Jurnal Informatika Ekonomi Bisnis

http://www.infeb.org

2020 Vol. 2 No. 2 Hal: 32-38 ISSN: 2714-8491 (electronic)

Optimalisasi Penggunaan Lahan Perkebunan Kelapa Hibrida Menggunakan K-Means Clustering

Henky Andema[1]✉, Sarjon Defit[2], Yuhandri[3]

[1,2,3]Universitas Putra Indonesia YPTK Padang

henkyandema27@gmail.com

Abstract

Plantations are the main source of income for farmers in Indragiri Hilir Regency. This plantation is the plantation sector most widely cultivated by farmers is a coconut plantation. The best grouping of coconut cultivation areas is important in developing farmers' income. This study aims to help the Plantation Office in the process of making the best decision areas for planting coconut, especially hybrid coconut. The data used in this study is the data of hybrid coconut plantations in 2018. Data processing in this study uses the K-Means Clustering method with the number of 3 Clusters namely Cluster 0 (C0) Less Potential, Cluster 1 (C1) Enough Potential, Cluster 2 (C2) Very Potential for planting hybrid coconuts. The results of the clustering process with 2 iterations stated that for Cluster 0 there were 7 village data, for Cluster 1 there were 1 village data, and for Cluster 2 there were 2 village data.

Keywords: Data Mining, K-Means, Clustering, Coconut Plantation, Copra.

Abstrak

Perkebunan merupakan sumber penghasilan utama bagi petani di Kabupaten Indragiri Hilir. Perkebunan ini merupakan sektor perkebunan yang paling banyak dibudidayakan petani adalah perkebunan kelapa. Pengelompokan daerah terbaik penanaman kelapa merupakan hal penting dalam perkembangan penghasilan petani. Penelitian ini bertujuan untuk membantu pihak Dinas Perkebunan dalam proses pengambilan keputusan daerah terbaik penanaman kelapa khususnya kelapa hibrida. Data yang digunakan dalam penelitian adalah data perkebunan kelapa hibrida tahun 2018. Pengolahan data dalam penelitian ini menggunakan metode K-Means Clustering dengan jumlah 3 Cluster yaitu Cluster 0 (C0) Kurang Berpotensi, Cluster 1 (C1) Cukup Berpotensi, Cluster 2 (C2) Sangat Berpotensi untuk penanaman kelapa hibrida. Hasil dari proses clustering dengan 2 kali iterasi menyatakan bahwa untuk Cluster 0 berjumlah 7 data desa, untuk Cluster 1 berjumlah 1 data desa, dan untuk Cluster 2 berjumlah 2 data desa.

Kata kunci: Data Mining, K-Means, Clustering, Perkebunan Kelapa, Kopra.

1. Pendahuluan

Kabupaten Indragiri Hilir merupakan salah satu daerah hasil pemekaran dengan pelaksanaannya terhitung pada tanggal 20 November 1965. Kabupaten Indragiri Hilir berada di Pulau Sumatera yang memiliki luas daratan sekitar 11.605 km^2 dan memiliki luas perairan sekitar 7.207 km^2, dengan data luas daratan yang begitu besar, Kabupaten Indragiri Hilir berpenduduk sekitar 703.734 jiwa pada tahun 2017 yang tercatat pada Badan Pusat Statistik (BPS) Kabupaten Indragiri Hilir.

Secara umum masyarakat Kabupaten Indragiri Hilir berusaha disektor pertanian yaitu sektor perkebunan, pada sektor perkebunan yang paling banyak dibudidayakan petani adalah perkebunan kelapa. Berdasarkan hasil data pada Dinas Perkebunan Kabupaten Indragiri Hilir luas tanaman kelapa mencapai 336,938 hektar pada tahun 2018. Banyaknya perkebunan kelapa yang diusahakan petani menjadikan perkebunan suatu faktor penting untuk pengembangan pertanian baik ditingkat nasional maupun ditingkat regional, hal ini sejalan dengan semakin meluasnya lahan perkebunan, maka dari itu Kabupaten Indragiri Hilir diberi julukan Negeri Hamparan Kelapa Dunia.

Pengoptimalan lahan perkebunan kelapa dengan mengelompokan desa yang memiliki potensi penanaman kelapa khususnya kelapa hibrida dalam penelitian ini menggunakan metode *clustering*, data-data lahan perkebunan diperlukan sebagai acuan untuk melakukan pengelompokan. Data yang sudah dikelompokan dengan menggunakan metode *k-means* diharapkan mempermudah masyarakat mengetahui lahan perkebunan disetiap desa agar masyarakat dapat mengetahui daerah mana yang sangat berpotensi, berpotensi sedang, dan kurang berpotensi dalam penanaman kelapa.

