



## *Implementation Of K-Means Algorithm For Palm Oil Productivity Data Clustering*

### **Implementasi Algoritma K-Means Untuk Klustering Data Produktivitas Kelapa Sawit**

**Indah Ramadhani<sup>1</sup>, Megawati<sup>2</sup>**

<sup>1,2</sup>Program Studi Sistem Informasi Fakultas Sains dan Teknologi  
Universitas Islam Negeri Sultan Syarif Kasim Riau

E-Mail : [12050320362@students.uin-suska.ac.id](mailto:12050320362@students.uin-suska.ac.id), [megawati@uin-suska.ac.id](mailto:megawati@uin-suska.ac.id)

*Makalah: Diterima 15 Desember 2022; Diperbaiki 28 Desember 2022; Disetujui 09 Januari 2023  
Corresponding Author: Indah Ramadhani*

#### **Abstrak**

Hasil produksi kelapa sawit merupakan salah satu hal yang sangat diperhatikan pada perusahaan PT Adei Plantation dan industry. Mengelompokkan kualitas bibit sawit perlu dilakukan untuk mengetahui kelompok jenis bibit kelapa sawit yang berkualitas baik dan tidak baik sehingga proses penanaman kelapa sawit dapat dilakukan dengan baik pula. Belum diterapkannya pengelompokan buah yang berpotensi baik untuk diproduksi oleh perusahaan yang terdapat pada daerah Mandau, serta perusahaan masih belum bisa membuat perencanaan merupakan kasus yang dialami perusahaan ini. Oleh karena itu, diperlukan alternatif untuk mempermudah dalam pengelompokan buah bersumber pada blok panennya. Melalui pengelompokan menggunakan K-Means, pembagian kelompok terbagi atas produksi berdasarkan hasil panen, hasil biji buah (brondol) dan hasil produksi. Dalam penelitian ini dilakukan pengelompokan blok panen hasil produksi buah dengan menggunakan algoritma K-Means. Algoritma K-Means merupakan salah satu algoritma yang digunakan untuk mengelompokkan data ke dalam beberapa kelompok dan termasuk kedalam kelompok Unsupervised learning. Algoritma ini bekerja berdasarkan kemiripan data yang dimiliki antar kelompoknya. Hasil dari penelitian ini didapatkan C0 ialah 13 data, dan C1 ialah 35 data.

**Kata Kunci:** *Algoritma K-Means, pengelompokan, Data Mining, hasil produksi*

#### **Abstract**

*Palm oil production is one of the things that is of great concern to PT Adei Plantation and the industry. Classifying the quality of oil palm seeds needs to be done to find out which groups of types of oil palm seeds are of good and bad quality so that the process of planting oil palm can be carried out properly as well. Companies in the Mandau area have not implemented fruit grouping that has good potential to be produced, and companies have not been able to make plans, which is the case with this company. Therefore, an alternative is needed to facilitate the grouping of fruit based on the harvest block. Through grouping using K-Means, the division of groups is divided into production based on crop yields, fruit seed yields (brondol) and production yields. In this research, block grouping of harvested fruit production was carried out using the K-Means algorithm. The K-Means algorithm is one of the algorithms used to group data into several groups and is included in the Unsupervised learning group. This algorithm works based on the similarity of the data held between groups. The results of this study showed that C0 was 13 data, and C1 was 35 data.*

**Keywords :** *K-Means Algorithm, grouping, Data Mining, production results*

#### **1. PENDAHULUAN**

Indonesia termasuk negara agraris dengan hasil komoditas perkebunan sangat tinggi dan berperan penting untuk perekonomian negara [1]. Perkebunan adalah sektor penting bagi pembangunan negara[2][3]. Tanaman komoditas utama di perkebunan Indonesia yaitu kelapa sawit. Saat ini, Indonesia merupakan salah satu negara di dunia yang sangat banyak menghasilkan minyak kelapa sawit. Dalam perdagangan Internasional, minyak kelapa sawit merupakan komoditas utama dalam transaksi ekspor impor dan merupakan sumber

pendapatan devisa negara [4]. Kelapa sawit yang memiliki nama latin *Elaeis guineensis* Jacq yang memiliki artian golongan tanaman yang tumbuh subur dikawasan tropis[5]. Tanaman ini dapat menghasilkan minyak nabati yang berperan penting dalam kehidupan keseharian kita terutama dalam hal memasak [6].

