

Implementasi Algoritma K-Means Clustering Dalam Menentukan Blok Tanaman Sawit Produktif Pada PT Arta Prigel

Yesi Pitaloka Anggriani^{1*}, Alfis Arif², Febriansyah³

^{1,2} Program Studi Teknik Informatika, Institut Teknologi Pagar Alam

*email: yesipitalokaanggriani@gmail.com

DOI: <https://doi.org/10.31603/komtika.v8i1.11192>

Received: 23-03-2024, Revised: 26-04-2024, Accepted: 30-04-2024

ABSTRACT

The purpose of this study is to implement the K-Means Clustering method to determine the patterns of productive oil palm production based on their blocks at Pt Arta Prigel. The research is motivated by issues within the oil palm blocks, such as the absence of productive block summaries, insufficient plantation land analysis, and erroneous decision-making. The development method utilizes CRISP-DM, with data spanning 2 years from October 2021 to October 2023. From the 1275 production records, after cleaning, 1015 records remain. Filtering the initial 51 blocks results in 37 blocks for the years 2021 and 2022, and 46 blocks for the year 2023. After clustering, the production outcomes for the year 2021 are as follows: cluster_0 has 34 blocks, cluster_1 has 2 blocks, and cluster_2 has 10 blocks. For the year 2022, cluster_0 has 36 blocks, cluster_1 has 8 blocks, and cluster_2 has 28 blocks. In the year 2023, cluster_0 has 39 blocks, cluster_1 has 8 blocks, and cluster_2 has 33 blocks. The testing method employs the silhouette coefficient, and the silhouette score testing results indicate the formation of 3 clusters (K=3) with a value of 0.61. The findings of this study include patterns, graphs, and production tables generated using the K-Means Clustering method at PT Arta Prigel.

Keywords: *K-Means Clustering, Palm, Silhouette Coefficient, Rapidminer*

ABSTRAK

Tujuan dari penelitian ini adalah menerapkan metode *K-Means Clustering* untuk mengetahui pola produksi kelapa sawit produktif berdasarkan blok PT Arta Prigel. Penelitian ini dilatarbelakangi oleh permasalahan pada blok kelapa sawit seperti kurangnya ringkasan blok produktif, kurangnya analisis perkebunan dan pengambilan keputusan yang salah. Metode pengembangan menggunakan CRISP-DM, datanya 2 tahun dari Oktober 2021 hingga Oktober 2023. Dari 1275 catatan produksi setelah proses *cleaning* menjadi 1015. 51 blok setelah proses filter menjadi 37 blok pada tahun 2021, 37 blok pada tahun 2022 dan 46 blok pada tahun 2023. Setelah dilakukan proses *clustering* diperoleh hasil produksi tahun 2021, cluster_0 sebanyak 34 blok, cluster_1 sebanyak 2 blok, cluster_2 sebanyak 10 blok. Pada tahun 2022, cluster_0 36 blok, cluster_1 8 blok, cluster_2 28 blok. Pada tahun 2023 cluster_0 39 blok, cluster_1 8 blok, cluster_2 33 blok. Metode pengujian menggunakan *silhouette coefficient*, hasil *silhouette coefficient* menghitung *silhouette score* membentuk 3 cluster (K=3) dengan nilai 0,61. Hasil dari penelitian ini berupa gambar, diagram dan tabel produksi dengan menggunakan metode *K-Means Clustering* dari PT Arta Prigel.

Kata Kunci: *K-Means Clustering, Sawit, Silhouette Coefficient, Rapidminer*

PENDAHULUAN

Teknologi merupakan satu kata yang sangat berperan penuh dalam perkembangan kehidupan manusia saat ini, maraknya kemajuan dan kecanggihan teknologi pada saat ini memberikan perubahan yang sangat signifikan [1]. *Data Mining* atau penambangan data melibatkan analisis data dari berbagai sudut dan mengubahnya menjadi informasi penting yang dapat digunakan untuk meningkatkan keuntungan [2]. Data yang diolah dengan teknik *data mining* akan menciptakan pengetahuan ilmiah baru dari data lama, hasil yang diperoleh dari pengolahan data tersebut dapat digunakan untuk menentukan keputusan di masa depan

