

Penerapan Algoritma K-Means Dalam Mengclustering Kualitas Bibit Kelapa Sawit Di PPKS Marihat

Rizal Efendi Marpaung^{1,*}, Jaya Tata Hardinata², Rizki Alfadillah Nasution³

^{1,2,3}Program Studi Sistem Informasi, STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

Email: ^{1*}rizalefendi5482@gmail.com, ²jayatatahardinata@amiktunasbangsa.ac.id, ³rizkialfadillah@amiktunasbangsa.ac.id

(* : rizalefendi5482@gmail.com)

Abstract

Grouping the quality of oil palm seeds needs to be done to find out the groups of types of oil palm seeds that are of good and bad quality so that the process of planting oil palm can be done well too. This study discusses classifying the quality of oil palm seeds by using the K-Means Clustering algorithm as a case resolution solution. The source of data in this study was obtained from the Marihat Oil Palm Research Center (PPKS) with 23 data processing data. The analysis in this study uses 2 (two) cluster levels, namely the good cluster (C1) and the bad cluster (C2). The results obtained are the good clusters are at the age (5-25 months), and the bad clusters are at the age (3-4 months). It is hoped that the results of the research can be input, suggestions, and efforts for PPKS Marihat to improve the quality of oil palm seeds so that they can support a more optimal strategy for producing oil palm seeds.

Keywords: Data mining, K Means, Clustering, Seeds, Oil Palm

Abstrak

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. Penelitian ini membahas tentang mengelompokkan kualitas bibit kelapa sawit dengan menggunakan algoritma K-Means Clustering sebagai solusi penyelesaian kasus. Sumber data pada penelitian ini diperoleh dari Pusat Penelitian Kelapa Sawit (PPKS) Marihat dengan data olah sebanyak 23 data. Analisa pada penelitian ini menggunakan 2 (dua) tingkat cluster yaitu cluster baik (C1) dan cluster buruk (C2). Hasil penelitian yang diperoleh yaitu cluster baik berada di umur (5-25 bulan), dan untuk cluster buruk berada di umur (3-4 bulan). Diharapkan hasil penelitian dapat menjadi masukan, saran, dan upaya pada PPKS Marihat untuk meningkatkan kualitas bibit kelapa sawit sehingga dapat menunjang strategi produksi bibit kelapa sawit yang lebih optimal.

Kata Kunci: Data mining, K Means, Clustering, Bibit, Kelapa Sawit

1. PENDAHULUAN

Data Mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database [1]-[2]. Data Mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk menginteraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar [3], [4]. Definisi umum dari Data Mining itu sendiri adalah proses pencarian pola – pola yang tersembunyi (hidden pattern) berupa pengetahuan (konowlegde) yang diketahui sebelumnya dari suatu sekumpulan data yang mana data tersebut dapat berada didalam database, data warehouse, atau media penyimpanan informasi yang lain [5], [6]. Clustering adalah metode penganalisis data, yang sering dimasukkan sebagai salah satu metode Data Mining, yang bertujuan adalah untuk mengelompokkan data dengan karakteristik yang sama ke suatu ‘wilayah’ yang sama dan data dengan karakteristik yang berbeda ‘wilayah’ yang lain [7], [8]. Ada beberapa pendekatan yang digunakan dalam pengembangan metode clustering. Dua pendekatan utama adalah clustering dengan pendekatan partisi dan clustering dengan pendekatan hirarki. Salah satu algoritma dalam kategori clustering adalah K-Means [9]. Perkebunan sebagai salah satu sub sektor pertanian di Indonesia berpeluang besar dalam peningkatan perekonomian rakyat dan pembangunan perekonomian nasional. Adanya dukungan sumber daya dan orientasi pembangunan perekonomian yang cukup besar menjadikan sub sektor ini sebagai penggerak pembangunan nasional [10], [11]. Kelapa sawit merupakan komoditi perkebunan andalan sebagai sumber minyak nabati dan bahan agroindustri sehingga menjadi penghasil devisa negara di luar minyak dan gas bumi [12]. Tanaman kelapa sawit memiliki daya tarik untuk dikelola dan ditanam, baik oleh pihak perkebunan negara, perkebunan swasta milik nasional maupun asing, bahkan petani dalam skala perkebunan rakyat [13].