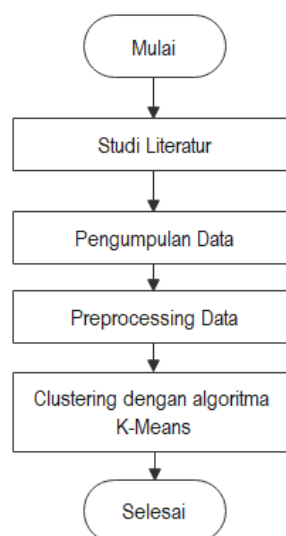
PT Adei Plantation & Industry merupakan Perusahaan Modal Asing (PMA) milik Malaysia yang beroperasi pada bagian perkebunan kelapa sawit yang dahulu bernama PT Adei Crumb Rubber Factory dan berdiri pada tahun 1954. Pada perusahaan ini belum dilakukannya pengelompokan terhadap hasil produksi dikarenakan belum adanya paerencanaan dalam menentukan potensi buah yang bagus akibat terlalu luasnya kebun pada wilayah Mandau. Untuk melakukan penelitian ini salah satunya dengan melakukan klasterisasi jumlah hasil produksi kelapa sawit dengan menggunakan teknik data mining. Data mining merupakan proses mengekstrasi dan mengidentifikasi pola untuk mendapatkan informasi yang berasal dari database yang besar. Teknik data mining etrbagi menjadi 5 yaitu teknik statistik, matematika, AI, dan machine learning[7]. Salah satu algoritma dalam metode data mining yaitu K-Means yang akan digunakan dalam mengelompokkan hasil panen kelapa sawit yang berpotensi baik.

Berdasarkan penelitian Deny Franata Pasaribu dkk pada tahun 2021 mengenai algoritma k-means dalam mengelompokkan hasil produksi kelapa sawit dapat memberikan informasi mengenai kelapa sawit yang baik dan kurang baik berdasarkan blok panen[5]. Sama halnya dengan penelitian Rizal Efendi Marpaung dkk pada tahun 2022, algoritma k-means dalam memetakan kualitas bibit kelapa sawit dapat memberikan informasi mengenai kelapa sawit yang baik dan kurang baik berdasarkan jumlah tandan[8]. Selanjutnya penelitian yang dilakukan oleh Dimas Alif dkk mengenai pemetaan lahan kopi di kabupaten malang yang menghasilkan 3 cluster yaitu rendah, sedang dan tinggi yang dapat membantu pada implementasi sistem[9]. Berikutnya penelitian yang dilakukan oleh Anderas Nugroho dkk mengenai pengklasteran covid-19 tingkat provinsi di Indonesia menjadi 2 kluster, yaitu sebanyak 30 provinsi termasuk cluster sedang dan 4 provinsi termasuk cluster tinggi [10]. Selanjutnya penelitian yang dilakukan oleh Dona dan Rifqi pada tahun 2022, mengenai K-means clustering untuk menentukan status gizi yaitu baik atau buruk pada balita. Hasil dari penelitian ini yaitu dapat memberikan pengetahuan mengenai status gizi yang dialami oleh balita di Kab. Rokan Hulu dan dapat menjadi acuan yang tepat dalam mengambil keputusan yang optimal[11].

Berdasarkan beberapa penelitian yang telah diuraikan diatas, menunjukkan algoritma K-Means sangat berperan penting dan berhasil dalam mengklaster/ mengelompokkan data- data menjadi beberapa kelompok. Penelitian-penelitian terkait inilah yang melatarbelakangi dilakukannya penelitian untuk mengelompokkan hasil panen kelapa sawit yang berpotensi pada PT Adei Plantation & Industry Kebun Mandau Selatan, Duri menggunakan Algoritma K-Means Klustering.

## 2. BAHAN DAN METODOLOGI PENELITIAN

Terdapat 5 tahapan yang dilakukan dalam penelitian ini, yaitu Perencanaan dan Studi Literatur, Pengumpulan data, preprocessing data, Analisa, dan Pembahasan. Tahapan awal yaitu Pengambilan data. Data yang dipakai adalah data sekunder, dataset yang digunakan adalah data produktivitas kelapa sawit yang diperoleh dari PT Adei Plantation and Industry pada bulan Februari-September sebanyak 194 item perhari dan dikelompokkan menjadi 8 bulan dengan menggunakan tools Ms. Excel dan Rapid Miner. Setelah itu dilakukan preprocessing data. Tahap preprocessing ini dengan melakukan cleaning dan normalisasi agar tidak ada kesetimpangan data. Lalu dilakukan proses cluster dengan menggunakan K-Means dan akan di analisa cluster buah berdasarkan blok panen dan selesai. Gambar 1 merupakan Metodologi Penelitian.