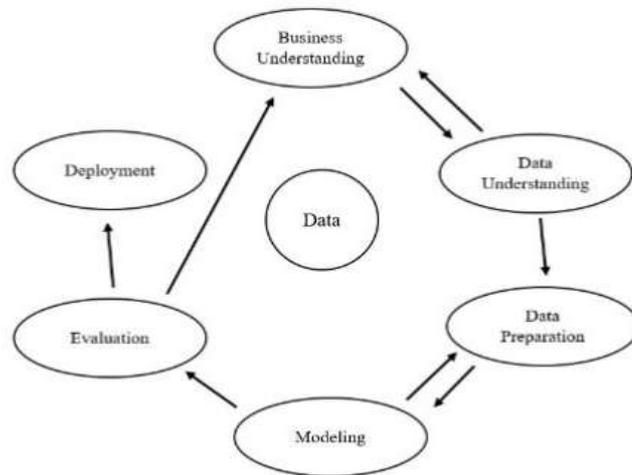
[3]. Algoritma *k-means* adalah algoritma dalam pembelajaran tanpa pengawasan (*unsupervised learning*), algoritma ini digunakan untuk mengelompokkan data menjadi beberapa kelompok. [4]. *K-Means* salah satu algoritma *clustering* yang banyak digunakan untuk mempartisi data ke dalam *cluster* [5]. *Clustering* merupakan metode pembelajaran mesin (*machine learning*) yang digunakan untuk mengelompokkan data ke dalam kelompok atau *cluster* yang sesuai [6]. *Rapid miner* menggunakan metode berorientasi objek dalam hierarki *Java* dan dapat digunakan di hampir semua *platform*. [7]. *Python* bahasa pemrograman tingkat tinggi yang dapat ditafsirkan, *multiplatform* (berfungsi pada berbagai platform termasuk *MS Windows*, *Linux*, dan *Macintosh*) [8]. *Silhouette Coefficient* merupakan metode evaluasi atau validasi algoritma *clustering* yang mengukur kualitas *cluster* yang terbentuk, memvalidasi *cluster* dengan menggabungkan metode separasi dan metode kohesi dan merupakan metode pengujian kualitas *cluster* yang digunakan untuk mengetahui tingkat kepemilikan setiap objek dalam suatu *cluster* [9], [10], [11].

Berdasarkan penelitian yang dilakukan oleh Pulungan dkk[12] hasil penelitian dari Implementasi Algoritma *K-Means Clustering* dalam Menentukan Blok Tanaman Sawit Paling Produktif menghasilkan 2 *cluster* yaitu *cluster* tinggi untuk blok paling produktif dan *cluster* rendah untuk blok tidak produktif, didapatkan 14 blok tanaman sawit paling produktif dan 26 blok tanaman sawit tidak produktif. Berdasarkan penelitian yang dilakukan oleh Alkhairi [3] hasil penelitian dari Penerapan *K-Means Cluster* Pada Daerah Potensi Pertanian Karet Produktif Di Sumatera Utara menghasilkan 3 *cluster* dengan kelompok tertinggi ada 1 item, *cluster* kelompok menengah ada 6 item, *cluster* terendah ada 19 item dengan jumlah data 26. Hasil *cluster* diatas didapatkan *cluster* terendah yang menggunakan algoritma *k-means* untuk menentukan konsistensi data Pertanian Perkebunan Karet. Sedangkan penelitian yang dilakukan oleh Ginting [13] hasil penelitian dari Implementasi Metode *K-Means Clustering* Dalam Pengelompokan Bibit Tanaman Kopi Arabika menghasilkan bibit kopi yang direkomendasikan (C1) terdiri dari 10 item, bibit kopi yang tidak direkomendasikan (C2) terdiri dari 7 jenis bibit kopi dan bibit kopi yang tidak layak (C3) terdiri dari 13 jenis bibit kopi.

Berdasarkan penelitian yang dilakukan oleh [14] hasil penelitian dari Klastering Sayuran Unggulan Menggunakan Algoritma *K-Means* menghasilkan C1 81%, C2 16% dan C3 3%. Cabai Rawit C1 29%, C2 48% dan C3 23%. Kacang Panjang C1 26%, C2 38% dan C3 36%. Kangkung C1 39%, C2 36% dan C3 25%. Terung C1 43%, C2 29% dan C3 28%. Tomat C1 41%, C2 58% dan C3 1%. Pengujian *silhouette coefficient* ketiga *cluster* didapatkan nilai *silhouette* = 0,6798, 3 klaster masih tergolong baik (medium structure) karena nilai *silhouette* berada diantara $0.5 < SC \leq 0.7$. Sedangkan penelitian yang dilakukan oleh [15] hasil penelitian dari Optimalisasi *K-Means* Berbasis *Particle Swarm Optimization* untuk Hasil Produksi Tanaman Sayuran di Indonesia menghasilkan produksi tanaman sayur di Indonesia dengan 3 *cluster*, *cluster* sedang 8,82%, *cluster* rendah 88,2%, *cluster* tinggi 2,94%. Hasil *silhouette coefficient* menunjukkan *k-means* berbasis PSO lebih baik dibanding *k-means* murni dengan nilai rata-rata *silhouette coefficient* sebesar 0,714114, tetapi rata-rata waktu komputasi *k-means* murni lebih cepat dibanding waktu komputasi *k-means* berbasis PSO.

