Jurnal Mantik, 7 (2) (2023), ISSN 2685-4236 (Online)

Published by:Institute of Computer Science (IOCS)

Jurnal Mantik

Journal homepage: www.iocscience.org/ejournal/index.php/mantik

Classification of books at SMP YPK Pematang Siantar using the k-means clustering method

Nani Hidayati[1], Kasini[2]
[1]Teknik Informatika, STIKOM Tunas Bangsa Pematangsiantar, Indonesia
[2]Teknik Infromatika, Universitas Pahlawan Tuanku Tambusai, Indonesia

A R T I C L E I N F O

Article history:

Accepted Jun 14, 2023
Revised Jul 27, 2023
Accepted Jul 26, 2023

Keywords:

Clustering;
K-Means;
SMP YPK.

ABSTRACT

The school library is a very important facility in supporting the process of improving the quality of education to produce quality young people. The YPK Pematang Siantar Private Middle School Library has ± 200 book titles in several categories, so that these books can be used optimally there must be a system that regulates the number of book stocks, the number of book loans each month, so that it can be seen which student's reading interest is the most popular in each category . YPK Pematang Siantar Middle School has not implemented an optimal computerized system or everything is still manual. By applying grouping of students' reading interest using the clustering method at SMP YPK Pematang Siantar, it is hoped that the process in the library will be more effective, fast, and precise. Clustering is the most suitable method for optimizing library services. The purpose of this research is to classify which category of books YPK SMP students are most interested in. After calculating the 20 book categories for 3 months, the final result is C1 or the most popular, namely the 3 book categories most in demand. most interested (C2) with 8 book categories, and finally C3 which is less desirable there are 9 book categories. By creating clusters of books which are the most desirable and not desirable, it can improve library services and students' interest in reading and also prevent accumulation of books that are not of interest every year.

Corresponding Author:

Nani Hidayati
Teknik Informatika,
STIKOM Tunas Bangsa Pematangsiantar, Indonesia,
Jl Sudirman Proklamasi Siantar Barat Kota Pematangsiantar 2114 Sumatra Utara.
Email: nanihidayati3@gmail.com

1. INTRODUCTION

The use of information systems in library processing is already familiar, in the activity of borrowing books or reading materials in the library it is necessaryan orderly, structured and systematic record management system (Tendra & Hafsah, 2021) (Menggunakan et al., 2023). At the YPK Private Middle School Library, a system is needed to manage book collections so that you can see which books are most in demand, in demand and not in demand so that there is no accumulation of books (Ani et al., 2021).

Good grouping is good grouping has high homogeneity. The library model studied, using the K-Means Clustering method. With the K-Means algorithm, it is expected to be

able to find good clusterization, so as to produce information in the form of reading interest values (Fitriani et al., 2020).

In this study the authors used the K-Means Clustering method. The data mining algorithm was chosen because it has been tested and is often used by researchers for grouping or clustering (Damanik et al., 2021) and data mining including methods that are often needed in data processing large scale (Saifullah & Hidayati, 2020). Data mining is a method for finding certain patterns from large data sets (Prastiwi et al., 2022). K-Means is one of the clustering algorithms, Clustering is a method widely used data grouping as one of the data mining methods (Suhartini & Yuliani, 2021). Clustering is a technique for one of the data mining functionalities (Hidayati & Rahmah, 2022) . The clustering algorithm is an algorithm for grouping a number of data into certain data groups (clusters) (Fakhriza & Umam, 2021). Therefore, this clustering method useful for finding groups that are unknown in the data (Ramadhan & Voutama, 2022). clustering is a method data grouping used for identify the cluster created with group smaller items based on similarities (Noor et al., 2021). Clustering is a suitable method for library services because it can effectively and efficiently cluster which books are most in demand according to their categories (Fakhri et al., 2021)(Kurniawan et al., 2020). K-means is a cluster analysis algorithm non-hierarchical. Cluster analysis is a tool for grouping data based on variables or characteristics (Lia Hananto et al., 2021). K-Means Is an algorithm that clusters through a continuous iteration process until it meets the final conditions, the iteration process stops and the clustering results are output (Wang et al., 2019).