Rendahnya produktivitas dan mutu produksi masih menjadi salah satu masalah dalam pengembangan kelapa sawit nasional. Rendahnya produktivitas perkebunan kelapa sawit nasional khususnya perkebunan rakyat, ditengarai karena teknologi produksi yang diterapkan masih relatif sederhana, mulai dari tahapan pembibitan sampai panen. Produksi minyak sawit Indonesia telah menyumbang 2 sekitar 30% dari produksi minyak nabati dunia dan ekspor minyak sawit Indonesia mencapai 60% dari permintaan pasar global. Peningkatan jumlah produksi kelapa sawit sejalan dengan meningkatnya permintaan global terhadap bahan mentah maupun hasil olahan kelapa sawit itu sendiri. Hal ini sesuai dengan pernyataan Corley (2009) bahwa dengan bertambahnya kebutuhan dan produksi kelapa sawit saat ini, konsumsi minyak kelapa sawit dunia diperkirakan akan mencapai 256 juta ton Crude Palm Oil (CPO) pada tahun 2050 (Iskandar, 2018).

Pusat Penelitian Kelapa Sawit (PPKS) Marihat Pematangsiantar merupakan salah satu perusahaan yang melakukan penelitian terhadap kecambah kelapa sawit untuk mendapatkan bibit kelapa sawit yang unggul. Dalam

mendukung pelaksanaan tugasnya sebagai lembaga riset. Kelapa sawit merupakan tanaman perkebunan yang menjadi faktor penting dalam perekonomian rakyat, penyerapan tenaga kerja dan sumber devisa bagi negara [14]. Permintaan kelapa sawit terus mengalami peningkatan sehingga menyebabkan produksi dan perluasan areal pertanaman kelapa sawit juga semakin meningkat. Dengan bertambahnya luas areal pertanaman kelapa sawit tersebut maka diperlukan pengadaan bibit dalam jumlah besar dan berkualitas. Dalam usaha membudidayakan kelapa sawit, masalah lain yang dihadapi oleh pengusaha atau petani adalah pengadaan bibit. Kualitas bibit yang sehat, berpotensi unggul dengan produksi yang tepat waktu dan tinggi, menjadi salah satu parameter dalam melihat kualitas komoditi ini.

2. METODOLOGI PENELITIAN

2.1. Metode Penelitian

Metode penelitian merupakan langkah yang dimiliki dan dilakukan oleh peneliti dalam rangka untuk mengumpulkan informasi atau data serta melakukan investigasi pada data yang telah didapatkan. Adapun metode penyelesaian yang digunakan dalam penelitian ini yaitu metode Data Mining dengan algoritma K-Means.

2.1.1. Metode Pengumpulan Data

Dalam melakukan penelitian untuk mendapatkan data dan informasi, maka metode yang digunakan dalam proses pengumpulan data sebagai berikut:

1. Metode Observasi, melihat serta mempelajari permasalahan yang ada di lapangan yang berkaitan dengan objek yang diteliti.
2. Metode Kajian Literatur, mencari bahan yang mendukung dalam pendefinisian masalah melalui buku-buku, makalah, *internet*, yang erat kaitannya dengan objek permasalahan.

2.1.2. Pengolahan Data

Analisis data yang digunakan penelitian ini menggunakan data kuantitatif dengan teknik analisis data yang menggunakan jenis statistik *deskriptif data*.

Tabel 1. Data Penelitian

Umur Jumlah Tandan	3 Bulan	4 Bulan	5 Bulan	22 Bulan	23 Bulan	24 Bulan	25 Bulan
RBT	S1	14	27	19	8	7	7
	S2	12	25	18	8	7	7
	S3	10	22	16	7	6	6
TBS	S1	5	6	12	27	28	28
	S2	4	6	12	26	27	28
	S3	4	5	11	26	27	29
TBS	S1	75	179	243	232	226	225
	S2	62	156	218	214	207	210
	S3	50	134	192	195	188	190
CPO	S1	10	24	33	31	30	30
	S2	8	20	28	27	26	26
	S3	6	17	24	2	24	24
PKO	S1	2,5	6,5	8,8	8,4	8,2	8,1
	S2	2	5,4	7,5	7,4	7,2	7,1
	S3	1,6	4,6	6,6	6,7	6,5	6,6
	S1	0,27	0,65	0,88	0,84	0,81	0,81
	S2	0,21	0,54	0,75	0,73	0,71	0,72
	S3	0,17	0,46	0,66	0,67	0,65	0,66

Keterangan variable Tabel 1

1. RBT : Rata Rata Berat Tandan
2. TBS 1 : Tandan Buah Segar 1
3. TBS 2 : Tandan Buah Segar 2
4. CPO : *Crude Palm Oil* (Sumber Bahan Baku)
5. PKO : *Palm Kernel Oil* (Hasil Daging Buah Sawit)

3.2. Analisis dan Perancangan Sistem

Di dalam analisis dan metode perancangan sistem terdapat rancangan yang digunakan dalam membangun sistem atau perhitungan-perhitungan algoritma yang digunakan.