Beberapa penelitian yang menggunakan algoritma *k-means clustering* diantaranya adalah Mar'i & Suprianto (2018) menggunakan algoritma *k-means* dalam pengklasteran *credit card* pembayaran tagihan dengan hasil penelitian didapati pengujian parameter Public

Diterima: 30-04-2020 | Revisi: 01-06-2020 | Diterbitkan: 30-06-2020 | DOI: 10.37034/infeb.v2i2.23

Service Obligation (PSO) dengan niai terbaik [1]. Penelitian Mahmuda dkk (2017) algoritma *k-means clustering* dalam pengklasteran pengunjung perpustakaan dengan hasil penelitian didapati beberapa *cluster* pengunjung sesuai *cluster* yang diinginkan [2]. Penelitian Ridho & Kusuma (2018) algorima *k-means clustering* dalam pendeteksi jaringan pada akses log dengan hasil yang memiliki keunggulan yang mendominasi untuk tiga *cluster* pengunjung [3]. Penelitian Putra & Wadisman (2018) algoritma *k-means clustering* dalam pengimplementasian pelanggan potensial dengan hasil keakuratan untuk kelompok pelanggan yang potensial [4].

Penelitian Widodo & Wahyuni (2017) algoritma *k-means clustering* dalam mengimplementasikan bidang sekripsi mahasiswa dengan hasil pengelompokkan yang baik dan akurat serta mendapatkan *cluster* yang diinginkan [5]. Penelitian Rosmini dkk (2018) algoritma *k-means clustering* dalam pengimplementasian data aktivitas kelompok siswa dengan hasil dapat menentukan kelompok aktivitas siswa sesuai *cluster* yang diharapkan [6]. Penelitian Rustam dkk (2018) algoritma *k-means clustering* dalam pengoptimalan penyakit menular pada daerah endemik dengan hasil data rekam medis menunjukkan hasil yang lebih akurat dan performance yang lebih baik [7]. Penelitian Afriyanti dkk (2018) algoritma *k-means clustering* dalam klasifikasi produk retur dengan hasil performa akurasi sebesar 95,6%, precision sebesar 0,943, dan recall sebesar 0,956 baik [8].

Penelitian Dewata dkk (2018) algoritma *k-means clustering* untuk memprediksi pengambilan jurusan siswa SMA kelas X dengan hasil menunjukkan bahwa nilai-nilai dari mata pelajaran ekstrak dan psikotes akan mempengaruhi untuk menentukan penjurusan siswa baru [9]. Penelitian Fajrin & maulida (2018) algoritma *k-means clustering* dalam analisa pembelian konsumen pada transaksi penjualan dengan hasil bahwa dengan penerapan algoritma *k-means clustering* didapati penjualan *spare part* mana yang paling banyak terjual [10]. Penelitian Fatmawati & Windarto (2018) algoritma *k-means clustering* dalam penerapan *rapidminer* untuk daerah terjangkau daerah terjangkit demam berdarah dengan hasil bahwa terdapat 4 provinsi dengan *cluster* tingkatan tinggi (C1), 13 provinsi dengan *cluster* tingkatan sedang (C2), dan 17 provinsi dengan *cluster* tingkatan rendah (C3) [11]. Penelitian Gustientiedina dkk (2019) algoritma *k-means clustering* dalam *clustering* data obat-obatan pada RSUD Pekan Baru dengan hasil kelompok obat yang termasuk pemakaian sedikit rata-rata permintaan obat setiap tahunnya kurang dari 18000 buah, dan obat yang termasuk pemakaian sedang rata-rata permintaan obat setiap tahunnya diantara 18000-70000 buah, sedangkan obat yang masuk kedalam kelompok obat yang pemakaian tinggi rata–rata permintaan obat setiap tahunnya diatas 70000 buah [12].