METODE

Penelitian ini menggunakan teknik data mining yaitu *k-means clustering* dan metode pengembangan CRISP-DM (*Cross Industry Standard Process for Data Mining*) merupakan standar pengolahan *data mining* yang dikembangkan untuk memastikan bahwa data yang ada melewati tahapan yang terstruktur dan terdefinisi dengan jelas dan efisien, selain menerapkan model dalam proses penambangan data, pemilihan algoritmanya sangat mempengaruhi komparasi dari kinerja metode *data mining* tersebut [16]. Gambar 1 menjelaskan metode pengembangan dengan CRISP-DM dengan 6 tahapan yaitu *business understanding*, *data understanding*, *data preparation*, *modelling*, *evaluation*, dan *deployment*. Metode pengujian *silhouette coefficient*.



Gambar 1 Metode Crisp-Dm

1. Business Understanding

Pada tahapan ini ada beberapa yang harus dilakukan yaitu, memahami kebutuhan dan tujuan dari sudut pandang bisnis, lalu mengartikan pengetahuan ke dalam bentuk pendefinisian masalah pada *data mining*, kemudian menentukan rencana serta strategi untuk mencapai tujuan *data mining*.

2. Data Understanding

Pada tahapan ini, Tabel 1 menjelaskan langkah pertama yang harus dilakukan yaitu pengumpulan data, lalu mendeskripsikan data, dan mengevaluasi kualitas data.

Tabel 1. Tabel Dataset

No	Blok	Ha	Hasil	Bulan	Tahun
1	1	21.09	5.772	10	2021
2	1	21.09	5.502	11	2021
3	1	21.09	9.120	12	2021
4	2	37.40	5.694	10	2021
5	2	37.40	7.921	11	2021
6	2	37.40	16.413	12	2021
7	13	39.97	8.899	10	2021
8	13	39.97	8.898	11	2021
9	13	39.97	11.131	12	2021
10	14	46.07	7.752	10	2021
.....
1015	51	20.12	17.322	10	2023

3. Data Preparation

Pada tahapan ini, membangun data set dari data mentah. Beberapa hal yang harus dilakukan yaitu, pembersihan data (*Data Cleaning*), pemilihan data (*Data Selection*), record, atribut-atribut, dan transformasi data (*Data Transformation*) yang akan dijadikan masukan dalam tahanan pemodelan.

4. Modeling

Pada tahap ini, *Machine Learning* berhubungan langsung dengan pendefinisian teknik, alat, dan algoritma data mining. Pemodelan yang digunakan pada penelitian ini adalah metode pengelompokan yang menggunakan algoritma *k-means*.

- a. Siapkan dataset
- b. Menentukan jumlah *Cluster*
- c. Titik pusat *cluster*
- d. Hitung jarak data
- e. Pusat *Cluster* berubah
- f. Mengulang langkah 3 sampai 5

5. Evaluation

Pada tahapan ini, melihat tingkat performa dari pola yang dihasilkan oleh algoritma. berfokus pada model yang dihasilkan sudah sesuai standar *K Means Clustering* dan tidak melewatkan satupun tahapan sampai selesai. Tahapan pengujian dilakukan menggunakan metode *Silhouette Coefficient*.

6. Deployment

Pada tahapan ini, melakukan pembuatan laporan dan artikel jurnal menggunakan model yang dihasilkan.

HASIL DAN PEMBAHASAN

Hasil dari penelitian ini adalah *clustering* pola hasil produksi sawit dengan algoritma *k-means* model yang dapat digunakan sebagai rekomendasi *cluster* yang ideal untuk melihat pola hasil produksi sawit selama 2 tahun dari bulan Oktober tahun 2021 sampai dengan bulan Oktober 2023 dengan jumlah blok 51. Hasil pola didapatkan melalui perhitungan algoritma *k-means*, dimulai dari menentukan jumlah *cluster* (*k*), menghitung jarak dengan *euclidean distance*, kemudian menentukan titik *centroid* secara acak, memperbaiki nilai *centroid*. Sebelum melakukan *clustering* data, peneliti menggunakan aplikasi *Rapid Miner* untuk menjadi bantuan bagi peneliti dalam mengelola data dan menghasilkan *cluster*. Proses pada penelitian ini menggunakan tahapan-tahapan yang ada pada metode *crisp-dm*, tahapan pertama yaitu:

1. Business Understanding

Pada tahap ini dilakukan proses menentukan tujuan dan lingkungan pencarian. Maksud dari pencarian ini yaitu mengelompokkan hasil produksi berdasarkan blok, proses pengelolaan data produksi masih dilakukan secara manual atau sederhana dimana masih menggunakan *excel* sehingga belum adanya penjelasan secara rinci mengenai blok yang produktif.