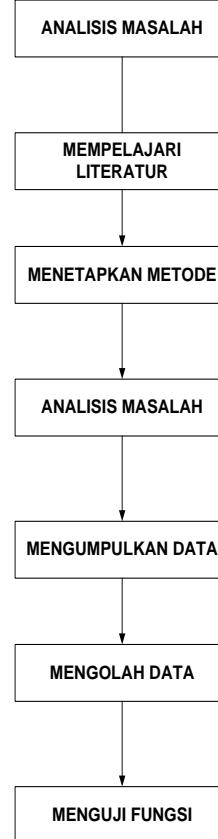
3.2.1. Analisa Permasalahan

Pusat Penelitian Kelapa Sawit (PPKS) Marihat Pematangsiantar merupakan salah satu perusahaan yang melakukan penelitian terhadap kecambah kelapa sawit untuk mendapatkan bibit kelapa sawit yang unggul. Dalam mendukung pelaksanaan tugasnya sebagai lembaga riset. Kelapa sawit merupakan tanaman perkebunan yang menjadi

faktor penting dalam perekonomian rakyat, penyerapan tenaga kerja dan sumber devisa bagi Negara. Permintaan kelapa sawit terus mengalami peningkatan sehingga menyebabkan produksi dan perlakuan areal pertanaman kelapa sawit juga semakin meningkat. Dengan bertambahnya luas areal pertanaman kelapa sawit tersebut maka diperlukan pengadaan bibit dalam jumlah besar dan berkualitas. Dalam usaha membudidayakan kelapa sawit, masalah lain yang dihadapi oleh pengusaha atau petani adalah pengadaan bibit. Kualitas bibit yang sehat, berpotensi unggul dengan produksi yang tepat waktu dan tinggi, menjadi salah satu parameter dalam melihat kualitas bibit kelapa sawit tersebut.

3.2.2. Perancangan Penelitian

Berikut ini perancangan penelitian dapat dilihat pada gambar 1 dibawah ini.



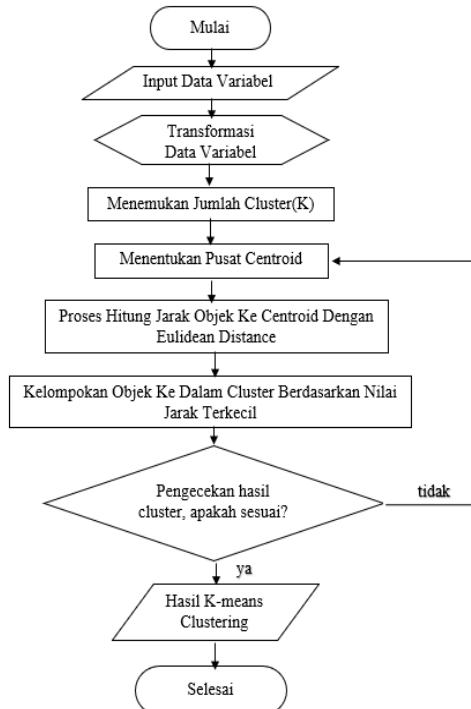
Gambar 1. Kerangka Penelitian

Diagram alir model penelitian seperti yang disajikan pada Gambar 1 menjelaskan rancangan penelitian yang dilakukan untuk menentukan lokasi dagangan pasar dengan menggunakan metode *K-Means* yang terdiri dari :