Penelitian Maulida (2018) algoritma *k-means clustering* dalam pengelompokkan kunjungan wisata dengan hasil pengelompokan C1 = Taman Impian Jaya Ancol, C2 = Taman Mini Indonesia Indah dan Kebon Binatang Ragunan dan C3 = Monumen Nasional, Museum Nasional, Museum Satria Mandala, Museum Sejarah Jakarta dan Pelabuhan Sunda Kelapa. Hasil pengelompokan C3 menjadi catatan bagi pemerintah Provinsi DKI. Jakarta [13]. Penelitian Silalahi (2018) algoritma *k-means clustering* untuk penjualan produk pada PT Batamas Niaga Jaya dengan hasil Kelompok C0 produk yang tidak laris terdiri dari 657 produk dan kelompok C1 terdiri dari 70 produk yang merupakan produk yang laris [14]. Penelitian Nur dkk (2017) algortima *k-means clustering* untuk *clustering* jurusan siswa baru dengan hasil diperoleh tiga kelompok yaitu kelompok tidak lulus, kelompok rekayasa perangkat lunak dan kelompok teknik komputer jaringan. Terdapat pusat *cluster* dengan *cluster*-1 = 1.4;2.2;2.2, *cluster*2 = 2.28;1.64;4 dan *cluster*-3 = 5;2;6 [15]. Penelitian Parlina dkk (2018) algortima *k-means clustering* dalam menentukan pegawai yang layak mengikuti *asessment center* dengan hasil dapat diambil pengelompokan dengan rata-rata data program SDP yang dapat melakukan *asessment center* lanjutan adalah yang lolos dan hasil klasifikasi program SDP yang hampir lolos harus memperbaiki administrasi seperti kedisiplinan dari bulan juni sampai bulan oktober agar dapat mengikuti *asessment center* lanjutan, sedangkan hasil klasifikasi daftar data program SDP yang Tidak lolos harus memperbaiki data kedisiplinannya selama 1 Tahun [16].

2. Metodologi Penelitian

Objek penelitian ini menggunakan data yang diperoleh dari laporan tahunan data perkebunan pada Dinas Perkebunan Kabupaten Indragiri Hilir tahun 2018.

Knowledge Discovery in Database (KDD) merupakan proses *non-trivial* yang berguna untuk mencari dan mengidentifikasi pola di dalam data, dimana pola yang ditemukan tersebut bersifat benar, agar dapat bermanfaat dan dapat dipahami. KDD juga dapat diartikan sebagai pengorganisasian proses yang bersifat otomanis untuk mengidentifikasi yang benar, yang mana proses tersebut dapat berguna dan dapat menemukan pola dari kumpulan data yang besar (*big data*) dan data kompleks. Data mining dapat diartikan sebagai teknologi baru yang berguna untuk membantu pihak perusahaan untuk menemukan informasi yang sangat penting dari gudang data mereka. Beberapa aplikasi data mining berfokus pada suatu prediksi, meramalkan apa yang akan terjadi pada situasi baru dari data yang menerangkan kejadian masa lalu.

Clustering merupakan sebuah metode pengelompokkan data berdasarkan informasi yang didapatkan dari data yang menguraikan objek tersebut serta jalinan diantaranya yang digunakan untuk mengklasifikasi data dengan cara otomatis untuk menghasilkan beberapa

kelompok yang diukur menggunakan asosiasi, agar menghasilkan data yang serupa berada pada kelompok data yang sama serta data yang berbeda berada pada kelompok data yang berbeda pula. Data masukan untuk *cluster* merupakan perangkat data yang memiliki kesamaan atau perbedaan ukuran antara dua data. Sedangkan untuk data keluaran dari data *cluster* adalah sejumlah kelompok data berbentuk partisi atau kelompok struktur partisi dari sekumpulan data. Hasil tambahan dari data *cluster* adalah gambaran umum dari setiap kelompok yang memiliki peran penting.

Algoritma *k-means* merupakan suatu algoritma yang mengelompokkan iteratif dengan tujuan untuk mempartisi set data pada sejumlah K *clustering* yang telah ditetapkan pada awal pengolahan. Algoritma *k-means* bersifat sederhana dalam pengimplementasian dan menjalankan, pengerjaan relatif cepat, dapat juga dapat beradaptasi dan sudah sering digunakan. Langkah-langkah yang terdapat pada Algoritma *K-means* yaitu dengan menetukan nilai K sebagai jumlah *cluster* yang akan dibuat, selanjutnya menentukan titik pusat *cluster* (K *centroid*) awal secara acak, setelah mendapatkan titik pusat *cluster* langkah selanjutnya ialah menghitung jarak setiap data ke masing-masing *centroid* dengan memanfaatkan rumus korelasi antar dua objek, seperti yang dilihat pada rumus *Euclidean Distance* berikut ini.

$$D_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + \ldots. (n_i - n_j)^2} \quad (1)$$

Dimana:

d_{ij} = Jarak antara i dan j

x_i = Koordinat x objek

x_j = Koordinat x pusat

y_i = Koordinat y objek

y_j = Koordinat y pusat

Setelah mendapatkan hasil dari perhitungan jarak setiap data, langkah selanjutnya menggabungkan setiap data yang dilihat berdasarkan jarak terdekat antara data dan *cluster*. Langkah selanjutnya yaitu menentukan tempat *centroid* baru dengan menghitung nilai rata-rata berdasarkan data pada *cluster* yang serupa, untuk menentukan tempat *centroid* baru untuk menghitung jarak antara data dan *centroid* dapat menggunakan rumus berikut ini [17].

$$Ci = (x1+x2+x3+\cdots+xn)/(\sum x) \quad (2)$$

Dimana:

x1 = nilai record data ke 1

x2 = nilai record data ke 2

$\sum x$= jumlah record data

Algoritma *k-means* merupakan algoritma yang dilakukan secara berulang-ulang hingga menghasilkan *cluster* sesuai kelompok pada data-data yang sama, dan yang berbeda di tempatkan pada kelompok yang berbeda.