2. Data Understanding

Pada tahap kedua tahap ini tahap proses pemahaman data, dimana data didapat dari bagian KrEstate Pt Arta Prigel yaitu data hasil produksi selama 2 tahun dimulai pada

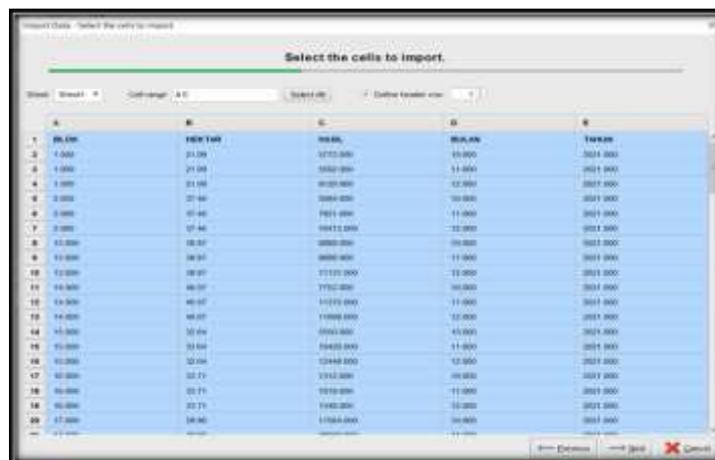
bulan Oktober tahun 2021 sampai dengan bulan Oktober 2023 dengan jumlah 1.275 *record* dengan 5 atribut yaitu Blok, Hektar, produksi, dan total kategori data yang diterima dalam bentuk *excel* dan data yang sudah didapatkan perlu dilakukan proses *cleaning* dan akan dilakukan pemilihan data dengan atributnya.

3. Data Preparation

Tahap ketiga proses pengolahan data dengan langsung mengimplementasikan pada *Rapid Miner* dengan melakukan tiga tahapan yang sudah ditetapkan yaitu:

a. Data Selection

Pemilihan (seleksi) adalah proses pengumpulan informasi tentang data harus dilakukan sebelum memulai penambangan data. Dilakukan pemilihan atribut dan sebagian data dalam atribut akan dikonversikan agar proses data *mining* menjadi lebih mudah. Gambar 2 menjelaskan alur proses seleksi data atau data *selection*. Pada tahapan ini dari 5 atribut yang didapat dari bagian KrEstate Pt Arta Prigel diantaranya Divisi, Blok, Luas Lahan, Produksi Perbulan dan Total. Dilakukan seleksi menggunakan *excel* dengan cara difilter. Setelah dilakukan proses *filter*, data yang peneliti gunakan ada 5 atribut yaitu Blok, Hektar, Hasil, Bulan dan Tahun dengan total jumlah blok 51 setelah di *selection* menjadi 37 blok. Dan hasil *record* produksi dari 1.275 data setelah di *cleaning* menjadi 1.015 data yang akan dijadikan *dataset*.



	A	B	C	D	E
1	26.000	148.100	98.800	96.000	1.000
2	1.000	21.000	1270.000	15.000	3001.000
3	1.000	21.000	1600.000	11.000	3001.000
4	1.000	21.000	1620.000	12.000	3001.000
5	2.000	27.000	1600.000	10.000	3001.000
6	2.000	27.000	1610.000	11.000	3001.000
7	2.000	27.000	16410.000	12.000	3001.000
8	12.000	18.50	1000.000	10.000	3001.000
9	12.000	18.50	1000.000	11.000	3001.000
10	12.000	18.50	1110.000	12.000	3001.000
11	14.000	40.00	1120.000	11.000	3001.000
12	14.000	40.00	1100.000	12.000	3001.000
13	14.000	40.00	1000.000	13.000	3001.000
14	14.000	40.00	1000.000	14.000	3001.000
15	14.000	40.00	1000.000	15.000	3001.000
16	14.000	40.00	1000.000	16.000	3001.000
17	14.000	40.00	1000.000	17.000	3001.000
18	14.000	40.00	1000.000	18.000	3001.000
19	14.000	40.00	1000.000	19.000	3001.000
20	17.000	50.00	1100.000	10.000	3001.000

Gambar 2. Proses Data Selection

b. Data Processing

Processing adalah proses yang melibatkan *cleaning* dan *transformasi* data. Pada tahap ini, peneliti memastikan tidak ada lagi data yang kosong. Jika masih ada data yang kosong pada data yang diproses maka proses akan gagal dan data kosong harus dihapus. Gambar 3 menjelaskan bahwa jika atribut berupa data kosong dengan angka 0 pada data set setiap bulannya, maka dapat digunakan pada langkah berikutnya.



