teknologi Informasi*, *11*(1), 32. https://doi.org/10.24843/lkjiti.2020.v11.i01.p04.

Seta, P. T., & Hartomo, K. D. (2020). Mapping Land Suitability for Sugar Cane Production Using K-means Algorithm with Leaflets Library to Support Food Sovereignty in Central Java. *Khazanah Informatika: Jurnal Ilmu Komputer Dan Informatika*, *6*(1), 15–25. https://doi.org/10.23917/khif.v6i1.9027.

Sugriyono, S., & Siregar, M. U. (2020). Preprocessing kNN algorithm classification using K-means and distance matrix with students' academic performance dataset. *Jurnal Teknologi Dan Sistem Komputer*, *8*(4), 311–316. https://doi.org/10.14710/jtsiskom.2020.13874.

Sulistiyawati, A., & Supriyanto, E. (2020). Implementasi Algoritma K-means Clustring dalam Penetuan Siswa Kelas Unggulan. *Jurnal Tekno Kompak*, *15*(2), 25–36.

Tempola, F., Muhammad, M., & Mubarak, A. (2020). Penggunaan Internet Dikalangan Siswa SD di Kota Ternate: Suatu Survey, Penerapan Algoritma Clustering dan Validasi DBI. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *7*(6), 1153. https://doi.org/10.25126/jtiik.2020722370.

Triayudi, A., & Fitri, I. (2021). Comparison Of The Feature Selection Algorithm In Educational Data Mining. *Telkomnika (Telecommunication Computing Electronics and Control)*, *19*(6), 1865–1871. https://doi.org/10.12928/TELKOMNIKA.v19i6.21594.

Uly Aldini, & Wara Pramesti. (2020). Pengelompokan Provinsi Di Indonesia Berdasarkan Indikator Mutu Pendidikan Sekolah Menengah Pertama Tahun 2016 – 2018 Menggunakan Model Based Clustering. *J Statistika: Jurnal Ilmiah Teori Dan Aplikasi Statistika*, *13*(2), 25–38. https://doi.org/10.36456/jstat.vol13.no2.a2905.

Wibowo, A., Moh Makruf, Inge Virdyna, & Farah Chikita Venna. (2021). Penentuan Klaster Koridor TransJakarta dengan Metode Majority Voting pada Algoritma Data Mining. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *5*(3), 565–575. https://doi.org/10.29207/resti.v5i3.3041.

*Corresponding author