Adapun kelebihan yang terdapat pada algoritma *K-means* adalah sebagai berikut [18].

- Sangat mudah dalam pengimplementasian dan proses menjalankan.
- Kebutuhan waktu yang sangat singkat dalam menjalankan pembelajaran.
- Mudah untuk beradaptasi.
- Sangat umum digunakan.

3. Hasil dan Pembahasan

Data yang diambil merupakan data perkebunan kelapa hibrida yang didapat dari Dinas Perkebunan Kabupaten Indragiri Hilir melalui bidang sekretariat. Data tersebut dapat dilihat berikut ini:

Tabel 1. Sample Data Perkebunan Kelapa Hibrida

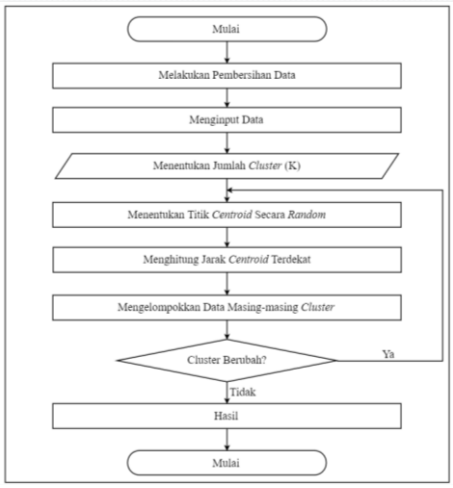
No	Nama Desa	Luas Area Perkebunan (Ha)				Produksi/Tahun (Ton)		Wujud Produksi	Jumlah Petani (orang)
		Kondisi				Pada Tahun Laporan			
		TBM	TM	TTM	Jumlah	Jumlah	Rata-rata		Pemilik
1	Sialang Jaya	7	214	186	407	141,240	660	Kopra	165
2	Sungai Raya	-	7	1	8	4,620	660	Kopra	3
3	Sungai Luar	-	15	21	36	9,900	660	Kopra	15
4	Simpang Jaya	5	15	16	36	9,900	660	Kopra	15
5	Enok	-	2	5	7	1,852	926	Kopra	3
6	Rantau Panjang	-	185	196	381	171,310	926	Kopra	154
7	Pengalihan	8	15	5	28	13,890	926	Kopra	11
8	Sungai Rukam	-	5	-	5	4,630	926	Kopra	2
9	Bagan Jaya	-	35	20	55	32,410	926	Kopra	22
10	Teluk Pinang	-	2	-	2	1,000	500	Kopra	1

Pada penelitian ini peneliti menggunakan analisa sistem dengan metode *k-means* untuk mengelompokkan daerah potensial untuk penanaman kelapa hibrida berdasarkan luas area perkebunan, hasil produksi, dan jumlah petani dengan menjadi 3 kelompok, yaitu kelompok C0 merupakan kelompok daerah kurang berpotensi, C1 merupakan kelompok daerah cukup berpotensi, dan C2 merupakan kelompok daerah sangat berpotensi. Berikut ini merupakan langkah-langkah

yang ditetapkan pada sistem kerja algoritma *k-means*, yaitu:

1. Melakukan pembersihan data
2. Menginput data
3. Menentukan jumlah *cluster* (K)
4. Menentukan titik *centroid* secara random
5. Menghitung jarak *centroid* terdekat
6. Mengelompokkan data masing-masing *cluster*
7. Hasil *cluster*

Dalam penerapan algoritma *k-means* yang diterapkan pada penelitian ini, maka akan digambarkan dalam bentuk *flowchart* agar terlihat lebih jelas dan terarah. *Flowchart* untuk analisa sistem pada penelitian ini dapat dilihat pada gambar berikut:



Gambar 2. *Flowchard* Proses *K-means*

a. Pembersihan Data

Tahapan ini merupakan suatu proses pengoreksian, pengurangan data sebelumnya atau membuang sebagian data yang ganda, menghilangkan data yang kotor, serta memverifikasi data yang dianggap tidak konsisten dan memverifikasi kesalahan yang ada pada data. Hasil pembersihan dapat dilihat berikut ini:

Tabel 2. Sample Pembersihan Data

No	Nama Desa	Luas Lahan (ha)	Produksi (kg)	Jumlah Petani (orang)
1	***2***	***3***	***4***	***5***
1	Sialang Jaya	407	141,240	165
2	Sungai Raya	8	4,620	3
3	Sungai Luar	36	9,900	15
4	Simpang Jaya	36	9,900	15
5	Enok	7	1,852	3
6	Rantau Panjang	381	171,310	154
7	Pengalihan	28	13,890	11
8	Sungai Rukam	5	4,630	2
9	Bagan Jaya	55	32,410	22
10	Teluk Pinang	2	1,000	1

b. Menginput Data

Proses penginputan data akan dilaksanakan kedalam sistem setelah melakukan pembersihan data (*data cleaning*). Data yang digunakan dalam penginputan pada penelitian ini adalah data traning yang sudah dilakukan pembersihan yang diperoleh dari Dinas Perkebunan Kabupaten Indragiri Hilir dengan jumlah sampel yang akan digunakan hanya sebanyak 10 data.

c. Menentukan Jumlah *Cluster*

Proses penerapan metode *k-means* harus menentukan terlebih dahulu berapa banyak jumlah *cluster* yang diinginkan, dimana nantinya *cluster* tersebut akan menjadi acuan pada hasil pengambilan keputusan. Pada proses penelitian ini jumlah *cluster* akan dibagi menjadi 3 *cluster* dimana nantinya ke-3 *cluster* ini akan menentukan hasil untuk penelitian mengenai daerah mana yang masuk dalam kategori daerah kurang berpotensi, cukup berpotensi, dan kategori daerah yang sangat berpotensi untuk penanaman kelapa hibrida di daerah Kabupaten Indragiri Hilir.

d. Menentukan Titik *Centroid* Acak

Pada penentuan *centroid* awal dalam perhitungan pada penelitian ini dilaksanakan dengan cara random, dengan data *centroid* sebagai berikut:

Tabel 3. Sample *Centroid* Acak

Centroid	Luas Lahan (ha)	Produksi (ton)	Jumlah Petani (orang)
C1	8	4,620	3
C2	55	32,410	22
C3	381	171,310	154

e. Menghitung Jarak *Centroid* terdekat

Rumus yang digunakan untuk menghitung jarak setiap data ke masing-masing *centroid* menggunakan rumus korelasi antara tiga objek (*euclidean distance*). Sehingga dengan rumus ini akan ditemukan jarak yang paling terdekat dari setiap data dengan masing-masing *centroid*. Proses perhitungan jarak masing-masing data ke titik pusat *cluster* berikut ini:

1. Proses Iterasi 1

a. Proses menghitung *cluster* 0 (C0)

Proses perhitungan diambil dari data desa dengan nomor urut 1, yaitu:

$$D1 = \sqrt{(407-8)^2 + (141{,}240-4{,}620)^2 + (165-3)^2}$$

$$= 136620.6787$$

b. Proses menghitung *cluster* 1 (C1)

$$D1 = \sqrt{(407-55)^2 + (141{,}240-32{,}410)^2 + (165-22)^2}$$

$$= 108830.6632$$

c. Proses menghitung *cluster* 2 (C2)

$$D1 = \sqrt{(407-381)^2 + (141{,}240-171{,}310)^2 + (165-154)^2}$$

$$= 30070.01325$$

f. Pengelompokkan Data Masing-Masing *Cluster*

Tabel 4. Sample Pengelompokkan Data

No	Nama Desa	Luas Area (ha)	Produksi (ton)	Jumlah Petani (orang)	C1	C2	C3	Jarak Terdekat	C0	C1	C2
1	Sialang Jaya	407	141,240	165	136620.679	108830.663	30070.0133	30070.01325			1
2	Sungai Raya	8	4,620	3	0	27790.0462	166690.486	0	1		
3	Sungai Luar	36	9,900	15	5280.08788	22510.0091	161410.429	5280.087878	1		
4	Simpang Jaya	36	9,900	15	5280.08788	22510.0091	161410.429	5280.087878	1		
5	Enok	7	1,852	3	2768.00018	30558.0436	169458.48	2768.000181	1		
6	Rantau Panjang	381	171,310	154	166690.486	138900.445	0	0			1
7	Pengalihan	28	13,890	11	9270.02503	18520.0229	157420.461	9270.025027	1		
8	Sungai Rukam	5	4,630	2	10.4880885	27780.0522	166680.493	10.48808848	1		
9	Bagan Jaya	55	32,410	22	27790.0462	0	138900.445	0		1	
10	Teluk Pinang	2	1,000	1	3620.00552	31410.0517	170310.49	3620.005525	1		
								Jumlah	7	1	2

Tabel diatas merupakan hasil *clusterisasi* dari masing-masing data sample penelitian, dimana untuk *cluster* 0 atau kurang berpotensi berjumlah 7 data, *cluster* 1 atau kategori cukup berpotensi berjumlah 1 data, dan *cluster* 2 atau kategori sangat berpotensi berjumlah 2 data.