The K-Means method has been widely used in various fields such as health (Sugara et al., 2020), education(Azmi et al., 2022), economics (Arifiansyah Ayub, 2021) and other fields.

In this study the authors used YPK Pematang Siantar Middle School library data, by grouping the book categories into 3 clusters, namely the most popular (C1), the most popular (C2) and the least popular (C3). The purpose of this research is to classify which book categories are the most popular, desirable and less desirable in order to obtain data that can help the YPK Pematang Siantar Middle School library in providing book categories so that the library can provide stock of books according to the needs of the most interested so there is no stock buildup unused book.

2. RESEARCH METHOD

Data mining is a way to automate the process of finding patterns from large-scale data sets (. et al., 2020) . The K-Means algorithm performs point-based clustering (centroid) by determining three parameters, namely the number of clusters and system distance (Tri et al., 2022), which is adjusted to the research object, both manually and with the help of rapid miner software, along with the steps (Figure 1).

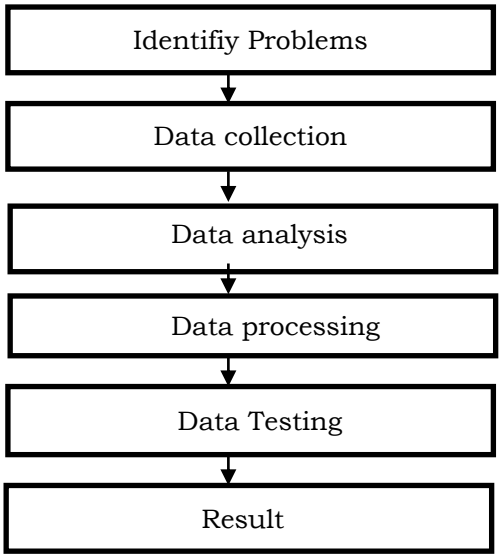

Figure 1. Research Stages

a. Identifiy Problems

Identify the problems in the YKP Middle School library, namely to find out which books are the most popular, desirable and not desirable

b. Data collection

Collecting data on all books and all visitors for 3 months in the YPK Middle School Library. The selected data is then entered into Excel and saved in .xls format and also taking literature studies from various sources, namely articles, scientific journals on the K-Means Clustering method, as well as other supporting reading materials (Aswan et al., 2021).

c. Data analysis

Analyzing the problem is analyzing library visitor data and library book stock data for 3 months

d. Data processing

The data processing technique uses the K-means technique which is a grouping technique that works based on Partitioned Clustering (Nugroho et al., 2022). Namely by processing the data provided by the YPK Middle School Library, then the data is processed using the K-means clustering method which is calculated manually to classify books that are of great interest, less desirable and not desirable.

e. Data Testing

The last stage is testing the results of data processing using RapidMiner Software to find out the equations for calculating the manual method and using Rapid Miner software (Hidayati, 2022).

f. Result

At the results stage, the author will find out the results of calculating the K-Means algorithm using excel and the RapidMiner application. From these results, conclusions and suggestions are obtained in classifying student characters according to the attributes studied and these results will be given to the school for further policies (Mawaddah et al., 2022)

3. RESULTS AND DISCUSSIONS

Clustering is a data mining method that functions to group some data to produce one data that is similar to one new data (Amanda & Veronica Sitorus, 2021).