**Gambar 1.** Metodologi Penelitian

## 2.1 Data Mining

Data mining merupakan adalah proses pengolahan data yang berkaitan langsung dengan big data yang tersimpan dalam database yang besar. Ada beberapa teknik yang termasuk ke dalam data mining. Teknik tersebut yaitu teknik Artificial Intelligence (AI), machine learning, matematika dan statistika. Data mining ini digunakan untuk mengekstrasi sejumlah informasi yang berguna untuk pengetahuan [12]. Fase- fase didalam data mining yaitu klustering, prediksi, klasifikasi, Estimasi dan Asosiasi [13].

## 2.3 Clustering

Clustering merupakan cara untuk mengelompokkan beberapa record data dengan memperhatikan atau membentuk kelas objek ke dalam kelompok dimana didalam kelompok tersebut diisi oleh data yang semirip mungkin dan memiliki perbedaan dari kelompok lainnya. [14]. Clustering ini termasuk kedalam sifat tanpa arahan (unsupervised learning) dimana data-data yang memiliki kemiripan akan membentuk kelompok sendiri sesuai dengan kesamaan polanya, sedangkan data yang berbeda akan dimasukkan dalam cluster yang berbeda pula. [15] [16]. Metode k-means sangat efisien untuk kumpulan data yang lebih kecil [17]. Secara umum, analisis Cluster dapat dikatakan bahwa [18] [19]:

1. Data pada suatu cluster yang sama, dapat memiliki kemiripan yang tinggi
2. Data pada suatu cluster yang berbeda, dapat memiliki tingkat kemiripan yang rendah

## 2.4 K-Means

K-Means adalah algoritma yang populer dalam dunia riset dan masuk kedalam kelompok unsupervised learning yang dilihat dari kemiripan data yang dimiliki [20]. Tujuan dari algoritma ini adalah memaksimalkan keragaman antar kluster atau kelompok dan meminimalkan suatu fungsi objektif dalam satu kelompok[21]. Kelebihan dari algoritma ini yaitu sederhana dalam pengimplementasiannya [3]. secara matematis, rumus K-Means yaitu[22]:

$$C_i = \frac{1}{M} \sum_{j=1}^M X_j$$

Berdasarkan rumus yang dapat dilihat diatas, dapat diketahui  $C_i$  adalah Centroid fitur ke-1,  $M$  merupakan banyak data dalam suatu kelompok, adapun  $i$  adalah fitur ke- $i$  dalam suatu kelompok. Berikut ini merupakan tahapan untuk menyelesaikan algoritma K-means [23] :

1. Menetapkan banyak data untuk diklaster
2. Menetapkan nilai  $k$  jumlah
3. Menentukan nilai centroid awal dengan random berdasarkan data variabel berdasarkan nilai yang telah ditentukan sebelumnya dengan rumus :

$$D_{(i,j)} = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2}$$

Berdasarkan rumus diatas, dapat dielaskan bahwa  $D_{(i,j)}$  yaitu jarak data ke  $i$  ke pusat *cluster*  $j$ ,  $X_{ki}$  adalah data ke  $i$  pada atribut data ke  $k$  dan  $X_{kj}$  = titik pusat ke  $j$  pada atribut ke  $k$

4. Menghitung jarak antar data terhadap centroid.
5. Menentukan posisi klaster data berdasar pada jarak minimal pada pusat centroid
6. Menghitung nilai hasil centroid berdasarkan hasil iterasi. Bila hasilnya sama dengan perhitungan sebelumnya, maka proses iterasi selesai. Namun bila masih terdapat pergantian posisi terhadap data, maka proses iterasi masih dilanjutkan ke tahap selanjutnya.

## 2.5 Rapidminer

Rapidminer adalah salah satu tools yang digunakan dalam mengolah data untuk data mining. Rapidminer digunakan pada bidang machine learning, text mining, information mining, serta analisa prediktif. Tujuannya adalah untuk mendapatkan data paling bermutu tinggi dari bacaan yang diolah [24].

