International Journal of Basic and Applied Science 10 (2) (2021) 60-68

Published by: IOCSCIENCE

International Journal of Basic and Applied Science

Journal homepage: www.ijobas.pelnus.ac.id

A Oil Palm Harvest Grouping Using K-Medoids Algorithm

Dessy Dwi Angraini[1], M. Safii[1], Fitri Anggraini[2]

[1]Informatics Engineering Study Program, STIKOM Tunas Bangsa, Pematangsiantar

[2]Computerized Accounting Study Program, Amik Tunas Bangsa, Indonesia

Article Info

Article history:

Received Jun 20, 2021
Revised Jul 22, 2021
Accepted Sep 17, 2021

Keywords:

Oil Palm
Janjang
Data Mining
K-Medoids
RapidMiner

ABSTRACT

Oil palm (Elaies Guinnnsiss Jacq) is one of the important industrial crops producing cooking oil, industrial oil, and fuel. Indonesia is the largest palm oil producer in the world. The rest of the processing of oil palm fruit is called janjang. Janjang also serves to be used as compost. The data that is processed in this research is the harvest data at PT. Surya Intisariraya Mandau. Data mining is the process of looking for patterns or information in selected data using certain techniques or methods. The processing steps are grouped using the K-Medoids method and then the data will be processed using RapidMiner tools. Where this grouping is done to minimize the amount of similarity of data and appropriate so that it becomes more valid data. This study aims to simplify the grouping of harvest data based on high, medium and low clusters.

This is an open access article under the CC BY-NC license.

Corresponding Author:

M.Safii,
Department of Information System,
Department of Computerized Accounting,
AMIK dan STIKOM Tunas Bangsa Pematangsiantar,
Jl. Jend. Sudirman Blok A-B No.1-3 Pematangsiantar, North Sumatera, Indonesia.
Email: m.safii@amiktunasbangsa.ac.id

1. INTRODUCTION

Oil palm *(Elaies Guinnnsiss Jacq)* is one of the important industrial crops producing cooking oil, industrial oil, and fuel [1]. Indonesia is the world's largest producer of palm oil [2][3][4] . The most important part to be processed from the oil palm plant is the fruit. The rest of the processing of oil palm fruit (janjang) is very potential which can be used as compost [5][6][7]. As one of the largest agricultural export commodities in Indonesia, palm oil has an important role as a source of foreign exchange and large tax earners.

Several previous studies that were used as references in this study conducted by [8] explained that forecasting oil palm yields consisted of several factors, namely monthly oil palm yields, land area, plant age, and principal amount. Palm oil. Subsequent research by [9][10]The process of identifying the maturity of oil palm fruit using the *K-Means Clustering method is* able to recognize oil palm fruit image objects based on the level of maturity, namely raw, moderately ripe, and ripe and the results of the identification of oil palm fruit maturity with K-means Clustering algorithm obtained an accuracy rate of 79.16% for test data and a level 2 accuracy for training data of 50%, so that the total accuracy of both is 64.58%. Harvesting is one of the important activities in the management of mature oil palm plantations [11]. In addition to planting material (seeds) and plant maintenance, harvesting is also an important factor in achieving productivity. In making harvest data at PT. Intisariraya Solar Saber, Riau where the processing of data not harvest bunch of data grouped by *cluster* of high, medium, and low on-AFD3 AFD1 per month. Therefore, the author uses the *k-medoids* method as a method for classifying harvest data. The purpose of this grouping is to facilitate PT. Surya Intisariraya Mandau, Riau in knowing the number of harvest bunch in AFD1-AFD3 which consists of several blocks that get results bunch production to *cluster* high, *cluster* medium, and *the cluster* is low.