Selanjutnya untuk menghitung iterasi ke-2 harus melalui perhitungan dari menghitung rata-rata hasil data yang sama pada *cluster* yang posisi sama, untuk perhitungan mencari *centroid* baru sebagai berikut:

1. *Cluster* 0 (C0) berjumlah 7 data

C1=(8+36+36+7+28+5+2)/7

= 17.42857143

C2=(4,620+9,900+9,900+1,852+13,890+4,630+1,00)/7

= 6541.714286

C3=(3+15+15+3+11+2+1)/7

= 7.142857143

2. *Cluster* 1 (C1) berjumlah 1 data

C1=(55)/1

= 55

C2=(32,410)/1

= 32,410

C3=(22)/1

= 22

3. *Cluster* 2 (C2) berjumlah 2 data

C1=(407+381)/2

= 394

C2=(141,240+171,310)/2

= 156275

C3=(165+154)/2

= 159.5

Setelah perhitungan selesai, didapat *centroid* baru untuk perhitungan iterasi selanjutnya, proses ini bertujuan untuk memastikan hasil perhitungan sebelumnya masih terdapat perubahan posisi data *cluster* atau tidak, dimana data *centroid* baru dapat dilihat berikut:

Tabel 5. *Centroid* Baru

Centroid	Luas Lahan (ha)	Produksi (ton)	Jumlah Petani (orang)
C1	17.42857143	6541.714286	7.142857143
C2	55	32410	22
C3	394	156275	159.5

3.7. Menghitung Jarak *Centroid* Baru

Rumus yang digunakan untuk menghitung jarak setiap data ke masing-masing *centroid* menggunakan rumus korelasi antara tiga objek (*euclidean distance*). Sehingga dengan rumus ini akan ditemukan jarak yang paling terdekat dari setiap data dengan masing-masing *centroid*. Data yang dirunakan untuk *centroid* baru didapat dari hasil perhitungan sebelumnya. Proses perhitungan jarak masing-masing data ke titik pusat *cluster* berikut ini:

2. Proses Iterasi 2

a. Proses menghitung *cluster* 0 (C0)

Proses perhitungan diambil dari data desa dengan nomor urut 1, yaitu:

$D1 = \sqrt{(407-17.42857143)^2 + (141{,}240-6541.714286)^2 + (165-7.142857143)^2}$

= 134698.9416

b. Proses menghitung *cluster* 1 (C1)

$D1 = \sqrt{(407-55)^2 + (141{,}240-32{,}410)^2 + (165-22)^2}$

= 108830.6632

c. Proses menghitung *cluster* 2 (C2)

$D1 = \sqrt{(407-394)^2 + (141{,}240-156275)^2 + (165-159.5)^2}$

= 15035.00663

3.8. Pengelompokkan Data Masing-Masing *Cluster*

Setelah proses perhitungan dilakukan pada tahap sebelumnya, maka jarak dari masing-masing data *cluster* 0 (C0), *cluster* 1 (C1), *cluster* 2 (C2) dapat ditampilkan. Hasil perhitungan seluruh data sumple penelitian dapat dilihat berikut ini:

Tabel 6. Sample Pengelompokkan Data

No	Nama Desa	Luas Area (ha)	Produksi (ton)	Jumlah Petani (orang)	C1	C2	C3	Jarak Terdekat	C0	C1	C2
1	Sialang Jaya	407	141,240	165	134698.9416	108830.6632	15035.00663	15035.00663			1
2	Sungai Raya	8	4,620	3	1921.741881	27790.04624	151655.572	1921.741881	1		
3	Sungai Luar	36	9,900	15	3358.346255	22510.00911	146375.5091	3358.346255	1		
4	Simpang Jaya	36	9,900	15	3358.346255	22510.00911	146375.5091	3358.346255	1		
5	Enok	7	1,852	3	4689.727711	30558.04361	154423.5642	4689.727711	1		
6	Rantau Panjang	381	171,310	154	164768.7523	138900.4453	15035.00663	15035.00663			1
7	Pengalihan	28	13,890	11	7348.294331	18520.02295	142385.5478	7348.294331	1		
8	Sungai Rukam	5	4,630	2	1911.761603	27780.0522	151645.5807	1911.761603	1		
9	Bagan Jaya	55	32,410	22	25868.31727	0	123865.5402	0		1	
10	Teluk Pinang	2	1,000	1	5541.739167	31410.05174	155275.5757	5541.739167	1		
								Jumlah	**7**	**1**	**2**