- Analisis Masalah
Menganalisis masalah yang terkait dengan mengelompokkan kualitas biji kelapa sawit dan menentukan parameter atau variabel apa saja yang digunakan.
- Mempelajari Literatur
Penelitian ini harus didasari rujukan yang digunakan untuk memperoleh informasi dalam penelitian.
- Menetapkan Metode
Menetapkan metode untuk memecahkan masalah. Pada penelitian ini metode yang digunakan yaitu *Data Mining* dengan algoritma *K-means*.
- Mengumpulkan Data
Pengumpulan data dalam penelitian ini dilakukan dengan mengambil data yang telah ada di instansi terkait dan melakukan wawancara dengan pihak PPKS Marihat Pematangsiantar selama 1 minggu yang dimulai pada 10 April 2021 sampai dengan 16 April 2021.
- Mengolah Data
Melakukan pengelolahan data dengan menggunakan algoritma *K-Means*. Pengelolahan data disini menggunakan Ms.Excel 2010.
- Menguji data
Pengujian data dilakukan dengan menggunakan aplikasi *RapidMiner*.
- Kesimpulan
Kesimpulan yang didapatkan dalam mengelompokkan kualitas bibit kelapa sawit. Jika diperoleh hasil kualitas rendah lebih banyak maka dapat menjadi informasi bagian pihak PPKS Marihat agar dapat mengelola kembali bibit kelapa sawit yang masih kualitas rendah.

3.2.3. Algoritma K-Means

Alur kerja proses pemodelan dan tahapan-tahapan algoritma *K-Means Clustering* dapat dilihat pada gambar 2 [15]:



Gambar 2. Flowchart Tahapan Algoritma K-Means

Permasalahan yang sudah dianalisa, selanjutnya penulis menganalisa data menggunakan *Data Mining* algoritma *K-Means Clustering*. Algoritma *K-Means* selanjutnya akan melakukan pengulangan langkah-langkah berikut sampai terjadi kestabilan :

- Tentukan berapa jumlah klaster yang ingin ditetapkan pusat *cluster* k.
- Menggunakan jarak *euclidean* dan kemudian menghitung setiap data ke pusat *cluster*.
- Kategorikan data ke dalam cluster dengan jarak yang terpendek dengan menggunakan persamaan
- Menghitung pusat *cluster* dengan menggunakan persamaan
- Silahkan ulangi langkah dua sampai empat sehingga sudah tidak ada lagi data yang berpindah ke kluster yang lain.

3. HASIL DAN PEMBAHASAN

3.1. Hasil

Hasil dan pembahasan pada bab ini disajikan sesuai penelitian yang dilakukan. Data yang digunakan dalam penelitian ini adalah data bibit kelapa sawit. Pada penelitian ini digunakan tools RapidMiner sebagai pengujian data yang telah di analisa. Berikut adalah bagian-bagian dari RapidMiner beserta isi dari data yang dibuat.

3.1.1. Perhitungan Menggunakan Algoritma K-Means

Data penelitian terdiri dari 23 Bulan umur dari bibit kelapa sawit dan digunakan data dari tahun 6 kriteria. Kualitas bibit kelapa sawit dapat dikelompokkan kedalam 2 cluster yaitu cluster baik dan tidak baik. Berikut adalah data penelitian yang digunakan. Berikut adalah data penelitian yang digunakan. Dapat dilihat pada tabel 3. berikut.

Tabel 2. Data Penelitian

Umur	Jumlah Tandan	RBT	TBS 1	TBS 2	CPO	PKO
3 Bulan	36	13	487	24	6,1	0,65
4 Bulan	58	17	369	61	16,5	1,65
5 Bulan	53	35	653	85	22,9	2,29
...
23 Bulan	20	82	621	80	21,9	2,17
24 Bulan	19	83	625	81	21,9	2,18
25 Bulan	20	87	618	80	21,7	2,17

Selanjutnya masuk dalam tahap perhitungan menggunakan metode *k-means*:

1. Penetuan Pusat *Cluster* Awal

Penetuan ini menggunakan 2 *cluster* (pengelompokan) ini dilakukan dengan mengambil nilai terbesar (maksimum) untuk *cluster* baik (C1), nilai rata-rata (*Average*) untuk *cluster* sedang, dan nilai terkecil untuk *cluster* rendah (C2) sebagai berikut:

Tabel 3. Nilai *Cluster*

Cluster	Jumlah Tandan	RBT	TBS 1	TBS 2	CPO	PKO
Max (C1)	58	87	719	93	25,6	2,51
Min (C2)	19	13	369	24	6,1	0,65