The object of this research uses data from the 2020 harvest at PT. Surya Intisariraya Mandau, Riau. The concept that will be used in this research is to use the concept of data mining. Data mining technique is a program for processing large volumes of data and has a high speed and can be used as a source of information and used to draw conclusions related to a research to be processed [12][13][14]. In the process, data mining uses statistical, mathematical, artificial intelligence and *machine learning* techniques that function to extract and identify useful information and related knowledge from various large *databases.*

2. RESEARCH METHOD

The design of this study was first carried out by observing data and analyzing existing problems. After that the data will be processed through the calculation process and follow the steps of the *k-medoids* algorithm calculation process. Furthermore, the results of the grouping will be applied to *Rapidminer* to see the results. The results of grouping the data on the harvest of janjang are the development of science which will later be able to provide solutions to problem solving in PT. Surya Intisariraya Mandau, Riau. Flowchart is a technique used to explain aspects of information systems clearly, precisely and logically.

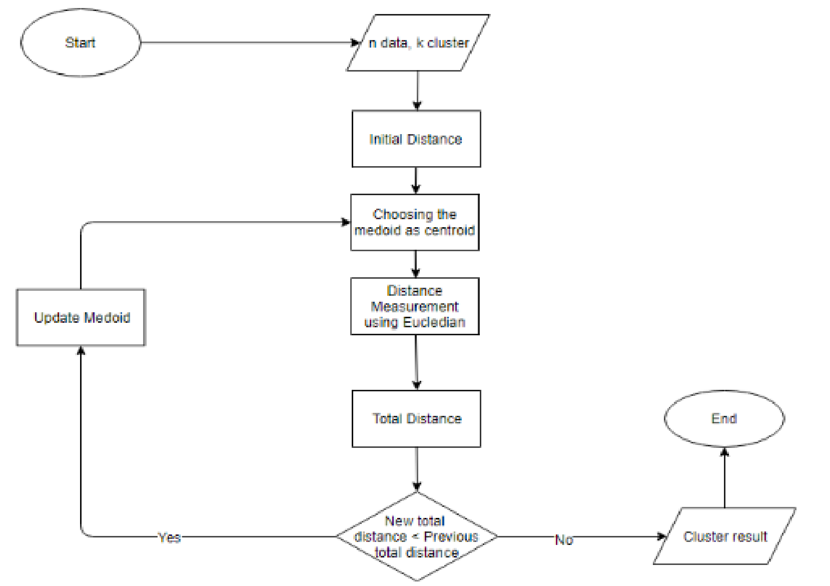


Figure 1. Flowchart K-medoids

The following are some examples of long harvest data samples per month in 2020, where the total number of perennial harvest data is 140 data.

Table 1. Data Sample
(Source: PT. Surya Intisariraya Mandau)

AF	Blok	Large	Harvest Janjang (kg) / Month

D		(HA)	Jan	Feb	Mar	April	May	Jun	Jul	Augst	Sep	Okt	Nov	Des
1	S34	28.36	2,099.00	2,191.00	1,761.00	3,456.00	1,718	3,224	3,694	4,419	3,246.00	3,331.00	2,738.00	1,653
1	S30	29.36	2,434.00	1,864.00	1,635.00	3,028.00	2,578	2,716	3,692	3,692	3,432.00	3,716.00	2,783.00	2,503
1	S31	28.69	2,277.00	1,803.00	2,105.00	2,815.00	2,267	3,280	3,713	3,459	3,778.00	3,120.00	2,617.00	2,824
1	U36	2,723.00	1,864.00	3,659.00	2,393.00	2,459	4,132	3,784	3,496	4,056.00	3,856.00	3,244.00	2,495	2,723.00
2	P32	27.33	2,368.00	1,989.0	2,189.00	2,971.00	3,035.00	2,818	3,050	4,095	3,420.00	3,409.00	2,831	2,955