Gambar 3. Proses Data Processing

c. Data Transformation

Tahap terakhir *coding* adalah proses mengubah data terpilih agar sesuai dengan proses *data mining*. Data yang diproses dalam proses transformasi data menampilkan atribut yang dipilih dan diproses di Rapid Miner. Atribut yang akan ditransformasi yaitu blok, sehingga proses penentuan blok dapat diproses oleh Rapid Miner, kemudian atribut hasil ditransformasi untuk memudahkan pemahaman hasil yang diproses oleh Rapid Miner. Tabel 2 menjelaskan kriteria yang digunakan dalam penelitian ini, sedangkan Tabel 3 menggambarkan hasil proses transformasi data 1275 diubah menjadi 1015 data, 111 pada tahun 2021, 444 pada tahun 2022, dan 460 pada tahun 2023. Dari 51 blok, 37 blok pada tahun 2021, 37 blok pada tahun 2022, 46 blok pada tahun 2023. Data produksi dikonversi berdasarkan kriteria yang telah ditentukan.

Tabel 2. Tabel Kriteria Nilai

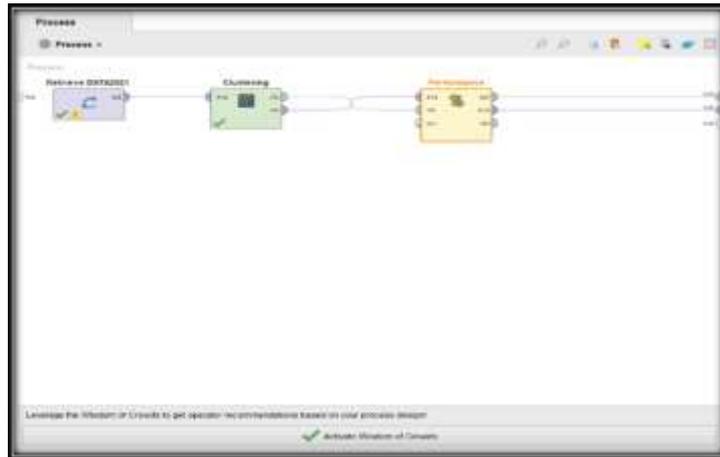
Kriteria	Nilai
Tidak Produktif	C0
Produktif	C1
Cukup Produktif	C2

Tabel 3 Tabel Transformasi Gabungan Tahun 2021-2023

No	Blok	Hektar	Hasil	Bulan	Tahun
1	1	21.09	0	10	2021
2	1	21.09	0	11	2021
3	1	21.09	0	12	2021
4	2	37.40	0	10	2021
5	2	37.40	0	11	2021
6	2	37.40	2	12	2021
7	13	39.97	0	10	2021
8	13	39.97	0	11	2021
9	13	39.97	0	12	2021
10	14	46.07	0	10	2021
....
1015	51	20.12	0	10	2023

4. Modeling

Pada proses ini, proses dimana model digunakan dengan algoritma yang digunakan dalam metode *clustering* yaitu operator *clustering* model *k-means* mengambil objek dari *port input* dan mengirimkan salinannya ke *output*. Gambar 4 menjelaskan *dataset* yang digunakan akan di proses modeling menggunakan algoritma *k-means clustering*, untuk proses *modeling*.



Gambar 4. Proses Model Algoritma K-means Pada Rapid Miner

Selanjutnya pada tahap *modeling* akan dilakukan proses tahapan algoritma k-means, tahap pertama:

1. Tentukan Jumlah Cluster

Untuk menentukan jumlah *cluster* perlu dilakukan percobaan 3 *cluster*, yaitu C0, C1, dan C2 dengan atribut Blok, Hektar, Hasil, Bulan dan tahun. Gambar 5 menjelaskan dengan pengukuran *performance vector* rata-rata dalam *centroid distance* 23000739.091, rata-rata *cluster_0* 28064575.850, rata-rata *cluster_1* 89406062.799, dan rata-rata dalam *centroid cluster_2* 13922135.927 dan nilai *Davies Bouldin Index* adalah 0.511

Attribute	cluster_0	cluster_1	cluster_2
HEKTAR	32.885	34.750	30.820
HASIL	24418.787	48831.707	9704.185
BULAN	6.882	6.881	6.306
TAHUN	2022.578	2022.568	2022.182