The data used in grouping library books is data on books in the YPK Junior High School library in Pematang Siantar, North Sumatra. Book data consists of several types which are then categorized based on the type of book. The data will be processed to create knowledge of book inventory management strategies. In the application of data mining, grouping of book data will be grouped into 3 clusters, namely the most desirable, desirable and less desirable, so the variables used as a reference for testing k means clustering must be applied. These variables are as follows:

Table 1. Variable data

NO	Variable name	Information
1	Book Category	Books name by category
2	Book Stock	Stock of book before borrowing
3	Book borrower	Number of books borrowed

Table 1 contains the variables used in this study. Based on the table there are 3 types of variables used, the primary data obtained from the school library:

Table 2 . Primary List of Book Categories

NO	Book Category	Book Stock	Borrower		
			August	September	October
1	Religion	20	3	4	2
2	B. England	25	1	0	3
3	Indonesian	3	0	1	2
4	Biography	5	2	2	1
5	Informatics	15	6	4	5
6	IPA	15	5	3	5
7	IPS	17	2	5	5
8	Dictionary	10	3	2	5
9	Skills	9	6	1	2
10	Comic	30	9	10	11
11	Literacy	15	2	5	3
12	Magazine	10	0	0	2
13	Mathematics	35	7	9	3
14	Novel	25	4	7	10
15	Sports physical Education and health	15	2	6	5
16	Pancasila and civic education	10	2	3	3
17	physical education	5	1	1	1
18	crafts	10	2	2	4
19	Art and culture	15	2	8	5

Based on table 2, primary data is data obtained from the school library which contains data on borrowing books for 3 months, namely August, September and October. This primary data is raw data that cannot be processed.

So that the primary data can be processed using the k-means clustering method, the data must be arranged according to the variables. The following is the data that will be used in this research, namely the latest borrow data, namely data from August to October 2022 which is then processed into a simpler form of data. The following data has been simplified:

Table 3. Data Table by Variable

NO	Book Category	Book Stock	Book borrower
1	Religion	20	9
2	B. England	25	4
3	Indonesian	3	3
4	Biography	5	5
5	Informatics	15	15
6	IPA	15	13
7	IPS	17	12
8	Dictionary	10	10
9	Skills	9	9
10	Comic	30	30
11	Literacy	15	10
12	Magazine	10	2
13	Mathematics	35	19
14	Novel	25	21
15	Sports physical Education and health	15	13
16	Pancasila and civic education	10	8
17	physical education	5	3
18	crafts	10	8
19	Art and culture	15	15
20	health tips	2	4

Based on table 3 which is primary data that has been processed into simpler data so that it can be processed based on the type of variable using k means clustering with 3 clusters. The application of the k-means clustering algorithm in grouping data on book lending can be described as follows:

a. Iteration 1 determines the initial centroid value In determining the initial centroid in iteration 1, it is determined randomly from existing data. In this study, the data taken were the 14th data, the 1st data, and the 12th data.

Table 4. Determination of initial centroid values

1. Determination Of Initial Centroid Values			
C1	25	21	Data 14
C2	20	9	Data 1
C3	10	2	Data 12

In table 4 the researcher selects 3 data based on 3 clusters of all the variables used, the selected data starts from the largest to the smallest data to facilitate calculation, then is calculated using the K means algorithm formula.

The following Calculation of the distance from data 1 to the cluster center point is as follows:

$D(1,1) = \sqrt{(20-25)^2 + (9-21)^2} = 169$

$D(1,2) = \sqrt{(20-20)^2 + (9-9)^2} = 0$

$D(1,3) = \sqrt{(20-10)^2 + (9-2)^2} = 149$

And so on, calculating the distance of the second data to the cluster center data, then determining the comparison between the 3 clusters with the smallest value being an option, if the smallest value is found then it can be grouped into that cluster. Following are the cluster results in iteration 1:

Table 5 Iteration calculation results 1

No	Book Category	Book Stock	book borrower	Cluster 1	Cluster 2	Cluster 3	shortest distance	Clustering
1	Religion	20	9	169	0	149	0	2
2	B. England	25	4	289	50	229	50	2
3	Indonesian	3	3	808	325	50	50	3