AFD	Blok	Large (HA)	Harvest Janjang (kg) / Month											
			Jan	Feb	Mar	April	May	Jun	Jul	Augst	Sep	Okt	Nov	Des
2	P33	28.89	1,743.00	2,687.00	2,316.00	2,091.00	2,925.00	2,325	2,666	2,895	4,628.00	3,360.00	2,190	2,871
2	P34	29.11	1,783.00	2,207.00	1,915.00	2,075.00	3,217.00	4,187	2,925	2,697	4,792.00	2,919.00	2,073	1,940
3	B07	2,337.00	2,810.00	2,300.00	2,753.00	2,201.00	2,833	3,595	2,757	3,561.00	3,320.00	3,009.00	4,072	2,337.00
3	B08	2,285.00	2,728.00	2,816.00	2,000.00	2,266.00	2,956	3,465	2,804	3,325.00	3,277.00	3,224.00	3,242	2,285.00

At this stage, the data transformation/ *preprocessing* process is carried out , where data showing numbers < 1000 = 1, > 1000 = 2, > 2000 = 3, > 3000 = 4, > 4000 = 5. The collected data will be processed into 3 *clusters*. with several stages and the calculation process using the *rapidminer version 5.3* application.

Table 2. Sample Data Used

AFD	Blok	Harvest Janjang (kg) / Month											
		Jan	Feb	Mar	Apr	Mei	Jun	Jul	Augst	Sep	Okt	Nov	Des
1	S34	2	2	2	3	2	3	4	4	3	3	3	2
1	S30	2	2	2	3	3	3	4	4	3	4	3	3
1	S31	2	2	2	3	2	3	4	3	4	3	3	3
1	S32	2	3	2	3	2	3	3	4	3	4	3	3
1	S36	2	3	2	2	3	3	4	4	4	3	3	2
1	T35	2	2	2	3	2	3	4	3	3	3	3	3
1	T36	3	2	2	2	5	4	4	4	3	3	2	3
1	U35	2	2	3	3	2	3	4	3	3	2	3	2
1	U36	3	2	4	2	2	4	4	3	4	4	3	2
2	P32	2	2	2	3	3	3	3	4	3	3	3	3
2	P33	2	3	2	2	3	2	3	3	5	3	2	3
2	P34	2	2	2	2	3	4	3	3	5	3	2	2
3	B07	2	3	2	3	2	3	4	3	4	3	3	4

AFD	Blok	Harvest Janjang (kg) / Month											
		Jan	Feb	Mar	Apr	Mei	Jun	Jul	Augst	Sep	Okt	Nov	Des
3	B08	2	3	3	2	2	3	3	3	3	3	3	3
3	B09	2	3	2	2	2	3	4	3	3	3	4	3
3	B10	2	2	2	2	2	2	3	2	2	2	3	2
3	B11	2	2	2	2	2	2	3	2	2	2	2	2
3	B12	1	1	1	1	1	2	2	1	2	2	2	2
3	C08	2	1	2	1	1	1	1	2	2	3	2	2

3. RESULTS AND DISCUSSIONS

To solve the problem in this research, data mining techniques with the k-medoids algorithm are used as follows: [15],[16],[17],[18],[19],[20]:

$$D_{ij} = \sqrt{\sum_{a=1}^{p} (x_{ia} - x_{ja})^2} = \sqrt{(x_i - x_j)'(x_i - x_j)}$$

3.1 Solution with *K-Medoids*

The *K-Medoids* method used in the clustering process of harvesting janjang data obtained from PT. Surya Intisariraya Mandau. The steps in completing manual data mining calculations using the *k-medois clustering method* are as follows:

a. Initialization *of k-medoids cluster* center (number of *clusters*)
   The number of *clusters* used in this study consisted of C2/high, C1/medium, C0/low. The *k-medoids* algorithm begins with the initial determination of the *cluster* center by randomly selecting among the objects in the panen janjang data

