



# Pengelompokan Data Janjang Panen Kelapa Sawit Menggunakan Algoritma *K-Medoids* Pada PT SIR MANDAU

Lila Agustini<sup>1</sup>, Sumarno<sup>2</sup>, Ika Okta Kirana<sup>3</sup>

<sup>123</sup>Program Studi Teknik Informatika, STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

[lila.agustini20@gmail.com](mailto:lila.agustini20@gmail.com)

**Abstrak**— Perkebunan kelapa sawit hampir menyebar di seluruh wilayah Indonesia, kelapa sawit juga merupakan tumbuhan tropis yang tergolong dalam family palmea dan berasal dari Afrika. Kelapa sawit merupakan tanaman dengan nilai ekonomis yang cukup tinggi karena satu tanaman penghasil minyak. Begitu juga dengan Indonesia yang merupakan salah satu penghasil kelapa sawit. penelitian ini menggunakan teknik *data mining* dalam pengolahan atau pengelompokan data dengan metode *k-medoids clustering*. Metode *k-medoids* merupakan metode *clustering* yang berfungsi untuk memecahkan data set menjadi beberapa kelompok. Kelebihan dari metode ini mampu mengatasi kelemahan dari metode *k-means* yang sensitif terhadap *outlier*. Kelebihan lain metode ini yaitu hasil proses *clustering* tidak bergantung pada urutan masuk data set. Metode *k-medoids clustering* dapat diterapkan pada data hasil janjang panen kelapa sawit berdasarkan Tingi, sedang dan rendahnya hasil panen, sehingga dapat diketahui pengelompokan panen berdasarkan data tersebut. Diharapkan penelitian ini dapat memberikan informasi kepada para karyawan tentang pengelompokan data janjang panen kelapa sawit yang berdampak pada hasil panen kelapa sawit untuk kedepannya.

**Kata Kunci:** Data janjang panen; K-Medoid.

**Abstract**— Oil palm plantations are almost spread throughout Indonesia. Oil palm is also a tropical plant belonging to the palmea family and comes from Africa. Oil palm is a plant with high economic value because it is an oil-producing plant. This research uses data mining techniques in processing or grouping data with the *k-medoids clustering* method. The *k-medoids* method is a clustering method that functions to break the dataset into several groups. The advantages of this method are able to overcome the weakness of the *k-medoids* method which is sensitive to outliers. Another advantage of this method is that the results of the clustering process do not depend on the order in which the dataset is entered. The *k-medoids clustering* method can be applied to data on oil palm harvest yields based on high, medium and low yields, so that harvest grouping can be known based on these data. It is hoped that this research can provide information to employees about the grouping of oil palm harvesting data which has an impact on oil palm yields in the future.

**Keywords:** Data of harvesting age; *K-Medoids*.

## 1. PENDAHULUAN

Panen adalah serangkaian kegiatan mulai dari memotong tandan matang panen sesuai kriteria, mengutip dan mengumpulkan brondolan, menyusun tandan di tempat pengumpulan hasil. Sebaliknya kegagalan panen akan menghambat pencapaian produksi perkebunan kelapa sawit. Pengelolaan tanaman yang sudah baku dan potensi produksi di pokok tinggi tidak ada artinya jika panen tidak dilaksanakan secara optimal. Standar kematangan berikut ini berdasarkan jumlah brodolan yang ada di permukaan tanah. Sangat penting untuk mempertahankan panen pada interval yang pendek pada tanaman muda, karena buah akan membrondol lebih dari 10% dalam waktu 5-7 hari, interval panen yang lama mengakibatkan banyaknya buah busuk dan jumlah brondolan yang banyak. Upaya untuk mengoptimalkan produksi maka kegiatan panen, angkut harus dikelola dengan baik, karena kegiatan panen, angkut, olah merupakan kegiatan yang tidak terpisah dan menentukan kualitas panen.