4	Biography	5	5	656	241	34	34	3
5	Informatics	15	15	136	61	194	61	2
6	IPA	15	13	164	41	146	41	2
7	IPS	17	12	145	18	149	18	2
8	Dictionary	10	10	346	101	64	64	3
9	Skills	9	9	400	121	50	50	3
10	Comic	30	30	106	541	1184	106	1
11	Literacy	15	10	221	26	89	26	2
12	Magazine	10	2	586	149	0	0	3
13	Mathematics	35	19	104	325	914	104	1
14	Novel	25	21	0	169	586	0	1
15	Sports physical Education and health	15	13	164	41	146	41	2
16	Pancasila and civic education	10	8	394	101	36	36	3
17	physical education	5	3	724	261	26	26	3
18	crafts	10	8	394	101	36	36	3
19	Art and culture	15	15	136	61	194	61	2
20	health tips	2	4	818	349	68	68	3

Based on table 5 above, the shortest distance is obtained, then the shortest distance data is grouped into clustering and obtained data in cluster 1 there are 3 data, cluster 2 has 8 data and cluster 3 has 9 data, but iteration 1 data cannot be used as the final result, therefore it is necessary to do the next step in the k means algorithm.

b. Then the next step determines the new centroid value, this value is determined by the data that enters the cluster, the following data enters the cluster based on the table above (data 1-20) and then is calculated using the following formula:

Table 6. Iteration Cluster Result 1

C1	3 data (data 10,13, and 14)
C2	8 data (data 1,2, 5,6,7,11,15 and 19)
C3	9 data (data 3,4,8,9,,12,16,17,18,and 20)

In table 6 it can be seen the calculated data from iteration 1 which is then calculated using the following formula to determine the next iteration, namely iteration 2.

$$Ck = \frac{The\ sum\ of\ the\ value\ entered\ into\ the\ cluster}{The\ amount\ of\ data\ entered} \quad (1)$$

So that the new centroid value is obtained in iteration 2 as follows:

Table 7 The new centroid value iteration 2

Cluster 1	30	23
cluster 2	17	12
Cluster 3	7	6

Next , table 7 displays the new centroid value in which iteration 2 to find the next centroid value, repeat step 1 above. After the new centroid value is found, repeat the distance calculation steps, namely in the previous step, until you enter data into the cluster.

c. If the above steps are repeated with the same steps until the data in a cluster is exactly the same as the previous data with the data in the next step, or in other words the data does not change its position in the cluster, the calculation of the centroid value

can be stopped or terminated. In calculating the research data, the cluster position does not change with the 2nd iteration cluster position, so that it can be seen in the following table with the 2nd iteration calculation:

Table 8 Iteration cluster result 2

No	Book Category	Book Category	Book Stock	book borrower	Cluster 1	Cluster 2	Cluster 3	shortest distance
1	Religion	20	9	296	18	178	18	2
2	B. England	25	4	386	128	328	128	2
3	Indonesian	3	3	1129	277	25	25	3
4	Biography	5	5	949	193	5	5	3
5	Informatics	15	15	289	13	145	13	2
6	IPA	15	13	325	5	113	5	2
7	IPS	17	12	290	0	136	0	2
8	Dictionary	10	10	569	53	25	25	3
9	Skills	9	9	637	73	13	13	3
10	Comic	30	30	49	493	1105	49	1
11	Literacy	15	10	394	8	80	8	2
12	Magazine	10	2	841	149	25	25	3
13	Mathematics	35	19	41	373	953	41	1
14	Novel	25	21	29	145	549	29	1
15	Sports physical Education and health	15	13	325	5	113	5	2
16	Pancasila and civic education	10	8	625	65	13	13	3
17	physical education	5	3	1025	225	13	13	3
18	crafts	10	8	625	65	13	13	3
19	Art and culture	15	15	289	13	145	13	2
20	health tips	2	4	1145	289	29	29	3

From the grouping results of table 8 above, it can be seen that there is no longer a change in members for each cluster. Therefore, the iteration process is sufficient until iteration 2, and is stopped here.