2. Perhitungan Jarak *Cluster*

Untuk menghitung jarak antara data dengan pusat *cluster* memakai persamaan (1):

$$d_{ik} = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2} \quad (1)$$

a. *Cluster* Baik

$$C1,3 = \sqrt{(36 - 58)^2 + (13 - 87)^2 + (487 - 719)^2 + (24 - 93)^2 + (6,1 - 25,6)^2 + (0,65 - 2,51)^2} = 254,81$$

$$C1,4 = \sqrt{(58 - 58)^2 + (17 - 87)^2 + (369 - 719)^2 + (61 - 93)^2 + (16,5 - 25,6)^2 + (1,65 - 2,51)^2} = 358,48$$

$$C1,5 = \sqrt{(53 - 58)^2 + (35 - 87)^2 + (653 - 719)^2 + (85 - 93)^2 + (22,9 - 25,6)^2 + (2,29 - 2,51)^2} = 84,60$$

$$C1,6 = \sqrt{(51 - 58)^2 + (38 - 87)^2 + (684 - 719)^2 + (88 - 93)^2 + (25,1 - 25,6)^2 + (2,39 - 2,51)^2} = 60,83$$

$$C1,7 = \sqrt{(42 - 58)^2 + (43 - 87)^2 + (702 - 719)^2 + (91 - 93)^2 + (24,7 - 25,6)^2 + (2,45 - 2,51)^2} = 49,86$$

$$C1,8 = \sqrt{(43 - 58)^2 + (45 - 87)^2 + (692 - 719)^2 + (91 - 93)^2 + (25,6 - 25,6)^2 + (2,42 - 2,51)^2} = 52,17$$

$$C1,9 = \sqrt{(39 - 58)^2 + (48 - 87)^2 + (691 - 719)^2 + (90 - 93)^2 + (25,6 - 25,6)^2 + (2,41 - 2,51)^2} = 51,72$$

$$C1,10 = \sqrt{(41 - 58)^2 + (48 - 87)^2 + (692 - 719)^2 + (90 - 93)^2 + (24,7 - 25,6)^2 + (2,42 - 2,51)^2} = 50,49$$

Dan begitu seterusnya hingga diperoleh hasil dari cluster baik dan cluster buruk untuk setiap alternatif. Berikut adalah hasil dari perhitungan jarak cluster pada iterasi pertama:

Tabel 4. Hasil Iterasi 1

Umur	Baik	Tidak Baik	Jarak	C1	C2
3 Bulan	254,81	119,22	119,22		1
4 Bulan	358,48	54,91	54,91		1
5 Bulan	84,60	293,77	84,60	1	
....
22 Bulan	92,04	282,71	92,04	1	
23 Bulan	106,09	267,68	106,09	1	
24 Bulan	102,62	271,91	102,62	1	
25 Bulan	108,76	266,19	108,76	1	

Berdasarkan matrik yang didapatkan pada tabel di atas maka didapatkan pengelompokan sebagai berikut :

$$C1 = 5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25$$

$$C2 = 1,2$$

Setelah didapatkan hasil dari setiap *cluster* kemudian pusat *cluster* baru dihitung berdasarkan data member tiap – tiap *cluster* yang sudah didapatkan dengan menjumlahkan setiap member dari tiap cluster dan dibagi dengan jumlah dari setiap cluster yang diperoleh. Berikut adalah pusat cluster baru yang diperoleh dan akan digunakan dalam perhitungan *cluster* 2 :

Tabel 5. Pusat Cluster Baru

Cluster	Jumlah Tandan	RBT	TBS 1	TBS 2	CPO	PKO
Max (C1)	32,33	62,67	679,24	87,00	23,86	2,37
Min (C2)	47	15	428	42,5	11,3	1,15

Iterasi selanjutnya dilakukan dengan cara yang sama hingga tidak ada perubahan data dalam suatu cluster