Metode *k-medoids* dapat diterapkan pada pengelompokan data janjang panen berdasarkan hasil produksi panen per bulannya, sehingga dapat diketahui pengelompokan hasil produksi berdasarkan data tersebut. Dari data pengelompokan tersebut dilihat karakteristiknya sehingga diketahui *cluster* rendah, *cluster* sedang dan *cluster* tinggi dengan presentasi target pada setiap hasil panen per bulannya. Algoritma *clustering* salah satunya adalah *K-Medoids* yang termasuk salah satu metode pengelompokan dalam data mining. Algoritma ini menggunakan objek pada kumpulan objek untuk mewakili sebuah cluster. Metode *K-Medoids* cukup efisien untuk data set yang kecil. Kelebihan dari metode *K-Means* yang *sensitive* terhadap outlier dan hasil *clustering* tidak tergantung pada urutan masuk dataset [1].

*K-Medoids* merupakan suatu algoritma yang digunakan untuk menemukan medoids di dalam sebuah kelompok (*cluster*) yang merupakan titik pusat dari suatu kelompok (*cluster*). Algoritma *K-Medoids* lebih baik dibandingkan dengan *K-Means* karena pada *K-Medoids* kita menemukan *k* sebagai objek yang representatif untuk meminimalkan jumlah ketidaksamaan objek data, sedangkan pada *K-Means* menggunakan jumlah jarak *euclidean distances* untuk objek data[2].

Salah satu cara mengatasi masalah tersebut adalah dengan melakukan pengklasteran terhadap hasil produksi panen kelapa sawit. Tujuannya adalah untuk mengetahui hasil panen kelapa sawit per bulannya dengan potensi

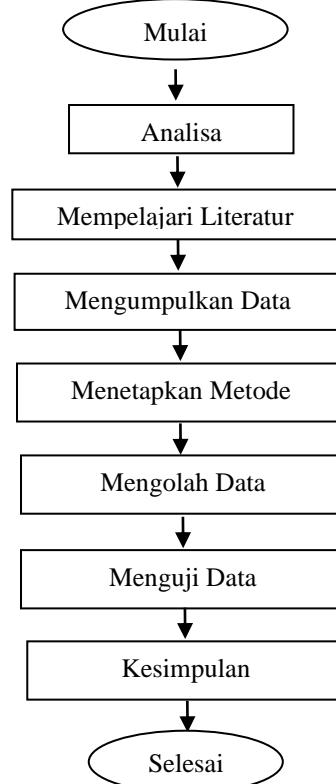


panen berdasarkan rendah, sedang dan tingginya hasil panen yang diperoleh oleh PT SIR MANDAU. Dari pengelompokan data tersebut dilihat karakteristiknya sehingga diketahui jumlah banyak sedikitnya hasil panen per bulannya.

## 2. METODOLOGI PENELITIAN

### 2.1 Rancangan Penelitian

Rancangan penelitian ini pertama kali penulis melakukan pengamatan dan menganalisis permasalahan yang ada setelah itu data akan diolah melalui proses perhitungan dan mengikuti langkah-langkah proses perhitungan Algoritma *K-Medoids*. Selanjutnya hasil perhitungan tersebut akan diaplikasikan ke *RapidMiner* untuk melihat keakuratan hasil yang diperoleh. Flowchart Rancangan Penelitian dapat dilihat pada gambar berikut ini:



**Gambar 1.** Perancangan Penelitian

- Analisis Masalah Menganalisis masalah yang terkait hasil janjang panen kelapa sawit di PT SIR MADAU.
- Mempelajari Literatur Penelitian ini harus didasari rujukan yang digunakan untuk mendapatkan rujukan yang digunakan untuk mendapatkan informasi dalam penelitian.
- Menetapkan Metode Menetapkan metode untuk memecahkan masalah. Pada penelitian ini metode yang digunakan algoritma *K-Medoids*.
- Mengumpulkan Data di ambil melalui riset yang telah dilakukan penulis di PT SIR MANDAU.
- Mengelola Data Melakukan pengelolahan data dengan menggunakan *RapidMiner* versi 5.3.
- Menguji Data Pengujian data dilakukan dengan menggunakan aplikasi *RapidMiner* versi 5.3. sebagai sistem untuk mengcluster jumlah janjang panen kelapa sawit.
- Kesimpulan yang didapatkan adalah hasil *cluster* tertinggi dan terendah panen kelapa sawit setiap bulannya.