The application of the K Means calculation method stops at iteration 2 so that it can be concluded that clustering of borrowing books for 3 months results in Cluster 1 with 3 categories of books that are most in demand, namely Comics, Mathematics and Novels, while in cluster 2 with 8 categories of books that are of interest, namely religious books, English, Literacy, Science, IPS, Physical *Education and Cultural Arts. And finally cluster 3 with 9 book categories.*

In this study the researchers implemented and tested library data using rapid miner 7.5 to ensure the accuracy of manual data processing with the results of data processing using a software.

The results based on the implementation of rapid miner 7.5 with book data in the SMP YPK Pematang Siantar library are as follows:

ExampleSet (20 examples, 2 special attributes, 2 regular attributes)

Row No.	id	cluster	Book Stock	book borro...
3	3	cluster_0	3	3
4	4	cluster_0	5	5
8	8	cluster_0	10	10
9	9	cluster_0	9	9
12	12	cluster_0	10	2
16	16	cluster_0	10	8
17	17	cluster_0	5	3
18	18	cluster_0	10	8
20	20	cluster_0	2	4
1	1	cluster_1	20	9
2	2	cluster_1	25	4
5	5	cluster_1	15	15
6	6	cluster_1	15	13
7	7	cluster_1	17	12
11	11	cluster_1	15	5
15	15	cluster_1	15	13
19	19	cluster_1	15	15
10	10	cluster_2	30	30
13	13	cluster_2	35	19
14	14	cluster_2	25	21

Figure 2 Result of K-Means Clustering Rapid Miner

Based on the tests implemented using the rapid miner, the processed data is primary data which has been simplified and entered into the application and then processed according to k means with 3 clusters so that the grouping of books of interest can be seen in Figure 2 above. In the picture above it can be seen that the grouping of books that are of interest and those that are of less interest is the same as the grouping in the previous manual calculation.

Cluster Model

```
Cluster 0: 9 items
Cluster 1: 8 items
Cluster 2: 3 items
Total number of items: 20
```

Figure 3 Result Model Clustering Rapid Miner

In figure 3 the Cluster Model (Clustering) you can see several views of the cluster results, namely the Text View which is a display of grouping results based on clusters and the number of book data (Azhami & Fauziah, 2020). The final results of this test can be seen in the grouping model using a rapid miner where the processed data has similar results to data using manual calculations. Therefore, overall manual calculations and using applications show high accuracy and effectiveness in grouping some data.

4. CONCLUSION

Based on the results of research using the k-means clustering method in the YPK Pematangsiantar Private Middle School library of 20 book categories for 3 months, the final result is C1 or the most popular, namely 3 book categories, the most popular (C2)

with 8 book categories, and finally C3 which less desirable there are 9 categories of books.

The YPK Pematang Siantar Private Middle School Library still has limited resources in managing library book data that has not implemented an optimal computerized system, this can be seen from the use of manual calculations and compilation of statistical data on book lending. With the implementation of classifying students' reading interest using the clustering method in the SMP YPK Pematang Siantar library, it is hoped that the management and communication process in the library will become more effective, fast and precise. The results of clustering books that are in great demand by students can be used as a reference to add to the collection of books in the library.

In the process of this research there are limitations that might affect the results of the research including limited time, manpower and research abilities and this research only covers groupings of students' reading interests so that further research needs to be developed with groupings of reading interests which are developed with a larger number of clusters and an additional number of attributes as an assessment parameter using different methods as a comparison material so that the results obtained are more optimal and accurate.

The results of this study are expected to be used as information, contributions of thoughts and considerations for schools to make decisions in an effort to increase students' reading interest in the YPK Pematang Siantar Private Middle School Library.

REFERENCES

F., Kesuma, F. T., & Tamba, S. P. (2020). Penerapan Data Mining Untuk Menentukan Penjualan Sparepart Toyota Dengan Metode K-Means Clustering. *Jurnal Sistem Informasi Dan Ilmu Komputer Prima(JUSIKOM PRIMA)*, *2*(2), 67–72. https://doi.org/10.34012/jusikom.v2i2.376

Amanda, & Veronica Sitorus, M. (2021). Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Konsumsi Produk Kosmetik milik PT Cedefindo. *Jurnal Ilmiah MIKA AMIK Al Muslim*, *V*(2), 63–68.