Tabel 3. Initial *Cluster* Center Initialization

Blok / Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Augst	Sep	Okt	Nov	Des
Blok N31 (C1)	1	2	1	1	2	2	2	2	3	2	2	1
Blok S34(C2)	2	2	2	3	2	3	4	4	3	3	3	2
Blok N30 (C3)	2	2	2	2	2	2	3	3	2	3	2	1

b. Calculate the value of the closest distance with the *Euclidean Distance* equation . To calculate the distance between the *centroid* point with the point of each object using *Euclidean Distance*. Then the manual calculation to calculate the distance of each object with the initial *medoids* is as follows:

$$D_{U31,\,C1} = \sqrt{\begin{matrix}(1-1)^2+(1-2)^2+(1-1)^2+(1-1)^2\\ +(1-2)^2+(1-2)^2+(1-2)^2+(1-2)^2\\ +(1-3)^2+(1-2)^2+(1-2)^2+(1-1)^2\end{matrix}}$$

$= 3$

$$D_{S30,\,C2} = \sqrt{\begin{matrix}(2-2)^2+(2-2)^2+(2-2)^2+(3-3)^2\\ +(3-2)^2+(3-3)^2+(4-4)^2+(4-4)^2\\ +(3-3)^2+(4-3)^2+(3-3)^2+(3-2)^2\end{matrix}}$$

$= 2$

$$D_{N29,\,C3} = \sqrt{\begin{array}{l}(1-2)^2+(2-2)^2+(1-2)^2+(2-2)^2\\+(2-2)^2+(2-2)^2+(3-3)^2+(3-3)^2\\+(2-2)^2+(2-3)^2+(2-2)^2+(1-1)^2\end{array}}$$
$= 2$

Tabel 4. Medoids Distance Iteration 1

C1	C2	C3	Nearest Distance	*Cluster*
4	0	3	0	2
5	2	3	2	2
5	3	4	3	2
3	2	3	2	2
6	3	4	3	2
3	7	4	3	1
3	2	3	2	2
6	3	5	3	2
4	2	3	2	2
3	6	4	3	1

c. Calculating *Cost* Value, the *cost* value is obtained from the total sum of the *medoid* closest distance values.

d. Repeat the previous steps with the new *medoids* center value. The *k-medoids* algorithm begins with the initial determination of the *cluster* center by randomly selecting among the objects in the long harvest data for each object as a candidate *for* new *medoids*.

Tabel 4. New Medoids Cluster Center Iteration 2

Nama / Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Augst	Sep	Okt	Nov	Des
Blok V33 (C1)	1	1	1	1	1	1	1	2	1	1	1	1
Blok U35 (C2)	2	2	3	3	2	3	4	3	3	2	3	2
Blok N30 (C3)	2	2	2	2	2	2	3	3	2	3	2	1

$$D_{B04,\,C1} = \sqrt{\begin{array}{l}(1-1)^2+(1-1)^2+(1-1)^2+(1-1)^2\\+(1-1)^2+(1-2)^2+(1-1)^2+(1-2)^2\\+(1-1)^2+(2-1)^2+(1-1)^2+(2-1)^2\end{array}}$$
$= 2$

$$D_{T33,\,C2} = \sqrt{\begin{array}{l}(3-2)^2+(2-2)^2+(2-3)^2+(4-3)^2\\+(4-2)^2+(4-3)^2+(5-4)^2+(4-3)^2\\+(4-3)^2+(4-2)^2+(3-3)^2+(4-2)^2\end{array}}$$
$= 4$

$$D_{V34,\,C3} = \sqrt{\begin{array}{l}(1-2)^2+(1-2)^2+(1-2)^2+(2-2)^2\\+(2-2)^2+(2-2)^2+(2-3)^2+(2-3)^2\\+(2-2)^2+(2-3)^2+(2-2)^2+(2-1)^2\end{array}}$$
$= 3$