### 3. ANALISA DAN HASIL

#### 3.1 Pengumpulan Data

Penelitian ini menerapkan menggunakan clustering dengan algoritma K-Means dalam mengelompokkan data produktivitas kelapa sawit pada PT Adei Plantation & Industry. Analisa menggunakan algoritma K-Means digunakan untuk mendapatkan cluster hasil produksi buah kelapa sawit. Dalam penelitian ini data yang digunakan yaitu hanya data pada tahun 2022.

**Tabel 1.** Hasil cleaning data

BLOK PANEN	BULAN	BUAH (TON)	BRONDOLAN (TON)	PRODUKSI (TON)
KM1	FEBRUARI	41.254	2.978	44.232
KM2		44.465	5.352	49.817
KM3		32.298	3.105	35.403
KM4		39.021	2.497	41.518
KM5		40.094	3.625	43.719
KM6		43.014	3.463	46.604
...	...	...	...	...
KM1	SEPTEMBER	31.365	3.178	34.543
KM2		37.475	5.529	43.004
KM3		29.069	4.521	33.590
KM4		30.778	2.725	33.503
KM5		33.307	3.590	36.897
KM6		35.926	3.569	39.495

#### 3.2 Pre- processing

Pada tahap ini, dataset akan dilakukan pembersihan dan dilakukan normalisasi agar tidak terjadinya kesetimpangan data antara nilai yang rendah dengan nilai yang tinggi, supaya tidak adanya jarak antar data dan parameter yang mendominasi saat perhitungan dilakukan. Tabel 2 Hasil Normalisasi Data.

**Tabel 2.** Hasil Normalisasi Data

BLOK PANEN	BULAN	BUAH (TON)	BRONDOLAN (TON)	PRODUKSI (TON)
KM1	FEBRUARI	0,322	0,082	0,281
KM2		0,407	0,486	0,427
KM3		0,085	0,103	0,050
KM4		0,263	0,000	0,210
KM5		0,291	0,192	0,267
KM6		0,369	0,164	0,343
...	...	...	...	...
KM6	SEPTEMBER	0,181	0,182	0,157

#### 3.3 Pengolahan Data

##### 1. Menentukan banyak kelompok yang akan diklaster dan centroid

Data yang telah dibersihkan dari noise-noise selanjutnya akan diclustering menjadi 2 buah cluster atau kelompok. Penentuan titik tengah atau nilai centroid ini dibagi atas 2 bagian yakni C0 dan C1

Pembagian centroid atau cluster ini dilakukan dengan menentukan nilai tertinggi untuk C0 dan nilai terkecil untuk C1 berdasarkan atribut nya masing masing. Untuk atribut Buah yang tertinggi (C0) yaitu 1,000 sedangkan yang terkecil (C1) yaitu 0,000. Kemudian Untuk brondolan yang tertinggi (C0) adalah 1,000 sedangkan yang terkecil (C1) adalah 0,000 Kemudian Untuk hasil Produksi yang tertinggi (C0) adalah 1,000 sedangkan yang terkecil (C1) adalah 0,000. Untuk nilai centroid dapat dilihat pada tabel berikut.

**Tabel 3.** Nilai Centroid

C0	1,000	1,000	1,000
C1	0,000	0,000	0,000

Berdasarkan tabel 3 diatas, nilai centroid yang diperoleh berdasarkan atribut masing-masing dapat diketahui yaitu C0 bernilai 1,000 dan C1 bernilai 0,000

**2. Perhitungan jarak ke centroid**

Dalam menghitung jarak antara centroid objek data digunakan rumus Euclidean Distance sebagai berikut [25]:

$$De = \sqrt{(xi - si)^2 + (yi - ti)^2} \quad (1)$$

Perhitungan jarak data ke centroid-1 yaitu sebagai berikut:

$$D_{A1,C0} = \sqrt{(0,322 - 1,000)^2 + (0,082 - 1,000)^2 + (0,281 - 1,000)^2} = 1,349$$

$$D_{A2,C0} = \sqrt{(0,407 - 1,000)^2 + (0,486 - 1,000)^2 + (0,427 - 1,000)^2} = 0,972$$

$$D_{A3,C0} = \sqrt{(0,085 - 1,000)^2 + (0,103 - 1,000)^2 + (0,050 - 1,000)^2} = 1,595$$