Ani, H., Nofriansyah, D., & Mariami, I. (2021). Implementasi Data Mining Untuk Pengelempokan Buku Di Perpustakaan Yayasan Nurul Islam Indonesia Baru Dengan Metode K-Means Clustering. *Jurnal Cyber Tech*, *1*(1), 1–12. https://ojs.trigunadharma.ac.id/index.php/jct/article/view/315

Arifiansyah Ayub, M. (2021). Analisis Topik Ekonomi Dengan Algoritma K-Means Pada Media Online Era Pandemi Covid-19 Di Sulawesi Tenggara. *JIKO (Jurnal Informatika Dan Komputer)*, *4*(2), 133–138. https://doi.org/10.33387/jiko.v4i2.3235

Aswan, Y., Defit, S., & Nurcahyo, G. W. (2021). Algoritma K-Means Clustering dalam Mengklasifikasi Data Daerah Rawan Tindak Kriminalitas (Polres Kepulauan Mentawai). *Jurnal Sistim Informasi Dan Teknologi*, *3*, 245–250. https://doi.org/10.37034/jsisfotek.v3i4.73

Azhami, I., & Fauziah, R. (2020). Penerapan Rapidminer Pada Data Mining Klastering (Kasus: Distribusi Persentase Rumah Tangga Menurut Kabupaten/Kota Dan Bahan Bakar Untuk Memasak). *KESATRIA: Jurnal Penerapan Sistem Informasi (Komputer & Manajemen)*, *1*(2), 52–58. https://doi.org/10.30645/kesatria.v1i2.20

Azmi, N., Helmiah, F., & Sudarmin, S. (2022). Implementasi Metode K-Means Sebagai Upaya Penentuan Lokasi Promosi Penerimaan Siswa Baru. *Building of Informatics, Technology and Science (BITS)*, *3*(4), 649–660. https://doi.org/10.47065/bits.v3i4.1456

Damanik, Y. F. S. Y., Sumarno, S., Gunawan, I., Hartama, D., & Kirana, I. O. (2021). Penerapan Data Mining Untuk Pengelompokan Penyebaran Covid-19 Di Sumatera Utara Menggunakan Algoritma K-Means. *Jurnal Ilmu Komputer Dan Informatika*, *1*(2), 109–132. https://doi.org/10.54082/jiki.13

Fakhri, D. A., Defit, S., & Sumijan. (2021). Optimalisasi Pelayanan Perpustakaan terhadap Minat Baca Menggunakan Metode K-Means Clustering. *Jurnal Informasi Dan Teknologi*, *3*, 160–166. https://doi.org/10.37034/jidt.v3i3.137

Fakhriza, M. H., & Umam, K. (2021). Analisis Produk Terlaris Menggunakan Metode K-Means

Clustering Pada "Pt.Sukanda Djaya. *JIKA (Jurnal Informatika)*, *5*(1), 8. https://doi.org/10.31000/jika.v5i1.3236

Fitriani, S., Kom, S., Kom, M., Kom, S., & Kom, M. (2020). *Implementasi Data Mining Dalam Pengelompokkan Minat Baca Pengunjung Pada Perpustakaan STMIK Triguna Dharma MedanMenggunakan Metode K-Means*.

Hidayati, N. (2022). *Jurnal Mantik Implementation of data mining to determine stock inventory at kenza grocery stores using the k-means clustering method*. *6*(3).

Hidayati, N., & Rahmah, S. A. (2022). *Clasterization Of Zeeida Product Sales Using K - Means Method In Medan Distributors*. *6*(36), 1685–1692.