a. *Cluster Baik*

$$\begin{aligned} C1,3 &= \sqrt{(36 - 32,33)^2 + (13 - 62,67)^2 + \\ &\quad (487 - 679,24)^2 + (24 - 87,00)^2} = 209,1008 \\ C1,4 &= \sqrt{(58 - 32,33)^2 + (17 - 62,67)^2 + \\ &\quad (369 - 679,24)^2 + (61 - 87,00)^2} = 315,7889 \\ C1,5 &= \sqrt{(53 - 32,33)^2 + (35 - 62,67)^2 + \\ &\quad (653 - 679,24)^2 + (85 - 87,00)^2} = 43,42724 \\ C1,6 &= \sqrt{(51 - 32,33)^2 + (38 - 62,67)^2 + \\ &\quad (684 - 679,24)^2 + (88 - 87,00)^2} = 31,33844 \\ C1,7 &= \sqrt{(42 - 32,33)^2 + (43 - 62,67)^2 + \\ &\quad (702 - 679,24)^2 + (91 - 87,00)^2} = 31,85961 \\ C1,8 &= \sqrt{(43 - 32,33)^2 + (45 - 62,67)^2 + \\ &\quad (692 - 679,24)^2 + (91 - 87,00)^2} = 24,65316 \\ C1,9 &= \sqrt{(39 - 32,33)^2 + (48 - 62,67)^2 + \\ &\quad (691 - 679,24)^2 + (90 - 87,00)^2} = 20,24649 \\ C1,10 &= \sqrt{(41 - 32,33)^2 + (48 - 62,67)^2 + \\ &\quad (692 - 679,24)^2 + (90 - 87,00)^2} = 21,51263 \end{aligned}$$

Berikut adalah hasil pengelompokan dari iterasi 2 dapat dilihat pada Tabel 6:

Tabel 6. Hasil Iterasi 2

Umur	Baik	Tidak Baik	Jarak	C1	C2
3 Bulan	209,1008	63,05188	63,05188	1	
4 Bulan	315,7889	63,05188	63,05188	1	
5 Bulan	43,42724	230,224	43,42724	1	
....
22 Bulan	50,46695	224,6639	50,46695	1	
23 Bulan	63,01155	209,7299	63,01155	1	
24 Bulan	59,77357	214,0389	59,77357	1	
25 Bulan	67,43917	208,6347	67,43917	1	

Berdasarkan matrik yang didapatkan pada tabel di atas maka didapatkan pengelompokan sebagai berikut :

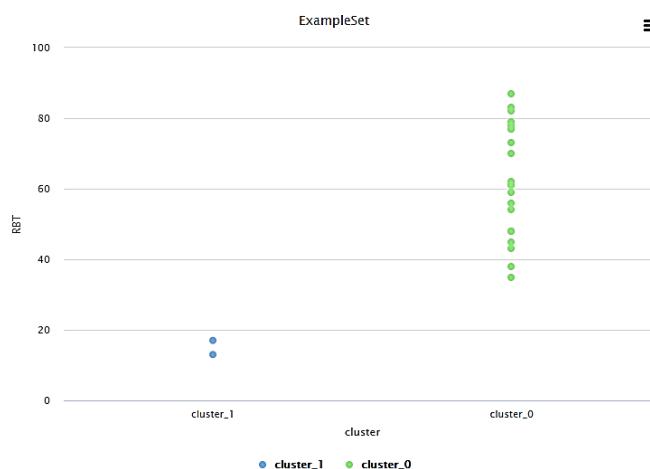
$$C1 = 5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25$$

$$C2 = 1,2$$

Dari hasil iterasi 2 dapat dilihat bahwa hasil yang diperoleh sama dengan iterasi 1 maka perhitungan dihentikan sampai iterasi 2. Untuk hasil pengelompokan dari kualitas bibit kelapa sawit pada cluster baik berada di umur (5-25 bulan), dan untuk cluster buruk berada di umur (3-4 bulan).

3.2. Implementasi

Setelah melakukan proses sistem seperti yang tertera pada gambar 4. maka akan muncul hasil dari RapidMiner sebagai berikut:



Gambar 3. Grafik Hail Cluster

Berdasarkan Gambar 3 dapat dilihat bahwa cluster 0 adalah cluster rendah dan cluster 1 adalah cluster tinggi. Dari Gambar 4 juga dapat dilihat bahwa terdapat 1 alternatif yang berada pada cluster baik dan 7 alternatif pada cluster sedang dan 4 alternatif pada cluster buruk seperti yang tertera pada Gambar 4.

Cluster Model

```
Cluster 0: 21 items
Cluster 1: 2 items
Total number of items: 23
```

Gambar 4. Cluster Model

Data di atas akan diuji dengan sampel sebanyak 4 tahun dan diselesaikan dengan menggunakan tools RapidMiner sehingga diperoleh hasil cluster seperti pada Gambar 5 berikut.