### 2.2 Data Mining

Data Mining dapat dibagi menjadi empat kelompok, yaitu model prediksi (prediction modelling), analisis kelompok (Cluster analysis), analisis asosiasi (association analysis) dan deteksi anomali (anomaly detection)[3]. Data mining merupakan proses analitik yang dirancang untuk memeriksa sejumlah data yang besar dalam mencari suatu pengetahuan tersembunyi yang berharga dan konsisten [4].

Proses KDD secara garis besar dapat dijelaskan sebagai berikut [5]:

#### a. *Data Selection*

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam *Knowledge Discovery in Database*(KDD) dimulai. Data hasil seleksi yang akan digunakan untuk proses data mining, disimpan dalam suatu berkas terpisah dari basis data operasional.



b. *Pre-processing / Cleaning*

Sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses cleaning pada data yang menjadi fokus *Knowledge Discovery in Database* (KDD). Proses cleaning mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak. Juga dilakukan proses enrichment, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk *Knowledge Discovery in Database* (KDD), seperti data atau informasi eksternal lainnya yang diperlukan.

c. *Transformation Coding*

Proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Proses *coding* dalam *Knowledge Discovery in Database* (KDD) merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

d. *Data Mining*

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik-teknik, metode-metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses *Knowledge Discovery in Database* (KDD) secara keseluruhan.

e. *Interpretation / Evaluation*

Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan.

### **2.3. Clustering**

*Clustering* merupakan beberapa data atau objek ke dalam *cluster* sehingga setiap dalam *cluster* lainnya. Analisa *cluster* adalah teknik *multivariant* yang memiliki tujuan utama sebagai pengelompokan objek-objek berdasarkan karakteristik yang dimilikinya. Banyak algoritma *clustering* yang pemakaianya tergantung pada tipe data yang akan dikelompokkan dan apa tujuan dari pembuatan aplikasinya. Algoritma ini dapat dipakai untuk pengelompokan objek ke dalam *cluster-cluster*, lalu hasil klastering akan di deteksi keberadaan *outlier* dalam data tersebut.

Algoritma *clustering* diklasifikasikan ke dalam 5 kategori[6]:

a. *Partitioning methods*

Pengelompokan objek dimana tiap objek dimiliki oleh 1 *cluster*. Algoritma partisi dapat meminimalkan mean kuadrat jarak dari setiap titik data ke pusat terdekatnya. Yang termasuk ke dalam metode ini adalah algoritma *k-means*, *k-medoids* atau *partitioning around medoids*, *clustering large application*, dan *clustering large application based on randomized*.

b. *Hierarchical method*

Objek dapat dikelompokkan dengan 2 cara, yaitu *agglomerative* dimulai dengan menyatukan beberapa *cluster* sampai menjadi satu, atau *divide* dimulai dengan *cluster* yang sama kemudian dibagi menjadi beberapa *clustering using representatives*, *balance iterative reducing and clustering*, dan *chameleon*.

c. *Density-based method*

Pengklasteran objek menurut tingkat kedekatan objek atau densitas. Algoritma yang terdapat di metode ini yaitu *density-based spatial clustering algorithm with noise*, *density based clustering*, dan *ordering points to identify the clustering structure*.

d. *Grid-based method*

Pengklasteran objek menggunakan struktur data *grid* multiresolusi yang mampu mengatasi data berdimensi tinggi. Algoritma yang terdapat di metode ini yaitu *clustering in quest*, *wave cluster*, dan *statistical information grid*.

e. *Model-based method*

Pengklasteran objek dengan memodelkan tiap *cluster* dan mencoba mengoptimalkan data yang sesuai dengan model matematika. Algoritma yang terdapat di metode ini yaitu algoritma *cobweb*.

### **2.4 Algoritma K-Medoids**

*K-Medoids* merupakan suatu algoritma yang digunakan untuk menemukan medoids didalam sebuah kelompok (*cluster*) yang merupakan titik pusat dari suatu kelompok (*cluster*). Algoritma *K-Medoids* lebih baik dibandingkan dengan *K-Means* karena pada *K-Medoids* kita menemukan *k* sebagai objek yang representatif untuk meminimalkan jumlah ketidaksamaan objek data, sedangkan pada *K-Means* menggunakan jumlah jarak *euclidean distances* untuk objek data[7].