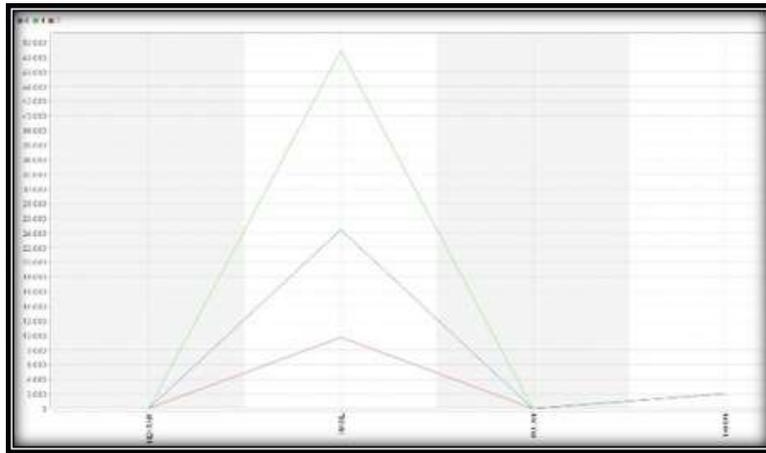
Gambar 5. Percobaan Jumlah Cluster

2. Titik Centroid Secara Acak

Titik *centroid* ditentukan secara acak berdasarkan beberapa eksperimen klaster di Rapid Miner. Dari ketiga pengujian tersebut dipilih titik pusat terkecil untuk mendapatkan hasil terbaik.

3. Jarak Data Ke Centroid

Setelah mengimpor data set dan algoritma *k-means* ke dalam *Rapid Miner*, proses data dengan mengklik Run untuk menampilkan jarak ke pusat data. Dari ketiga pengujian tersebut dipilih titik pusat terkecil untuk mendapatkan hasil terbaik. Gambar 6 menjelaskan hasil nilai jarak data *cluster_0* dimulai 0 sampai 10.00 untuk *cluster* tidak produktif, *cluster_1* dimulai 24.00 sampai 48.00 untuk *cluster* produktif, dan *cluster_2* dimulai 10.00 sampai 24.00 untuk *cluster* cukup produktif.



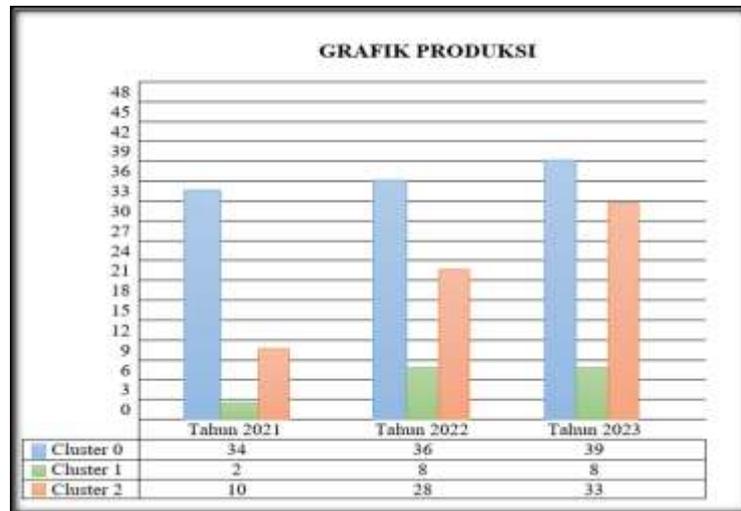
Gambar 6. Jarak Data ke Centroid

4. Perbarui Nilai Centroid

Setelah beberapa percobaan, *centroid* pada 3 cluster tidak berubah. Jika titik berat berubah maka iterasi diulangi, tetapi jika tidak berubah maka pengulangan dihentikan dan diperoleh hasil dari setiap kelompoknya. Berdasarkan perhitungan *Rapid Miner* maka diperoleh pola yang nantinya akan diimplementasikan dalam *Python*. Pola yang digunakan untuk mengelompokkan *Cluster* data berdasarkan hasil perhitungan jarak adalah :

- Jika $C_0 < C_1$ dan $C_0 < C_2$ maka *Cluster_0* dengan keterangan Tidak Produktif
- Jika $C_1 < C_0$ dan $C_1 < C_2$ maka *Cluster_1* dengan keterangan Produktif
- Jika $C_1 < C_0$ dan $C_2 < C_1$ maka *Cluster_2* dengan keterangan Cukup Produktif

Gambar 7 menjelaskan dari data produksi yang diolah dengan *Rapid Miner* yang digunakan menggunakan metode *Clustering K-Means* maka diperoleh pola hasil produksi perblok yaitu *Cluster_0* memiliki tingkat produksi tidak produktif, *Cluster_1* memiliki tingkat produksi produktif, dan *Cluster_2* memiliki tingkat produksi cukup produktif.



Gambar 7. Hasil Grafik

5. Evaluation

Pengujian dilakukan menggunakan *Silhouette Coefficient* terhadap cluster yang sudah terbentuk dengan seluruh data 1015 record. Gambar 8 menjelaskan hasil yang didapatkan dari *silhouette score* setelah dilakukan beberapa kali percobaan, didapatkan hasil *cluster* yang cocok berjumlah 3 *cluster* ($K=3$) dengan nilai kualitasnya yaitu 0,61.