Row No.	umur	cluster	jumlah tandan	RBT	TBS 1	TBS 2	CPO	PKO
1	3	cluster_1	36	13	487	24	6.100	0.650
2	4	cluster_1	58	17	369	61	16.500	1.650
3	5	cluster_0	53	35	653	85	22.900	2.290
4	6	cluster_0	51	38	684	88	25.100	2.390
5	7	cluster_0	42	43	702	91	24.700	2.450
6	8	cluster_0	43	45	692	91	25.600	2.420
7	9	cluster_0	39	48	691	90	25.600	2.410
8	10	cluster_0	41	48	692	90	24.700	2.420
9	11	cluster_0	36	54	703	91	24.900	2.460
10	12	cluster_0	34	56	707	92	23.700	2.470
11	13	cluster_0	34	59	717	93	24.100	2.510
12	14	cluster_0	34	61	719	93	23	2.510
13	15	cluster_0	33	61	715	93	25.200	2.490
14	16	cluster_0	32	62	710	92	24.900	2.490

Gambar 5. Hasil Cluster

Pada Gambar 5 dapat dilihat bahwa hasil dari perhitungan manual adalah sama dengan hasil perhitungan menggunakan tools *RapidMiner*. 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 dipakai pada tools *RapidMiner*. Berikut ditampilkan hasil yang didapatkan dari perhitungan algoritma dan pengujian pada *RapidMiner*.

Tabel 7. Perbandingan Hasil Manual dan *RapidMiner*

No	Bulan	Hasil Perhitungan Manual	Hasil RapidMiner
1	3 Bulan	cluster_1	cluster_1
2	4 Bulan	cluster_1	cluster_1
3	5 Bulan	cluster_0	cluster_0
4	6 Bulan	cluster_0	cluster_0
5	7 Bulan	cluster_0	cluster_0
6	8 Bulan	cluster_0	cluster_0
7	9 Bulan	cluster_0	cluster_0
8	10 Bulan	cluster_0	cluster_0
9	11 Bulan	cluster_0	cluster_0
10	12 Bulan	cluster_0	cluster_0
11	13 Bulan	cluster_0	cluster_0
12	14 Bulan	cluster_0	cluster_0
13	15 Bulan	cluster_0	cluster_0
14	16 Bulan	cluster_0	cluster_0
15	17 Bulan	cluster_0	cluster_0
16	18 Bulan	cluster_0	cluster_0
17	19 Bulan	cluster_0	cluster_0
18	20 Bulan	cluster_0	cluster_0
19	21 Bulan	cluster_0	cluster_0
20	22 Bulan	cluster_0	cluster_0
21	23 Bulan	cluster_0	cluster_0
22	24 Bulan	cluster_0	cluster_0
23	25 Bulan	cluster_0	cluster_0

Berdasarkan matrik yang didapatkan pada tabel di atas maka didapatkan pengelompokan sebagai berikut :

$$C1 = 5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25$$

$$C2 = 1,2$$

Dari hasil iterasi 2 dapat dilihat bahwa hasil yang diperoleh sama dengan iterasi 1 maka perhitungan dihentikan sampai iterasi 2. Untuk hasil pengelompokan dari kualitas bibit kelapa sawit pada cluster baik berada di umur (5-25 bulan), dan untuk cluster buruk berada di umur (3-4 bulan).

4. KESIMPULAN

Kesimpulan yang didapat dari hasil literasi 2 dapat dilihat bahwa hasil yang diperoleh sama dengan iterasi 1 maka perhitungan dihentikan sampai iterasi 2. Untuk hasil pengelompokan dari kualitas bibit kelapa sawit pada cluster baik berada di umur (5-25 bulan), dan untuk cluster buruk berada di umur (3-4 bulan). Cluster1=5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25 Cluster2 = 1,2. Hasil antara perhitungan yang dilakukan penulis dengan hasil yang diperoleh tools *RapidMiner* memiliki kesamaan hasil pengelompokan pengelompokan kualitas bibit sawit pada PPKS Marihat. Untuk hasil pengelompokan dari kualitas bibit kelapa sawit pada cluster baik berada di umur (5-25 bulan), dan untuk cluster buruk berada di umur (3-4 bulan). Berdasarkan hasil yang diperoleh pihak PPKS Marihat Pematangsiantar dapat memperoleh informasi baru berupa data kualitas bibit kelapa sawit dengan pengelompok baik dan kurang baik. Informasi yang dihasilkan dapat menjadi acuan untuk teknik pengiriman produksi, peningkatan pelayanan dan fasilitas, serta memenuhi kebutuhan dari PPKS Marihat Pematangsiantar.