Langkah-langkah algoritma K-Medoids [8]:

- a. Inisialisasi pusat cluster sebanyak *k* (jumlah cluster).
- b. Alokasikan setiap data (objek) ke cluster terdekat menggunakan persamaan ukuran jarak Euclidian Distance dengan persamaan.
- c. Pilih secara acak objek pada masing-masing cluster sebagai kandidat medoid baru.
- d. Hitung jarak setiap objek yang berada pada masing-masing cluster dengan kandidat medoid baru.



- e. Hitung total simpangan ( $S$ ) dengan menghitung nilai total distance baru – total distance lama. Jika  $S < 0$ , maka tukar objek dengan data cluster untuk membentuk sekumpulan k objek baru sebagai medoid.
- f. Ulangi langkah 3 sampai 5 hingga tidak terjadi perubahan medoid, sehingga didapatkan cluster beserta anggota cluster masing-masing.

#### 2.4.1 Flowchart

Flowchart adalah simbol dari suatu algoritma untuk menyelesaikan suatu permasalahan memudahkan melakukan pengecekan bagian yang dilupakan dalam menganalisis masalah. *Flowchart* berguna untuk berkomunikasi untuk pemrograman suatu proyek[9].

#### 2.4.2 Rapidminer

*RapidMiner* memudahkan penggunanya pada melakukan perhitungan data yang sangat banyak menggunakan memakai operator-operator. Operator ini berfungsi buat memodifikasi data. Data dihubungkan menggunakan node-node dalam operator lalu pengguna hanya tinggal menghubungkan ke node *output* buat melihat hasilnya. Hasil yang diperlihatkan *RapidMiner* dapat ditampilkan secara visual menggunakan grafik membuat *RapidMiner* merupakan satu perangkat lunak pilihan buat melakukan ekstraksi data menggunakan metode-metode data mining [10].

### 3. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan Data Mining metode *K-Medoids*, pada tahapan ini dilakukan analisis data hasil janjang panen kelapa sawit berdasarkan data yang ada dengan tool *RapidMin*. *RapidMiner* adalah sebuah *software* pilihan untuk melakukan ekstraksi data dengan metode-metode Data Mining.

#### 3.1 Penerapan Algoritma *K-Medoids Clustering*

Pada tahapan ini pengolahan data janjang panen kelapa sawit pada PT SIR MANDAU menggunakan algoritma *K-Medoids*. Data janjang panen kelapa sawit pada PT SIR MANDAU diambil penulis melalui penelitian. Langkah-langkah dalam menyelesaikan perhitungan manual data mining menggunakan *K-Medoids Clustering* yaitu:

- a. Inisialisasi pusat *cluster* sebanyak 3 *cluster* dari data sampel. Untuk pemilihan setiap *medoids* dipilih secara acak. Seperti pada tabel 1. berikut ini.

Tabel 1. Nilai Pusat *Cluster* Pada *Medoids Awal*

No.	AFD	Luas (Ha)	Janjang Panen											
			Jan	Feb	Mar	Apr	Mei	Jun	Jul	Agu	Sep	Okt	Nov	Des
1	1	30,27	3	2	3	4	4	5	5	4	4	3	3	4
2	1	19,17	1	1	1	1	1	1	1	1	1	1	1	1
3	1	10,41	2	1	2	2	3	3	2	2	2	2	2	2

- b. Menghitung nilai jarak (cost) dengan persamaan *Eulidion Distance*: untuk menghitung jarak antara titik *centroid* dengan titik tiap objek menggunakan *Eulidion Distance*. Rumus untuk menghitung jarak menggunakan persamaan dua (3). Maka perhitungan untuk jarak pada *medoids* ke -1 adalah sebagai berikut:

Perhitungan jarak C1:

$$C(30,27) = \sqrt{((3-3)^2 + (3-2)^2 + (2-3)^2 + (4-4)^2 + (2-4)^2 + (4-5)^2 + (4-5)^2 + (5-4)^2 + (4-4)^2 + (4-3)^2 + (3-3)^2 + (2-4)^2} = 3,605551275 \quad (1)$$

Perhitungan jarak C2:

$$C(19,17) = \sqrt{((3-1)^2 + (3-1)^2 + (2-1)^2 + (4-1)^2 + (2-1)^2 + (4-1)^2 + (4-1)^2 + (5-1)^2 + (4-1)^2 + (4-1)^2 + (3-1)^2 + (2-1)^2} = 8,717797887 \quad (2)$$