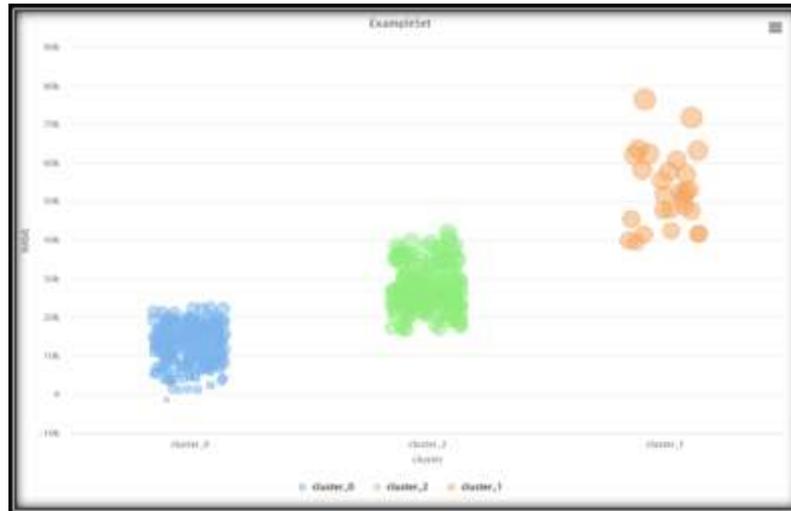
```
Kode + Teks
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:570: FutureWarning: The default value of 'n_init'
warnings.warn(
for n_clusters = 2, the average silhouette_score is: 0.8247551881690137
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:570: FutureWarning: The default value of 'n_init'
warnings.warn(
for n_clusters = 3, the average silhouette_score is: 0.6185667318598869
for n_clusters = 4, the average silhouette_score is: 0.584979471833381
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:570: FutureWarning: The default value of 'n_init'
warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:570: FutureWarning: The default value of 'n_init'
warnings.warn(
for n_clusters = 5, the average silhouette_score is: 0.5539685788468412
for n_clusters = 6, the average silhouette_score is: 0.5285787464354899
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:570: FutureWarning: The default value of 'n_init'
warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:570: FutureWarning: The default value of 'n_init'
warnings.warn(
for n_clusters = 7, the average silhouette_score is: 0.5306587881387137
for n_clusters = 8, the average silhouette_score is: 0.5424088935168519
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:570: FutureWarning: The default value of 'n_init'
warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:570: FutureWarning: The default value of 'n_init'
warnings.warn(
for n_clusters = 9, the average silhouette_score is: 0.5485963845957651
for n_clusters = 10, the average silhouette_score is: 0.547194246974793
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:570: FutureWarning: The default value of 'n_init'
warnings.warn(
```

Gambar 8. Hasil Pengujian

6. Deployment

Pada tahap *deployment* merupakan tahapan terakhir berupa pengetahuan atau informasi yaitu mengenai pola hasil produksi Pt Arta Prigel selama 2 tahun sehingga dapat diketahui ada 3 *cluster*. Gambar 9 menjelaskan pola yang dihasilkan yang pertama tahun 2021 *cluster_0* 34 blok, kemudian *cluster_1* 2 blok, dan *cluster_2* 10 blok. Tahun 2022 *cluster_0* 36 blok, *cluster_1* 8 blok, dan *cluster_2* 28 blok. Tahun 2023 *cluster_0* 39 blok, *cluster_1* 8 blok, dan *cluster_2* 33 blok. Dari proses *clustering* dapat dilihat tingkat produksi produktif memiliki peningkatan pada tahun

2022 dan 2023, dan untuk hasil produksi pada Pt Arta Prigel lebih dominan berada pada tingkat produksi tidak produktif.



Gambar 9. Hasil Pola

KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, maka diperoleh hasil sebagai berikut: Penelitian ini menghasilkan 3 Pola *Clustering* hasil produksi Pada Pt Arta Prigel yaitu *cluster_0* dengan keterangan Tidak Produktif, *cluster_1* dengan keterangan Produktif, dan *cluster_2* dengan keterangan Cukup Produktif. Dari hasil *Clustering* Produksi Sawit Pt Arta Prigel dapat dilihat pola hasil produksi selama 2 tahun dengan total 51 blok. Diketahui produksi tahun 2021 untuk *cluster_0* yaitu 34 blok, *cluster_1* yaitu 2 blok, dan *cluster_2* yaitu 10 blok. Selanjutnya produksi tahun 2022 untuk *cluster_0* yaitu 36 blok, *cluster_1* yaitu 8 blok, dan *cluster_2* 28 blok. Dan tahun 2023 *cluster_0* yaitu 39 blok, *cluster_1* yaitu 8 blok, dan *cluster_2* yaitu 33 blok. Kemudian hasil pada tahap pengujian menggunakan *Silhouette Coefficient* pada aplikasi *google colab* dengan bahasa pemrograman *python* untuk menghitung hasil dari *silhouette score* diperoleh jumlah *cluster* yang cocok yaitu $K=3$ dengan jumlah nilai 0.61. Nilai yang diperoleh pada pengujian *silhouette score* menunjukkan bahwa kualitas dari *cluster* tersebut tepat. Dari hasil tersebut akan didapatkan beberapa pengetahuan yang diharapkan dapat bermanfaat bagi pihak perusahaan Pt Arta Prigel terutama bagian Manager dan Asisten Kebun untuk penunjang keputusan dan melakukan tindakan selanjutnya.