Tabel 5. New Medoids Distance Iteration 2

C1	C2	C3	Nearest Distance	*Cluster*
6	2	3	2	2
7	2	3	2	2
7	2	3	2	2
7	3	4	3	2

7 3 4 3 2
7 3 4 3 2

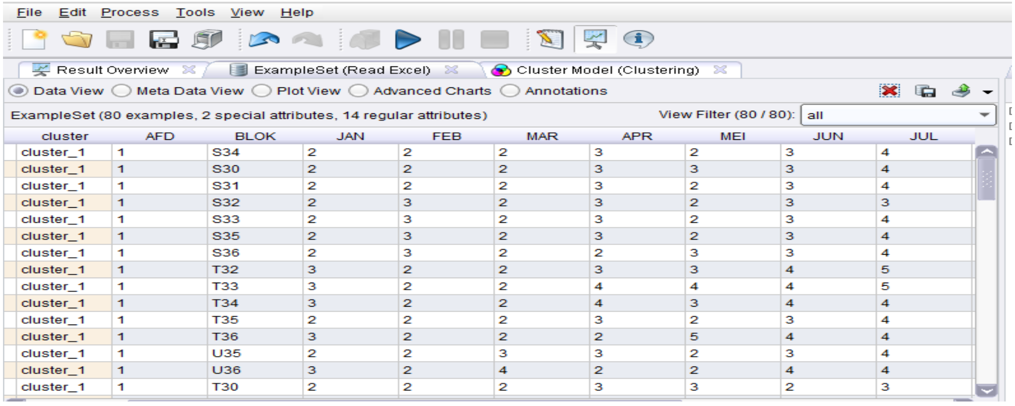
e. Calculate the total deviation (S) by reducing the *cost* value in iteration 2 (new) with iteration 1 (initial). If the value of S < 0 then the processing is continued by using the new *medoids* center value. If the value of S > 0 or the value of the *cost of* the new iteration is greater than the value of the *cost of* the old iteration, the process is stopped so that the value of S is obtained:

S = new cost value â€“ old cost value
= 182 - 177
= 5

Because the New *Cost* Value > Old *Cost* Value, the iteration is stopped and the *cluster* result is in iteration 1.

3.2 Results of Data Processing With RapidMiner

RapidMiner is software for data processing using data mining principles and algorithms [21],[22],[23],[24],[25].



File Edit Process Tools View Help

Result Overview | ExampleSet (Read Excel) | Cluster Model (Clustering)

Data View | Meta Data View | Plot View | Advanced Charts | Annotations

ExampleSet (80 examples, 2 special attributes, 14 regular attributes) View Filter (80 / 80): all

cluster	AFD	BLOK	JAN	FEB	MAR	APR	MEI	JUN	JUL
cluster_1	1	S34	2	2	2	3	2	3	4
cluster_1	1	S30	2	2	2	3	3	3	4
cluster_1	1	S31	2	2	2	3	2	3	4
cluster_1	1	S32	2	3	2	3	2	3	3
cluster_1	1	S33	2	3	2	3	2	3	4
cluster_1	1	S35	2	3	2	3	2	3	4
cluster_1	1	S36	2	3	2	2	3	3	4
cluster_1	1	T32	3	2	2	3	3	4	5
cluster_1	1	T33	3	2	2	4	4	4	5
cluster_1	1	T34	3	2	2	4	3	4	4
cluster_1	1	T35	2	2	2	3	2	3	4
cluster_1	1	T36	3	2	2	2	5	4	4
cluster_1	1	U35	2	2	3	3	2	3	4
cluster_1	1	U36	3	2	4	2	2	4	4
cluster_1	1	T30	2	2	2	3	3	2	3

Figure 2. *Cluster* Data Results

In figure 2. it can be seen that the results *of* data *clusters are* in AFD1-AFD3 and each block has been grouped based on *cluster* 0, *cluster* 1, and *cluster* 2.