REFERENCES

- [1] L. Felicia, "Penerapan Metode Clustering Dengan K-Means Untuk Memetakan Potensi Tanaman Padi Di Kota Semarang," Pp. 1–5, 2014.
- [2] R. Rosmini, A. Fadlil, And S. Sunardi, "Implementasi Metode K-Means Dalam Pemetaan Kelompok Mahasiswa Melalui Data Aktivitas Kuliah," *It J. Res. Dev.*, Vol. 3, No. 1, P. 22, 2018, Doi: 10.25299/Itjrd.2018.Vol3(1).1773.
- [3] A. P. Windarto, J. Naam, Y. Yuhandri, A. Wanto, And M. Mesran, "Bagian 2: Model Arsitektur Neural Network Dengan Kombinasi K-Medoids Dan Backpropagation Pada Kasus Pandemi Covid-19 Di Indonesia," *J. Media Inform. Budidarma*, Vol. 4, No. 4, Pp. 1175–1180, 2020, Doi: 10.30865/Mib.V4i4.2505.

- [4] A. P. Windarto, U. Indriani, M. R. Raharjo, And L. S. Dewi, "Bagian I: Kombinasi Metode Klastering Dan Klasifikasi (Kasus Pandemi Covid-19 Di Indonesia)," *J. Media Inform. Budidarma*, Vol. 4, No. 3, P. 855, 2020, Doi: 10.30865/Mib.V4i3.2312.
- [5] Mardalius, "Pemanfaatan Rapid Miner Studio 8.2 Untuk Pengelompokan Data Penjualan Aksesoris Menggunakan Algoritma K-Means," Vol. Iv, No. 2, Pp. 401–411, 2018.
- [6] 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.
- [7] F. Nasari And C. J. M. Sianturi, "Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Penyebaran Diare Di Kabupaten Langkat," *Cogito Smart J.*, Vol. 2, No. 2, P. 108, 2018, Doi: 10.31154/Cogito.V2i2.19.108-119.
- [8] F. Nurzaman, "Penerapan Algoritma K-Means Dalam Pengelompokan Lokasi Rumah Sakit Provider Pada Asuransi Kesehatan," Pp. 61–67, 2018.
- [9] W. A. Triyanto, "Algoritma K-Medoids Untuk Penentuan Strategi Pemasaran Produk," *J. Simetris*, Vol. 6, No. 1, Pp. 183–188, 2015.
- [10] L. Iswari And U. I. Indonesia, "Pemanfaatan Algoritma K-Means Untuk Pemetaan Hasil," Vol. 21, No. December, 2016, Doi: 10.20885/Teknoin.Vol21.Iss1.Art7.
- [11] Haviluddin, F. Agus, M. Azhari, And A. S. Ahmar, "Artificial Neural Network Optimized Approach For Improving Spatial Cluster Quality Of Land Value Zone," *Int. J. Eng. Technol.*, Vol. 7, No. 2, Pp. 80–83, 2018, Doi: 10.14419/Ijet.V7i2.2.12738.
- [12] P. Alkhairi, I. S. Damanik, And A. P. Windarto, "Penerapan Jaringan Saraf Tiruan Untuk Mengukur Korelasi Beban Kerja Dosen Terhadap Peningkatan Jumlah Publikasi," *Pros. Semin. Nas. Ris. Inf. Sci.*, Vol. 1, No. September, P. 581, 2019, Doi: 10.30645/Senaris.V1i0.65.
- [13] S. Haryati, A. Sudarsono, And E. Suryana, "Implementasi Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4.5 (Studi Kasus: Universitas Dehasen Bengkulu)," *J. Media Infotama*, Vol. 11, No. 2, Pp. 130–138, 2015.
- [14] M. L. Sibuea And A. Safta, "Pemetaan Siswa Berprestasi Menggunakan Metode K-Means Clustering," *Jurteksi*, Vol. 4, No. 1, Pp. 85–92, 2017, Doi: 10.33330/Jurteksi.V4i1.28.
- [15] S. Takalapeta, "Penerapan Data Mining Untuk Menganalisis Kepuasan Konsumen Menggunakan Metode Algoritma C4.5," *JIM P - J. Inform. Merdeka Pasuruan*, Vol. 3, No. 3, Pp. 34–38, 2018, Doi: 10.37438/Jimp.V3i3.186.