...

$$D_{A48,C0} = \sqrt{(0,181 - 1,000)^2 + (0,182 - 1,000)^2 + (0,157 - 1,000)^2} = 1,432$$

Selanjutnya, dilakukan perhitungan jarak data ke centroid-2 sebagai berikut :

$$D_{A1,C1} = \sqrt{(0,322 - 0,000)^2 + (0,082 - 0,000)^2 + (0,281 - 0,000)^2} = 0,435$$

$$D_{A2,C1} = \sqrt{(0,407 - 0,000)^2 + (0,486 - 0,000)^2 + (0,427 - 0,000)^2} = 0,764$$

...

$$D_{A48,C1} = \sqrt{(0,181 - 0,000)^2 + (0,182 - 0,000)^2 + (0,157 - 0,000)^2} = 0,301$$

Setelah dilakukan perhitungan untuk kedua centroid yang telah ditentukan, maka didapatkan tabel jarak data dari centroid sebagai berikut :

**Tabel 4.** Nilai Iterasi Ke 1

CO	C1	JARAK TERDEKAT	HASIL KLUSTER
1,349	0,435	0,435	C1
0,972	0,764	0,764	C1
1,595	0,143	0,143	C1
1,472	0,336	0,336	C1
1,301	0,440	0,440	C1
1,237	0,529	0,529	C1
0,805	0,935	0,805	C0
0,516	1,229	0,516	C0
0,994	0,754	0,754	C1
0,990	0,788	0,788	C1
0,604	1,469	0,604	C0
0,768	1,085	0,768	C0
1,039	0,693	0,693	C1
0,458	1,389	0,458	C0
0,597	1,148	0,597	C0
1,207	0,537	0,537	C1
0,924	0,810	0,810	C1
0,689	1,050	0,689	C0
1,185	0,548	0,548	C1

0,458	1,297	0,458	C0
1,232	0,502	0,502	C1
1,115	0,686	0,686	C1
0,637	1,291	0,637	C0
0,797	1,008	0,797	C0
1,025	0,710	0,710	C1
1,081	0,653	0,653	C1
1,043	0,691	0,691	C1
1,074	0,677	0,677	C1
1,064	0,679	0,679	C1
1,013	0,729	0,729	C1
1,352	0,414	0,414	C1
0,980	0,753	0,753	C1
0,897	1,196	0,897	C0
1,299	0,574	0,574	C1
1,230	0,525	0,525	C1
1,255	0,502	0,502	C1
0,957	0,782	0,782	C1
0,712	1,031	0,712	C0
0,969	0,802	0,802	C1
1,165	0,579	0,579	C1
1,029	0,705	0,705	C1
0,796	0,950	0,796	C0
1,616	0,134	0,134	C1
1,185	0,614	0,614	C1
1,557	0,344	0,344	C1
1,684	0,060	0,060	C1
1,510	0,235	0,235	C1
1,432	0,301	1,432	C1

### 3. Penentuan kelompok atau cluster

Untuk menentukan Cluster dengan nilai Cluster yang sesuai dapat dilihat pada hasil di iterasi 1. Perhitungan akan dihentikan jika hasil iterasi sudah sama pada proses iterasi sebelumnya. Selanjutnya untuk mendapatkan hasil centroid baru, dapat dilakukan dengan menjumlahkan nilai pada cluster tersebut dan membaginya sebanyak jumlah nilai itu. Hasil dari centroid baru adalah sebagai berikut:

#### Cluster 0 :

$$C_{0a} = (0,565 + 0,640 + 1,000 + 0,733 + 0,631 + 0,595 + 0,620 + 0,664 + 0,870 + 0,668 + 0,860 + 0,530 + 0,583 / 13) = \mathbf{0,689}$$

$$C_{0b} = (0,467 + 0,780 + 0,396 + 0,333 + 1,000 + 0,735 + 0,547 + 0,837 + 0,391 + 0,357 + 0,132 + 0,665 + 0,456 / 13) = \mathbf{0,546}$$

$$C_{0c} = (0,580 + 0,703 + 1,000 + 0,726 + 0,728 + 0,651 + 0,647 + 0,735 + 0,870 + 0,665 + 0,821 + 0,584 + 0,596 / 13) = \mathbf{0,716}$$