Tabel diatas merupakan hasil dari pengelompokkan yang dilakukan pada proses iterasi ke-2 dimana dari hasil yang diperoleh dapat disimpulkan bahwa data tersebut tidak didapati perubahan untuk setiap *cluster*. Dengan demikian proses iterasi hanya dilakukan sampai iterasi ke-2 dengan hasil untuk *cluster* 0 kategori kurang berpotensi berjumlah 7 data desa, *cluster* 1 kategori cukup berpotensi berjumlah 1 data desa, dan untuk *cluster* 2 kategori sangat berpotensi berjumlah 2 data desa, dengan demikian proses iterasi dihentikan. Hasil untuk setiap *cluster* dari penelitian ini dapat dilihat berikut:

Tabel 7. Hasil *cluster* 0

No	Nama Desa	Luas Lahan (ha)	Produksi (ton)	Jumlah Petani (orang)
1	Sungai Raya	8	4,620	3
2	Sungai Luar	36	9,900	15
3	Simpang Jaya	36	9,900	15
4	Enok	7	1,852	3
5	Pengalihan	28	13,890	11
6	Sungai Rukam	5	4,630	2
7	Teluk Pinang	2	1,000	1
	Jumlah	17.428571	6541.7143	7.142857143

Tabel diatas menunjukkan hasil pengelompokkan untuk kategori daerah kurang berpotensi untuk penanaman kelapa hibrida di Kabupaten Indragiri Hilir pada *cluster* 0 berjumlah 7 Desa.

Tabel 8. Hasil *cluster* 1

No	Nama Desa	Luas Lahan (ha)	Produksi (ton)	Jumlah Petani (orang)
1	Bagan Jaya	55	32,410	22
	Jumlah	55	32,410	22

Tabel diatas menunjukkan hasil pengelompokkan untuk kategori daerah cukup berpotensi untuk penanaman kelapa hibrida di Kabupaten Indragiri Hilir pada *cluster* 1 berjumlah 1 Desa.

Tabel 9. Hasil *cluster* 2

No	Nama Desa	Luas Lahan (ha)	Produksi (ton)	Jumlah Petani (orang)
1	Sialang Jaya	407	141,240	165
2	Rantau Panjang	381	171,310	154
	Jumlah	394	156274	159.5

Tabel diatas menunjukkan hasil pengelompokkan untuk kategori daerah sangat berpotensi untuk penanaman kelapa hibrida di Kabupaten Indragiri Hilir pada *cluster* 2 berjumlah 2 Desa.

4. Kesimpulan

Optimalisasi penggunaan lahan perkebunan kelapa hibrida menggunakan *K-means Clustering* dapat digunakan sebagai metode yang tepat untuk mengelompokkan daerah potensial dalam penanaman kelapa hibrida. Dari 10 data desa yang diolah menunjukkan daerah dengan kategori kurang potensial (C0) sebanyak 7 data desa, untuk daerah dengan kategori cukup berpotensi (C1) sebanyak 1 data desa, dan untuk daerah dengan kategori sangat berpotensi (C2) sebanyak 2 data desa.

Daftar Rujukan

[1] Mar'i, F., & Supianto, A. A. (2018). *Clustering* Credit Card Holder Berdasarkan Pembayaran Tagihan Menggunakan Improved *K-means* dengan Particle Swarm Optimization. *Jurnal Teknologi Informasi dan Ilmu Komputer, 5*(6), 737-744. http://dx.doi.org/10.25126/jtiik.201856858

[2] Mahmuda, F., Sitorus, M. A. R., Widyastuti, H., & Kurniawan, D. E. (2017). Clustering Profil Pengunjung Perpustakaan Menggunakan Algoritma K-means. *Journal of Applied Informatics and Computing, 1*(1), 14-21. https://doi.org/10.30871/jaic.v1i1.476

[3] Ridho, F., & Kusuma, A. A. (2019). Deteksi Intrusi Jaringan dengan *K-means clustering* pada Akses Log dengan Teknik Pengolahan Big Data. *Jurnal Aplikasi Statistika & Komputasi Statistik, 10*(1), 53-66. https://doi.org/10.34123/jurnalasks.v10i1.202.