Perhitungan jarak C3:

$$C(10,41) = \sqrt{((3-2)^2 + (3-1)^2 + (2-2)^2 + (4-2)^2 + (2-3)^2 + (4-3)^2 + (4-2)^2 + (5-2)^2 + (4-2)^2 + (4-2)^2 + (3-2)^2 + (2-2)^2} = 6,32455532 \quad (3)$$



Hasil dari seluruh dapat dilihat pada tabel 2 sebagai berikut:

**Tabel 2.** Hasil Perhitungan Algoritma *K-Medoids* Iterasi satu

No	Afd	Luas (Ha)	C1	C2	C3	Terdekat	Cluster yang diikuti
1.	1	28,36	3,605551275	8,717797887	6,32455532	3,605551275	0
2.	1	28,69	2,449489743	8,185352772	4,898979486	2,449489743	0
3.	1	28,61	2,236067977	8,366600265	5	2,236067977	0
4.	1	29,37	2,828427125	9,219544457	5,477225575	2,828427125	0
5.	1	27,97	2,645751311	9,591663047	5,744562647	2,645751311	0
6.	1	29,01	3,16227766	8,306623863	5,477225575	3,16227766	0
7.	1	28,11	4	9,643650761	6,782329983	4	0
8.	1	30,27	1	9,695359715	5,916079783	1	0
9.	1	29,96	2	10,34408043	5,830951895	2	0
10.	1	27,77	1,732050808	9,591663047	5,567764363	1,732050808	0
11.	1	29,68	1,414213562	8,888194417	5,477225575	1,414213562	0
12.	1	27,84	2,828427125	9,433981132	5,830951895	2,828427125	0
13.	1	29,55	2,828427125	8,185352772	5,656854249	2,828427125	0
14.	1	18,93	4,582575695	6,32455532	4,358898944	4,358898944	2
15.	1	28,00	3,31662479	7,348469228	4,582575695	3,31662479	0
16.	1	20,68	5,567764363	5,830951895	5	5	2
17.	1	30,64	3,464101615	11,18033989	6,633249581	3,464101615	0
18.	1	27,70	3	10,39230485	7,141428429	3	0
19.	1	15,78	6,403124237	3,16227766	5	3,16227766	1
20.	1	11,24	3,741657387	7,549834435	4	3,741657387	1
21.	1	18,97	7,348469228	5	4,898979486	4,898979486	2
22.	1	19,17	9,327379053	0	7	0	2
23.	1	18,72	4,69041576	6,403124237	4,69041576	4,69041576	2
24.	1	21,09	4,472135955	7,416198487	4,472135955	4,472135955	2
25.	1	24,89	3,16227766	8,306623863	5,656854249	3,16227766	0
26.	1	25,19	3	9,055385138	5,916079783	3	0
27.	1	25,67	2,645751311	9,695359715	6,08276253	2,645751311	0
28.	1	24,13	2,645751311	9,16515139	5,916079783	2,645751311	0
29.	1	7,72	6,92820323	4,123105626	4,242640687	4,123105626	1
30.	1	24,57	5,196152423	7,745966692	4,582575695	4,582575695	2
31.	1	24,52	2,828427125	9,643650761	6,92820323	2,828427125	0
32.	1	25,97	3	9,695359715	5,916079783	3	0
33.	1	25,27	2,828427125	9,643650761	5,656854249	2,828427125	0
34.	1	21,93	5,744562647	6	4,582575695	4,582575695	2
35.	1	10,05	6,92820323	4,472135955	5,196152423	4,472135955	1
36.	1	13,21	6,92820323	3,31662479	5,656854249	3,31662479	1
37.	1	10,41	5,196152423	4,472135955	4,582575695	4,472135955	1
38.	2	27,33	3,741657387	7,280109889	5,477225575	3,741657387	0
39.	2	28,89	2,645751311	8,485281374	5	2,645751311	0
40.	2	29,11	4	7,681145748	5,477225575	4	0
41.	2	28,66	3,872983346	8,246211251	6,557438524	3,872983346	0
42.	2	29,42	3,605551275	7,615773106	4,795831523	3,605551275	0
43.	2	28,96	4,69041576	5,744562647	4,69041576	4,69041576	0
44.	2	29,29	3	8,602325267	5,385164807	3	0
45.	2	28,77	3,31662479	8,485281374	5,744562647	3,31662479	0
46.	2	28,91	4,242640687	7,937253933	6,164414003	4,242640687	0
47.	2	11,34	8,717797887	2,236067977	6,32455532	2,236067977	1
48.	2	15,69	9,055385138	1	6,782329983	1	1
49.	2	17,71	5,567764363	8,246211251	4,358898944	4,358898944	2
50.	2	18,97	5,567764363	4,242640687	3,872983346	3,872983346	2