Berdasarkan perhitungan diatas, didapatkan data Centroid baru untuk Iterasi ke-2 yaitu :

**Tabel 5.** Centroid Baru

C0	0,689	0,546	0,716
C1	0,329	0,270	0,318

Berdasarkan nilai centroid baru seperti tabel 5 diatas, diperoleh tabel antara jarak data dari centroid dan nilai terendah atau minimal dari centroid. Seperti yang tertera pada tabel 6 berikut :

Tabel 6. Nilai Iterasi ke 2

C0	C1	JARAK TERDEKAT	HASIL KLUSTER
0,735	0,192	0,192	C1
0,409	0,253	0,253	C1
1,002	0,399	0,399	C1
0,858	0,299	0,299	C1
0,696	0,101	0,101	C1
0,623	0,116	0,116	C1
0,200	0,404	0,200	C0
0,239	0,710	0,239	C0
0,452	0,267	0,267	C1
0,372	0,280	0,280	C1
0,447	0,965	0,447	C0
0,217	0,578	0,217	C0
0,452	0,169	0,169	C1
0,458	0,890	0,458	C0
0,221	0,630	0,221	C0
0,600	0,050	0,050	C1
0,330	0,279	0,279	C1
0,097	0,519	0,097	C0
0,588	0,021	0,021	C1
0,293	0,779	0,293	C0
0,646	0,080	0,080	C1
0,498	0,225	0,225	C1
0,284	0,782	0,284	C0
0,196	0,493	0,196	C0
0,425	0,179	0,179	C1
0,503	0,148	0,148	C1
0,446	0,160	0,160	C1
0,462	0,159	0,159	C1
0,456	0,151	0,151	C1
0,405	0,199	0,199	C1
0,740	0,174	0,174	C1
0,405	0,234	0,234	C1
0,460	0,744	0,460	C0
0,683	0,285	0,285	C1
0,618	0,088	0,088	C1
0,643	0,098	0,098	C1
0,404	0,276	0,276	C1
0,239	0,516	0,239	C0
0,464	0,339	0,339	C1
0,557	0,065	0,065	C1
0,433	0,175	0,175	C1
0,184	0,420	0,184	C0

1,027	0,425	0,425	C1
0,661	0,277	0,277	C1
1,012	0,462	0,462	C1
1,088	0,485	0,485	C1
0,925	0,327	0,327	C1
0,838	0,236	0,236	C1

Dari tabel 6 diatas diperoleh hasil cluster pada proses iterasi 1 dan 2 memiliki hasil cluster yang tidak berubah. Karena hal tersebut maka proses iterasi dapat dihentikan. Dan diperoleh hasil yaitu :

1. Cluster 0 (C0) mendapatkan 13 data Panen
2. Cluster 1 (C1) mendapatkan 35 data panen.

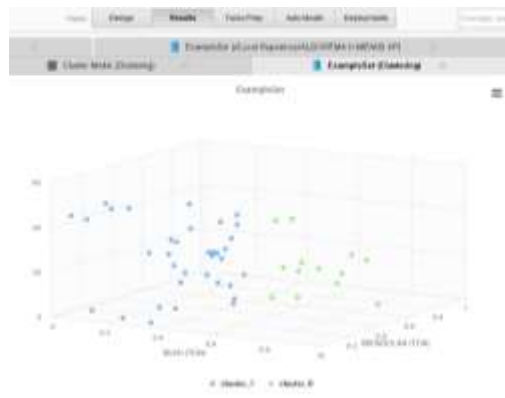
### 3.4. Hasil clustering dengan *RapidMiner*

Pada algoritma ini tahapan pengujian digunakan untuk melihat hasil kebenaran pada analisa pada tahap yang telah dilakukan sebelumnya serta diuji dengan cara manual, sehingga butuh pengujian lagi dalam pengelompokan data. Untuk melihat pengolahan menggunakan tools, Gambar 2 merupakan hasil dari pengujian yang telah dilakukan.



**Gambar 2.** Data hasil clustering

Dari hasil yang telah diproses pada rapidminer seperti gambar 2 diatas, didapatkan bahwa Cluster\_0 merupakan kelompok hasil panen sebanyak 13 data dan Cluster\_1 merupakan kelompok hasil panen sebanyak 35 data. Maka didapatkan hasil visualisasi menggunakan rapidminer seperti gambar berikut :



**Gambar 3.** Grafik clustering

Berdasarkan gambar 3 diatas, dapat terlihat bahwa penyebaran cluster 0 yaitu sebanyak 13 item yang ditandai dengan warna hijau, dan cluster 1 sebanyak 35 item yang ditandai dengan warna biru.