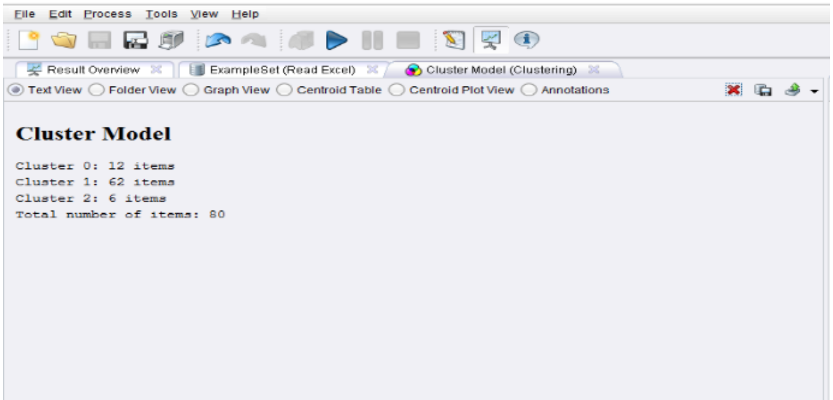


Figure 3. Cluster Model Results

In figure 3. it can be seen the results of *clustering* that have been processed, where *cluster* 0 consists of 12 data, *cluster* 1 consists of 62 data, and *cluster* 2 consists of 6 data.

Attribute	cluster_0	cluster_1	cluster_2
AFD	3	3	3
BLOK	79	77	78
JAN	2	2	2
FEB	1	2	1
MAR	1	2	2
APR	1	2	2
MEI	1	3	2
JUN	2	2	2
JUL	1	3	2
AGST	2	3	3
SEP	2	3	2
OKT	2	3	3
NOV	2	3	2
DES	2	3	2

Figure 4. *Centroid* Table Results on *RapidMiner*

In figure 3.3. the results of the *centroid* table from the calculation of *k-medoids* using *rapidminer tools* have the same results as the results of manual calculations.

3.3 Final Result of *K-Medoids* and *RapidMiner Tools*

Tabel 6. Final Result of *K-Medoids* and *RapidMiner Tools*

AFD	BLOCK	*K-Medoids* Algorithm Calculation	Testing With *RapidMiner* App
1	S34	Cluster_2	Cluster_1
1	S30	Cluster_2	Cluster_1
1	S31	Cluster_2	Cluster_1
1	S32	Cluster_2	Cluster_1
1	S33	Cluster_2	Cluster_1
1	S35	Cluster_2	Cluster_1
1	S36	Cluster_2	Cluster_1
1	T32	Cluster_2	Cluster_1
1	T33	Cluster_2	Cluster_1
1	T34	Cluster_2	Cluster_1
1	T35	Cluster_2	Cluster_1
1	T31	Cluster_2	Cluster_1
2	P34	Cluster_2	Cluster_1
2	P35	Cluster_2	Cluster_1
2	P36	Cluster_2	Cluster_1
2	P37	Cluster_1	Cluster_0
2	P34	Cluster_2	Cluster_1
3	B07	Cluster_2	Cluster_1
3	B08	Cluster_2	Cluster_1
3	B09	Cluster_2	Cluster_1
3	B10	Cluster_2	Cluster_1
3	B11	Cluster_2	Cluster_2

4. CONCLUSION

Based on the results of the research that has been done, the following conclusions can be drawn, namely, the clustering of harvest data using the *k-medoids clustering method* resulted in 3 *clusters*, namely *cluster* 1 (low) consisting of 12 data, *cluster* 2 (medium) consisting of 62 data, and *cluster* 3 (high) consisting of 6 data. Knowing the process of testing the *k-medoids* method on harvested harvest data using *rapidminer tools*.

ACKNOWLEDGEMENTS
Thank you to the supervisors as lecturers at AMIK and STIKOM Tunas Bangsa Pematangsiantar who have assisted in the preparation of this research so that it has reached the publication stage. Thanks to all parties, especially PT. Intisariraya Mandau, Riau and it is hoped that this research can be used as a source of data that can assist in decision making. For the development and progress of this research title, we expect constructive input and criticism so that this research can be continued.