Kurniawan, R., Defit, S., & Sumijan, S. (2020). Prediksi Tingkat Kerugian Peternak Akibat Penyakit pada Sapi Menggunakan Algoritma K-Means Clustering. *Jurnal Informasi Dan Teknologi*, *3*. https://doi.org/10.37034/jidt.v3i1.87

Lia Hananto, A., Assiroj, P., Priyatna, B., Nurhayati, Fauzi, A., Yuniar Rahman, A., & Shofiah Hilabi, S. (2021). Analysis of Drug Data Mining with Clustering Technique Using K-Means Algorithm. *Journal of Physics: Conference Series*, *1908*(1). https://doi.org/10.1088/1742-6596/1908/1/012024

Mawaddah, W. J., Gunawan, I., & Sari, I. P. (2022). Implementation of Data Mining Algorithm for Clustering of Palm Oil Harvested Data. *JOMLAI: Journal of Machine Learning and Artificial Intelligence*, *1*(1), 43–54. https://doi.org/10.55123/jomlai.v1i1.163

Menggunakan, P., Berbasis, R., & Via, I. O. T. (2023). *Inventarisasi peminjaman buku secara otomatis pada perpustakaan menggunakan rfid berbasis iot via telegram*. *8*(1), 1–16.

Noor, H., Dharmawati, A., & Qur'ana, T. W. (2021). Penerapan Algoritma K-Means Clustering Analysis Pada Kasus Penderita Hiv/Aids (Studi Kasus Kabupaten Banjar). *Technologia: Jurnal Ilmiah*, *12*(2), 72. https://doi.org/10.31602/tji.v12i2.4573

Nugroho, M. R., Hendrawan, I. E., & Purwantoro, P. P. (2022). Penerapan Algoritma K-Means Untuk Klasterisasi Data Obat Pada Rumah Sakit ASRI. *Nuansa Informatika*, *16*(1), 125–133. https://doi.org/10.25134/nuansa.v16i1.5294

Prastiwi, H., Pricilia, J., & Raswir, E. (2022). Implementasi Data Mining Untuk Menentuksn Persediaan Stok Barang Di Mini Market Menggunakan Metode K-Means Clustering Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM). *Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM)*, *1*(April), 141–148.

Ramadhan, V., & Voutama, A. (2022). Clustering Menggunakan Algoritma K-Means Pada Penyakit ISPA di Puskesmas Kabupaten Karawang. *Jurnal Pendidikan Dan Konseling*, *4*, 462–473.

Saifullah, S., & Hidayati, N. (2020). Pengelompokan Persentase Buta Huruf Umur 15-44 Menurut Provinsi Menggunakan Algoritma K-Means. *Klik - Kumpulan Jurnal Ilmu Komputer*, *7*(3), 230. https://doi.org/10.20527/klik.v7i3.329

Sugara, H., Hanafiah, M. A., Sekretari, P., Indonesia, P. B., Akuntasi, P. K., & Indonesia, P. B. (2020). *JurnalMantik*. *4*(3), 2225–2231.

Suhartini, S., & Yuliani, R. (2021). Penerapan Data Mining untuk Mengcluster Data Penduduk Miskin Menggunakan Algoritma K-Means di Dusun Bagik Endep Sukamulia Timur. *Infotek : Jurnal Informatika Dan Teknologi*, *4*(1), 39–50. https://doi.org/10.29408/jit.v4i1.2986

Tendra, G., & Hafsah, H. (2021). *Perancangan Sistem Informasi Pendaftaran Anggota Pada Perpustakaan Soeman HS Pekanbaru Berbasis Web*. *10*, 152–161.

Tri, M., Rahmayani, I., & Hidayati, N. (2022). *Implemention K-Means Algorithm Determine the Recovery Rate Of Covid- 19 Patients In Indonesia*. *6*(36), 127–135.

Wang, Y., Li, D., & Wang, Y. (2019). Realization of remote sensing image segmentation based on K-means clustering. *IOP Conference Series: Materials Science and Engineering*, *490*(7), 1–8. https://doi.org/10.1088/1757-899X/490/7/072008