## 4. KESIMPULAN

Berdasarkan penelitian dalam mengelompokkan hasil panen kelapa sawit ini menggunakan 2 cluster, yang mana algoritma ini dapat diterapkan terhadap pengelompokan hasil panen kelapa sawit. Dan juga dapat difungsikan kedalam tools rapidminer. Hasil akhir dari penelitian ini yaitu didapatkan C0 sebanyak 13 item informasi Blok Panen, serta C1 sebanyak 35 item informasi Blok Panen. Dalam hal ini pengelompokan hasil pembuatan kelapa sawit yang diterapkan dalam rapidminer sama dengan hasil manual

## REFERENSI

- [1] T. Yuniarti and D. Hayati, "Segmentasi Perkebunan Kelapa Sawit dengan Data Mining Teknik K-Means Clustering Berdasarkan Luas Areal, Produksi dan Produktivitas," *Invent. Ind. Vocat. E-Journal Agroindustry*, vol. 2, no. 2, p. 56, 2021, doi: 10.52759/inventory.v2i2.47.
- [2] M. Sitorus, H. Vande, N. Sinaga, R. Pramono, and D. Kusumo, "Clusterisasi Perkebunan Kelapa Sawit



- Berdasarkan Luas Area Produksi Dengan Algoritma K-MEANS,” *J. Informatics ...*, vol. 3, no. 1, 2022, [Online]. Available: <https://journal.univpancasila.ac.id/index.php/jiac/article/view/3660%0Ahttps://journal.univpancasila.ac.id/index.php/jiac/article/download/3660/1934>
- [3] 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 ),” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 12, 2019, [Online]. Available: <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/6841>
  - [4] S. Hajar, A. A. Novany, A. P. Windarto, A. Wanto, and E. Irawan, “Penerapan K-Means Clustering pada ekspor minyak kelapa sawit menurut negara tujuan,” *Semin. Nas. Teknol. Komput. Sains 2020*, pp. 314–318, 2020.
  - [5] 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.
  - [6] W. J. Mawaddah, I. Gunawan, and I. P. Sari, “Implementation of Data Mining Algorithm for Clustering of Palm Oil Harvested Data,” *JOMLAI J. Mach. Learn. Artif. Intell.*, vol. 1, no. 1, pp. 43–54, 2022, doi: 10.55123/jomlai.v1i1.163.
  - [7] I. M. Pulungan, S. Saifullah, M. Fauzan, and A. P. Windarto, “Implementasi Algoritma K-Means Clustering dalam Menentukan Blok Tanaman Sawit Paling Produktif,” *Pros. Semin. Nas. Ris. Inf. Sci.*, vol. 1, no. September, p. 338, 2019, doi: 10.30645/senaris.v1i0.39.
  - [8] R. E. Marpaung, “Penerapan Algoritma K-Means Dalam Mengclustering Kualitas Bibit Kelapa Sawit Di PPKS Marihat,” *Bul. Big Data, Data Sci. Artif. Intell. Penerapan*, vol. 1, no. 1, pp. 7–15, 2022.
  - [9] D. Alif Fajar Fadhillah, A. Faisol, and N. Vendyansyah, “Penerapan Metode K-Means Clustering Pada Pemetaan Lahan Kopi Di Kabupaten Malang,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 6, no. 1, pp. 162–170, 2022, doi: 10.36040/jati.v6i1.4617.
  - [10] A. N. Sihananto, A. Puspita Sari, H. Khariono, R. Akhmad Fernanda, and D. Cakra Mudra Wijaya, “Implementasi Metode K-Means Untuk Pengelompokan Kasus Covid-19 Tingkat Provinsi Di Indonesia,” *J. Inform. dan Sist. Inf.*, vol. 3, no. 1, pp. 76–85, 2022, doi: 10.33005/jifosi.v3i1.472.