1. Mengitung Nilai *Cost*

Nilai *cost* diperoleh dari total penjumlahan nilai jarak dekat *medoids* yang di peroleh : 163,0791384

2. Lakukan ulang 3 langkah di atas dengan nilai pusat *medoids* baru. Algoritma *Medoids* dimulai pada penentuan awal pusat *cluster* dengan memilih secara acak di antara objek yang ada pada dataset. Jumlah objek yang dipilih sesuai dengan jumlah *cluster* yang diinginkan, hal ini dikarenakan objek yang terpilih akan merepresentasikan pusat dari *cluster* sebagai *medoids*, untuk kasus ini akan dipilih tiga buah objek sebagai *medoids* iterasi 2.

Seperti pada tabel 3. berikut ini:

**Tabel 3.** Nilai Pusat *Cluster* pada Medoids Awal

No	AFD	Luas (Ha)	Janjang Panen											
			Jan	Feb	Mar	Apr	Mei	Jun	Jul	Agu	Sep	Okt	Nov	Des
1	1	29,37	3	3	3	3	3	4	4	5	4	4	4	3
2	2	17,71	3	3	3	1	1	2	5	4	2	5	3	4
3	1	10,05	4	3	2	2	1	1	4	2	1	2	1	2



### 3. Mencari Nilai S

Nilai S diperoleh dengan cara mengurangi nilai *Cost* pada iterasi yang baru kepada iterasi awal. Jika nilai S < 0 maka pengolahan diteruskan dengan menggunakan nilai pusat *medoids* baru. Jika nilai S > 0 atau nilai *Cost* iterasi baru lebih besar dari pada nilai *Cost* iterasi lama maka proses dihentikan. Sehingga nilai S diperoleh :

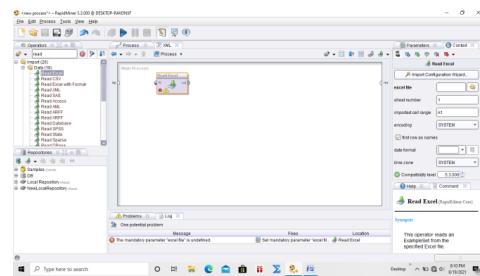
$$\begin{aligned} S &= \text{Nilai } Cost \text{ Baru} - \text{Nilai } Cost \text{ Lama} = 167,8677198 - 163,0791384 \quad (4) \\ &\qquad\qquad\qquad = 47.885.814 \end{aligned}$$

Karena Nilai *Cost* Baru > Nilai *Cost* Lama, maka iterasi dihentikan dan hasil *cluster* terletak di iterasi 1.

### 3.2. Implementasi pada aplikasi *rapidminer*

#### a. Memasukan (input) sistem

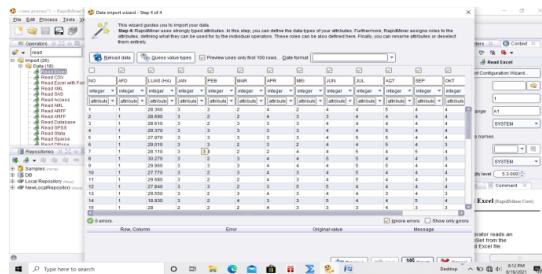
Dalam menjalankan proses clustering, terlebih dahulu import data yang berekstensi .xlsx menggunakan read excel pada tools RapidMiner. Proses dapat dilihat pada gambar 2. berikut ini:



**Gambar 2.** New Process untuk Import Data Excel

#### b. Import configurasi wizard

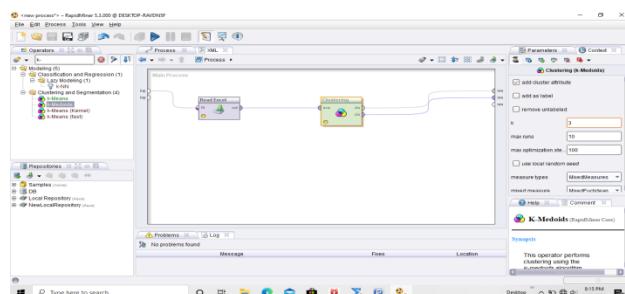
Step ini dilakukan langsung ke data import wizard step 4 dengan cara klik Next maka akan muncul form data import wizard step 4



**Gambar 3.** Process Import Data

#### c. Setelah konfigurasi wizard, pilih file sesuaikan dengan atribut dan tipe data

Pada saat import data selesai, maka selanjutnya drag and drop operator k-medoids ke dalam Main process. Hubungkan read excel dengan operator k-medoids dengan cara menarik garis diantara data dengan operator excel. Dapat dilihat pada gambar 4. berikut ini:

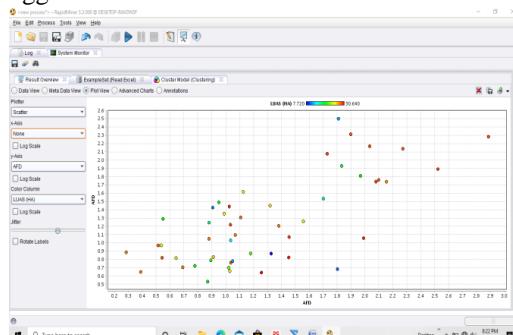


**Gambar 4.** Drag and Drop Operator K-Medoids



#### d. Keluaran (Output) Sistem

Pada tahap ini akan menampilkan hasil akhir serta langkah terakhir dalam penggunaan tools RapidMiner 5.3 ini. Hasil akhir yang akan ditampilkan adalah berupa pengelompokan dimana hasil dari pengujian data akan tampak cluster dengan masing-masing anggota.



Gambar 5. Hasil Pengelompokan Data

### 3.3. Pembahasan

Berdasarkan penjelasan diatas mengenai penggunaan serta hasil yang telah di tampilkan, maka pembahasan ini mengenai keterkaitan dari hasil yang diperoleh antara perhitungan algoritma dengan hasil pengujian pada *tools RapidMiner 5.3*.

#### 3.3.1 Validasi Data

Dalam melakukan validasi data perhitungan algoritma harus menghasilkan hasil akhir berupa pengelompokan dengan 3 *Cluster*, serta data yang digunakan merupakan data yang valid dan sama dengan yang di pakai pada *tools RapidMiner 5.3*. Berikut ditampilkan hasil yang didapatkan dari perhitungan algoritma dan pengujian pada *RapidMiner 5.3*.

Tabel 4. Perbandingan Hasil Akhir *K-Medoids*

No	Afd	Luas (Ha)	Perhitungan Algoritma <i>K-Medoids</i>	Pengujian Dengan <i>Rapid Miner 5.3</i>
1.	1	28,36	0	cluster_0
2.	1	28,69	0	cluster_0
3.	1	28,61	0	cluster_0
4.	1	29,37	0	cluster_0
5.	1	27,97	0	cluster_0
6.	1	29,01	0	cluster_0
7.	1	28,11	0	cluster_0
8.	1	30,27	0	cluster_0
9.	1	29,96	0	cluster_0
10.	1	27,77	0	cluster_0
11.	1	29,68	0	cluster_0
12.	1	27,84	0	cluster_0
13.	1	29,55	0	cluster_0
14.	1	18,93	2	cluster_2
15.	1	28,00	0	cluster_0
16.	1	20,68	2	cluster_2
17.	1	30,64	0	cluster_0
18.	1	27,70	0	cluster_0
19.	1	15,78	1	cluster_1
20.	1	11,24	1	cluster_1
21.	1	18,97	2	cluster_2
22.	1	19,17	2	cluster_2
23.	1	18,72	2	cluster_2
24.	1	21,09	2	cluster_2
25.	1	24,89	0	cluster_0