[4] Putra, R. R., & Wadisman, C. (2018). Implementasi Data mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K Means. *INTECOMS: Journal of Information Technology and Computer Science*, 1(1), 72-77. https://doi.org/10.31539/intecoms.v1i1.141

[5] Widodo, W., & Wahyuni, D. (2017). Implementasi Algoritma *K-means clustering* Untuk Mengetahui Bidang Skripsi Mahasiswa Multimedia Pendidikan Teknik Informatika Dan Komputer Universitas Negeri Jakarta. *PINTER: Jurnal Pendidikan Teknik Informatika dan Komputer*, *1*(2), 157-166. https://doi.org/10.21009/pinter.1.2.10

[6] Rosmini, R., Fadlil, A., & Sunardi, S. (2018). Implementasi Metode *K-means* Dalam Pemetaan Kelompok Mahasiswa Melalui Data Aktivitas Kuliah. *IT Journal Research and Development*, 3(1), 22-31. https://doi.org/10.25299/itjrd.2018.vol3(1).1773

[7] Rustam, S., Santoso, H. A., & Supriyanto, C. (2018). Optimasi *K-means clustering* Untuk Identifikasi Daerah Endemik Penyakit Menular Dengan Algoritma Particle Swarm Optimization di Kota Semarang. *ILKOM Jurnal Ilmiah*, *10*(3), 251-259.

[8] Arifiyanti, A. A., Pradana, R. M., & Novian, I. F. (2018). Klasifikasi Produk Retur dengan Algoritma Pohon Keputusan C4. 5. *Jurnal IPTEK*, *22*(1), 79-86. https://doi.org/10.31284/j.iptek.2018.v22i1.243

[9] Dewata, A. S. P., Ibrahim, P. Z., & Agung, H. Aplikasi Data mining Berbasis Android Menggunakan Algoritma *K-means clustering* dan Algoritma C4. 5 Untuk Memprediksi Pengambilan Jurusan Siswa SMA Kelas X Pada Sekolah Bunda Mulia.

[10] Fajrin, A. A., & Maulana, A. (2018). Penerapan data mining untuk analisis pola pembelian konsumen dengan algoritma fp-growth pada data transaksi penjualan spare part motor. *Kumpulan jurnaL Ilmu Komputer (KLIK)*, *5*, 27-36. http://dx.doi.org/10.20527/klik.v5i1.100

[11] Fatmawati, K., & Windarto, A. P. (2018). Data mining: Penerapan Rapidminer Dengan *K-means Cluster* Pada Daerah Terjangkit Demam Berdarah Dengue (DBD) Berdasarkan Provinsi. *Computer Engineering, Science and System Journal*, *3*(2), 173-178. https://doi.org/10.24114/cess.v3i2.9661

[12] Gustientiedina, G., Adiya, M. H., & Desnelita, Y. (2019). Penerapan Algoritma *K-means* Untuk *Clustering* Data Obat-Obatan. *Jurnal Nasional Teknologi dan Sistem Informasi*, *5*(1), 17-24. https://doi.org/10.25077/TEKNOSI.v5i1.2019.17-24

[13] Maulida, L. (2018). Penerapan Datamining Dalam Mengelompokkan Kunjungan Wisatawan Ke Objek Wisata Unggulan Di Prov. Dki Jakarta Dengan *K-means*. *JISKA (Jurnal Informatika Sunan Kalijaga)*, *2*(3), 167-174. http://dx.doi.org/10.14421/jiska.2018.23-06

[14] Yahya, K., & Mahpuz, M. M. (2019). Penggunaan Algoritma *K-means* Untuk Menganalisis Pelanggan Potensial Pada Dealer SPS Motor Honda Lombok Timur Nusa Tenggara Barat. *Infotek: Jurnal Informatika dan Teknologi*, *2*(2), 109-118.

[15] Nur, F., Zarlis, M., & Nasution, B. B. (2017). Penerapan Algoritma *K-Means* pada Siswa Baru Sekolah Menengah Kejuruan Untuk *Clustering* Jurusan. *InfoTekJar: Jurnal Nasional Informatika dan Teknologi Jaringan*, *1*(2), 100-105. https://doi.org/10.30743/infotekjar.v1i2.70

[16] Parlina, I., Windarto, A. P., Wanto, A., & Lubis, M. R. (2018). Memanfaatkan Algoritma *K-means* dalam Menentukan Pegawai yang Layak Mengikuti Asessment Center untuk *Clustering* Program SDP. *Computer Engineering, Science and System Journal*, *3*(1), 87-93. https://doi.org/10.24114/cess.v3i1.8192

[17] Santoso, S., & Nurmalina, R. (2017). Perencanaan dan Pengembangan Aplikasi Absensi Mahasiswa Menggunakan Smart Card Guna Pengembangan Kampus Cerdas. *Jurnal Integrasi*, *9*(1), 84-91. https://doi.org/10.30871/ji.v9i1.288

[18] Ayu, F., & Permatasari, N. (2018). Perancangan Sistem Informasi Pengolahan Data PKL (Praktek Kerja Lapangan) di Devisi Humas Pada Pt Pegadaian. *Jurnal Intra Tech*, *2*(2), 12-26.