  - [11] G. Baik, D. A. N. Gizi, B. Pada, S. Kasus, K. Rokan, and K. R. Miner, “Dona, 2) Mi’rajul Rifqi ABSTRAK,” vol. 7, no. 2, 2022.
  - [12] B. Lesmana, “Pengelompokan Pengiriman Hasil Kelapa Sawit Berdasarkan Tonase dan Kualitas Menggunakan Metode Clustering (Studi Kasus: KUD Bumi Pusaka),” *Semin. Nas. Inform.*, pp. 154–165, 2021.
  - [13] Y. Andini *et al.*, “PENERAPAN DATA MINING TERHADAP TATA LETAK BUKU,” vol. XI, no. 1, pp. 9–15, 2022.
  - [14] A. Nofiar, S. Defit, and Sumijan, “Penentuan Mutu Kelapa Sawit Menggunakan Metode K-Means Clustering,” *J. KomtekInfo*, vol. 5, no. 3, pp. 1–9, 2019, doi: 10.35134/komtekinfo.v5i3.26.
  - [15] W. D. H. Hendra Effendi, Ahmad Syahrial, Sefran Prayoga, “Penerapan Metode K-Means Clustering untuk Pengelompokan Lahan Sawit Produktif pada PT Kasih Agro Mandiri,” *Teknomatika*, vol. 11, no. 02, pp. 117–126, 2021.
  - [16] F. Yunita, “Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Pada Penerimaan Mahasiswa Baru,” *Sistemasi*, vol. 7, no. 3, p. 238, 2018, doi: 10.32520/stmsi.v7i3.388.
  - [17] Y. Hawari, J. T. Hardinata, and R. A. Nasution, “Implementasi K-Means Clustering Dalam Menentukan Kualitas Biji Kelapa Sawit ( Kasus PPKS Marihat ),” vol. 1, no. 1, pp. 16–23, 2022.
  - [18] N. Nuraisana, “Analisis Clustering Untuk Mengetahui Tingkat Potensi Tanaman Kelapa Sawit Berdasarkan Luas Tanaman Menggunakan Algoritma K ...,” *J. Mantik Penusa*, vol. 3, no. 1, pp. 110–123, 2019, [Online]. Available: <http://e-jurnal.pelitanusantara.ac.id/index.php/mantik/article/view/727>
  - [19] T. Pribadi, R. Irsyada, H. Audytra, and D. A. Fatah, “9006-24189-2-Pb,” vol. 9, no. 1, pp. 20–28, 2020.
  - [20] N. Dwitri, J. A. Tampubolon, S. Prayoga, F. I. R.H Zer, and D. Hartama, “Penerapan Algoritma K-Means Dalam Menentukan Tingkat Penyebaran Pandemi Covid-19 Di Indonesia,” *J. Teknol. Inf.*, vol. 4, no. 1, pp. 128–132, 2020, doi: 10.36294/jurti.v4i1.1266.
  - [21] A. F. Tanjung, T. M. Diansyah, and R. Rismayanti, “Pemanfaatan Algoritma K-Means Clustering Sebagai Pengamanan Pencurian Buah Kelapa Sawit Se-Distrik Tandun PT. Perkebunan Nusantara V,” *J. Media Inform. Budidarma*, vol. 3, no. 4, p. 351, 2019, doi: 10.30865/mib.v3i4.1443.
  - [22] R. Supardi and I. Kanedi, “Implementasi Metode Algoritma K-Means Clustering pada Toko Eidelweis,” *J. Teknol. Inf.*, vol. 4, no. 2, pp. 270–277, 2020, doi: 10.36294/jurti.v4i2.1444.
  - [23] P. Alkhairi and A. P. Windarto, “Penerapan K-Means Cluster pada Daerah Potensi Pertanian Karet Produktif di Sumatera Utara,” *Semin. Nas. Teknol. Komput. Sains*, pp. 762–767, 2019.
  - [24] D. Suhendro, I. S. Damanik, and M. Fauzan, “Implementasi K-Means Clustering Untuk Mengelompokkan Hasil Pertanian Kacang Kedelai ( Ha ) Berdasarkan Provinsi,” *Senaris*, vol. 2, no. 5, pp. 192–199, 2020.
  - [25] E. Andriani, L. Zuni, A. Nasution, and R. A. Azdy, “KLASTERISASI LAHAN SAWIT PADA SRIWIJAYA PALM OIL PALEMBANG MENGGUNAKAN ALGORITMA K-MEANS,” pp. 1432–1440, 2022.