REFERENCES

[1] S. D. *et al.*, *The impacts and opportunities of oil palm in Southeast Asia: What do we know and what do we need to know?* 2009.
[2] J. A. Widians, M. Taruk, Y. Fauziah, and H. J. Setyadi, "Decision Support System on Potential Land Palm Oil Cultivation using Promethee with Geographical Visualization," *J. Phys. Conf. Ser.*, vol. 1341, no. 4, pp. 0–9, 2019, doi: 10.1088/1742-6596/1341/4/042011.
[3] A. Sequiño and J. Avenido, "The World's Leader in the Palm Oil Industry: Indonesia," *IAMURE Int. J. Ecol. Conserv.*, vol. 13, Jan. 2015, doi: 10.7718/ijec.v13i1.1074.
[4] D. Khatiwada, C. Palmén, and S. Silveira, "Evaluating the palm oil demand in Indonesia: production trends, yields, and emerging issues," *Biofuels*, vol. 12, no. 2, pp. 135–147, 2021, doi: 10.1080/17597269.2018.1461520.
[5] N. Khan, M. A. Kamaruddin, U. U. Sheikh, and Y. Yusup, "Oil Palm and Machine Learning : Reviewing One Decade of," pp. 1–26, 2021.
[6] M. S. M. Kassim, W. I. W. Ismail, A. R. Ramli, and S. K. Bejo, "Image Clustering Technique in Oil Palm Fresh Fruit Bunch (FFB) Growth Modeling," *Agric. Agric. Sci. Procedia*, vol. 2, pp. 337–344, 2014, doi: 10.1016/j.aaspro.2014.11.047.
[7] V. Issue, N. S. Astuti, I. Karim, and M. Arafat, "Anjoro : International Journal of Agriculture and Business Contribution of Oil Palm (Elaeis guineensis J .) Plantations to Farmers ' Income in West Sulawesi," vol. 1, no. 2, pp. 45–51, 2020, doi: 10.31605/anjoro.v1i2.892.
[8] Y. I. Febiola, I. Cholissodin, and C. Dewi, "Peramalan Hasil Panen Kelapa Sawit Menggunakan Metode Multifactors High Order Fuzzy Time Series yang Dioptimasi dengan K-Means Clustering (Studi Kasus : PT . Sandabi Indah Lestari Kota Bengkulu)," vol. 3, no. 12, 2019, [Online]. Available: http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/6841.
[9] E. F. Himmah, M. Widyaningsih, and M. Maysaroh, "Identifikasi Kematangan Buah Kelapa Sawit Berdasarkan Warna RGB Dan HSV Menggunakan Metode K-Means Clustering," *J. Sains dan Inform.*, vol. 6, no. 2, pp. 193–202, 2020, doi: 10.34128/jsi.v6i2.242.
[10] E. Fauzi *et al.*, "Cluster Analysis of Determining the Location of Oil Palm Replanting in Mukomuko District," *AGRITROPICA J. Agric. Sci.*, vol. 3, no. 1, pp. 38–45, 2020, doi: 10.31186/j.agritropica.3.1.38-45.
[11] D. F. Pasaribu, I. S. Damanik, E. Irawan, Suhada, and H. S. Tambunan, "Memanfaatkan Algoritma K-Means Dalam Memetakan Potensi Hasil Produksi Kelapa Sawit PTPN IV Marihat," *BIOS J. Teknol. Inf. dan Rekayasa Komput.*, vol. 2, no. 1, pp. 11–20, 2021, doi: 10.37148/bios.v2i1.17.
[12] V. Marriboyina and L. C. Reddy, "A Review on Data mining from Past to the Future," *Int. J. Comput. Appl.*, vol. 15, Feb. 2011, doi: 10.5120/1961-2623.
[13] I. B. A. Peling, I. N. Arnawan, I. P. A. Arthawan, and I. G. N. Janardana, "Implementation of Data Mining To Predict Period of Students Study Using Naive Bayes Algorithm," *Int. J. Eng. Emerg. Technol.*, vol. 2, no. 1, p. 53, 2017, doi: 10.24843/ijeet.2017.v02.i01.p11.
[14] A. Palmer, R. Jiménez, and E. Gervilla, "Data Mining: Machine Learning and Statistical Techniques," 2011, pp. 373–396.
[15] D. F. Pramesti, M. Tanzil Furqon, and C. Dewi, "Implementasi Metode K-Medoids Clustering Untuk Pengelompokan Data," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 1, no. 9, pp. 723–732, 2017, doi: 10.1109/EUMC.2008.4751704.
[16] V. A. P. Sangga, "Perbandingan Algoritma K-Means dan Algoritma K-Medoids dalam Pengelompokan Komoditas Peternakan di Provinsi Jawa Tengah Tahun 2015," *Tugas Akhir Jur. Stat. Fak. Mat. dan Ilmu Pengetah. Alam Univ. Islam Inndonesia Yogyakarta*, vol. 53, no. 9, pp. 1689–1699, 2018.
[17] A. D. Andini and T. Arifin, "Implementasi Algoritma K-Medoids Untuk Klasterisasi Data Penyakit Pasien Di Rsud Kota Bandung," *J. RESPONSIF Ris. Sains ...*, vol. 2, no. 2, pp. 128–138, 2020, [Online]. Available: http://ejurnal.ars.ac.id/index.php/jti/article/view/247.
[18] B. Wira, A. E. Budianto, and A. S. Wiguna, "Implementasi Metode K-Medoids Clustering Untuk Mengetahui Pola Pemilihan Program Studi Mahasiwa Baru Tahun 2018 Di Universitas Kanjuruhan