26.	1	25,19	0	cluster_0
27.	1	25,67	0	cluster_0
28.	1	24,13	0	cluster_0
29.	1	7,72	1	cluster_1
30.	1	24,57	2	cluster_2
31.	1	24,52	0	cluster_0
32.	1	25,97	0	cluster_0
33.	1	25,27	0	cluster_0
34.	1	21,93	2	cluster_2
35.	1	10,05	1	cluster_1
36.	1	13,21	1	cluster_1
37.	1	10,41	1	cluster_1
38.	2	27,33	0	cluster_0
39.	2	28,89	0	cluster_0
40.	2	29,11	0	cluster_0
41.	2	28,66	0	cluster_0
42.	2	29,42	0	cluster_0
43.	2	28,96	0	cluster_0
44.	2	29,29	0	cluster_0
45.	2	28,77	0	cluster_0
46.	2	28,91	0	cluster_0
47.	2	11,34	1	cluster_1
48.	2	15,69	1	cluster_1
49.	2	17,71	2	cluster_2
50.	2	18,97	2	cluster_2

## 4. KESIMPULAN

Perhitungan algoritma *K-Medoids* dengan pengujian *tools RapidMiner 5.3*. diperoleh hasil yang sama dengan perhitungan manual. Hasil *clustering* yang diperoleh, dimana *cluster 0* terdiri dari 32 data, *cluster 1* terdiri dari 8 data, dan *cluster 2* terdiri dari 10 data. Pengelompokan data janjang panen kelapa sawit menggunakan algoritma *k-medoids* menunjukan bahwa data janjang panen pada PT SIR MANDAU pada tahun 2020 berhasil mengalami peningkatan panen yang tinggi.

## REFERENCES

- [1] G. Abdillah *et al.*, “Penerapan Data Mining Pemakaian Air Pelanggan Untuk Menentukan Klasifikasi Potensi Pemakaian Air Pelanggan Baru Di Pdam Tirta Raharja Menggunakan Algoritma K-Means,” *Sentika* 2016, vol. 2016, no. Sentika, pp. 18–19, 2016.
- [2] Y. Mardi, “Data Mining : Klasifikasi Menggunakan Algoritma C4.5,” *Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2017, doi: 10.22202/ei.2016.v2i2.1465.
- [3] 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.
- [4] D. F. Pramesti, Lahan, 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.
- [5] N. Pulungan, S. Suhada, and D. Suhendro, “Penerapan Algoritma K-Medoids Untuk Mengelompokkan Penduduk 15 Tahun Keatas Menurut Lapangan Pekerjaan Utama,” *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, pp. 329–334, 2019, doi: 10.30865/komik.v3i1.1609.
- [6] S. Santoso and R. Nurmalina, “Perencanaan dan Pengembangan Aplikasi Absensi Mahasiswa Menggunakan Smart Card Guna Pengembangan Kampus Cerdas (Studi Kasus Politeknik Negeri Tanah Laut),” *J. Integr.*, vol. 9, no. 1, pp. 84–91, 2017.
- [7] M. Silalahi, “Analisis Clustering Menggunakan Algoritma K-Means Terhadap Penjualan Produk Padapt Batamas Niaga Jaya,” *Comput. Based Inf. Syst. J.*, vol. 02, pp. 20–35, 2018.



- [8] S. Sindi, W. R. O. Ningse, I. A. Sihombing, F. Ilmi R.H.Zer, and D. Hartama, “Analisis Algoritma K-Medoids Clustering Dalam Pengelompokan Penyebaran Covid-19 Di Indonesia,” *Jti (Jurnal Teknol. Informasi)*, vol. 4, no. 1, pp. 166–173, 2020.
- [9] S. Sundari, I. S. Damanik, A. P. Windarto, H. S. Tambunan, J. Jalaluddin, and A. Wanto, “Analisis K-Medoids Clustering Dalam Pengelompokkan Data Imunisasi Campak Balita di Indonesia,” *Pros. Semin. Nas. Ris. Inf. Sci.*, vol. 1, no. September, p. 687, 2019, doi: 10.30645/senaris.v1i0.75.
- [10] D. Ardiansyah and W. Walim, “Algoritma c4.5 untuk klasifikasi calon peserta lomba cerdas cermat siswa smp dengan menggunakan aplikasi rapid miner,” *J. Inkofar*, vol. 1, no. 2, pp. 5–12, 2018.