DAFTAR PUSTAKA

- [1] P. Amita Tri Prasasti and C. Dewi, "Pengembangan Assesment of Inovation Learning Berbasis Revolusi Industri 4.0. untuk Guru Sekolah Dasar," *J. Ilm. Sekol. Dasar*, vol. 4, no. 1, p. 66, 2020, doi: 10.23887/jisd.v4i1.24280.
- [2] R. I. O. Limabri, F. Putrawansyah, and A. Arif, "Penerapan Data Mining Untuk Mengklasifikasi Nasabah Bank Sumsel Babel Menggunakan Algoritma C4. 5," *Escaf*, pp. 1101–1108, 2023.
- [3] P. Alkhairi and A. P. Windarto, "Penerapan K-Means Cluster Pada Daerah Potensi Pertanian Karet Produktif di Sumatera Utara," pp. 762–767, 2019.

- [4] & Z. 2021 Zulfa, Auliya, Permata, “Analisis Data Mining Untuk Clustering Kasus COVID-19 di Provinsi Lampung Dengan Algoritma K-Means,” vol. 2, no. 2, pp. 100–108, 2021.
- [5] F. Febriansyah and S. Muntari, “Penerapan Algoritma K-Means untuk Klasterisasi Penduduk Miskin pada Kota Pagar Alam,” vol. 8, no. 1, pp. 66–77, 2023.
- [6] M. R. Nahjan, N. Heryana, A. Voutama, F. I. Komputer, U. S. Karawang, and R. Miner, “Implementasi Rapidminer Dengan Metode Clustering K-Means Untuk Analisa Penjualan Pada Toko Oj Cell,” vol. 7, no. 1, pp. 101–104, 2023.
- [7] S. Rokhanah, A. Hermawan, and D. Avianto, “Pengaruh Principal Component Analysis Pada Naïve Bayes dan K-Nearest Neighbor Untuk Prediksi Dini Diabetes Melitus Menggunakan Rapidminer,” vol. 11, no. 1, 2023.
- [8] K. Ademariana, F. R. Lumbanraja, and R. Andrian, “Jurnal Pepadun Clustering K-Means Jenis Kata Pada Laporan Kegiatan Kuliah Kerja Nyata (KKN) Universitas Lampung Menggunakan Word2vec © 2021 Ilmu Komputer Unila Publishing Network all rights reserved Jurnal Pepadun,” vol. 3, no. 2, pp. 221–228, 2021.
- [9] M. Astiningrum, M. Mentari, and Y. G. Maranatha, “Mutu Buah Salak Menggunakan Pengolahan Citra Digital,” pp. 205–210.
- [10] N. Afdhaliah, “Perbandingan kinerja algoritma ward dan algoritma k-means dengan uji silhouette coefficient,” 2020.
- [11] F. N. Dhewayani, D. Amelia, D. N. Alifah, B. N. Sari, and M. Jajuli, “Implementasi K-Means Clustering untuk Pengelompokan Daerah Rawan Bencana Kebakaran Menggunakan Model CRISP-DM,” vol. 12, no. 1, pp. 64–77, 2022, doi: 10.34010/jati.v12i1.6674.
- [12] I. M. Pulungan, M. Fauzan, and A. P. Windarto, “Implementasi Algoritma K-Means Clustering dalam Menentukan Blok Tanaman Sawit Paling Produktif,” no. September, pp. 338–348, 2019.
- [13] B. Ginting and F. Riandari, “Implementasi Metode K-Means Clustering Dalam Pengelompokan Bibit Tanaman Kopi Arabika,” vol. 3, no. 2, pp. 151–157, 2020.
- [14] A. K-means, “Klastering Sayuran Unggulan Menggunakan Clustering of Featured Vegetables Using the K-Means Algorithm,” vol. 8, pp. 567–579, 2022.
- [15] Budiman and T. Parama Yoga, “Optimalisasi K-Means Berbasis Particle Swarm Optimization untuk Hasil Produksi Tanaman Sayuran di Indonesia,” *J. Nuansa Inform.*, vol. 17, pp. 2614–5405, 2023.
- [16] M. A. Hasanah, S. Soim, and A. S. Handayani, “Implementasi CRISP-DM Model Menggunakan Metode Decision Tree dengan Algoritma CART untuk Prediksi Curah Hujan Berpotensi Banjir,” vol. 5, no. 2, 2021.