Malang," *J. Terap. Sains Teknol.*, vol. 1, no. 3, pp. 54–69, 2019.

[19] A. Y. Rofiqi, "CLUSTERING BERITA OLAHRAGA BERBAHASA INDONESIA MENGGUNAKAN METODE K-MEDOID," vol. 6, no. 1, 2017.

[20] I. Gunawan, G. Anggraeni, E. S. Rini, and Y. Mustofa, "Klasterisasi provinsi di Indonesia berbasis perkembangan kasus Covid-19 menggunakan metode K-Medoids," *Semin. Nas. Mat. dan Pendidik. Mat.*, pp. 301–306, 2020.

[21] K. Fatmawati and A. P. Windarto, "Data Mining: Penerapan Rapidminer Dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (Dbd) Berdasarkan Provinsi," *Comput. Eng. Sci. Syst. J.*, vol. 3, no. 2, p. 173, 2018, doi: 10.24114/cess.v3i2.9661.

[22] I. Werdiningsih, B. Nuqoba, and Muhammadun, *Data Mining Menggunakan Android, Weka, Dan Spss.pdf*. Jawa Timur: Airlangga University Press, 2020.

[23] M. North, *Data Mining for the Masses*. 2012.

[24] B. G. Sudarsono, M. I. Leo, A. Santoso, and F. Hendrawan, "Analisis Data Mining Data Netflix Menggunakan Aplikasi Rapid Miner," *JBASE - J. Bus. Audit Inf. Syst.*, vol. 4, no. 1, pp. 13–21, 2021, doi: 10.30813/jbase.v4i1.2729.

[25] Ardiyansyah, P. A. Rahayuningsih, and R. Maulana, "Analisis Perbandingan Algoritma Klasifikasi Data Mining Untuk Dataset Blogger Dengan Rapid Miner," *J. Khatulistiwa Inform.*, vol. VI, no. 1, pp. 20–28, 2018.