

Implementasi Algoritma K-Medoids pada Pengelompokan Keragaman Kelompok Tani

Dina Andriana^{1,*}, Eka Irawan², Rizky Khairunnisa Sormin³

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

Email: ¹dinaandriana182000@gmail.com, ²ekairawan@amiktunasbangsa.ac.id, ³rizkysormin@amiktunasbangsa.ac.id

(*: dinaandriana182000@gmail.com)

Abstrak

Kelompok tani memiliki peran dalam menggerakkan pembangunan pertanian di Perdesaan dan Pelaku utama dalam Pemberdayaan Ekonomi berbasis Pertanian. Penelitian ini bertujuan untuk membuat sebuah model pengelompokan dengan menggunakan algoritma K-Medoids. Algoritma K-Medoids atau dikenal pula dengan PAM (Partitioning Around Medoids) menggunakan metode partisi *clustering* untuk mengelompokkan sekumpulan n objek menjadi sejumlah k *cluster*, dengan melakukan pengelompokan kelompok tani sehingga diperoleh tingkat perkembangan dan kelompok tani sesuai komoditi unggulan. Penelitian ini dilaksanakan di Pematang Bandar Simalungun. Data dalam penelitian ini bersumber dari BBP (Balai Penyuluhan Pertanian). Data yang terkumpul dihitung sesuai dengan data tersebut sehingga diperoleh kelompok tani sesuai komoditi unggulan. Pengelompokan dilakukan berdasarkan jumlah data kelompok tani dari 4 desa yang ada di pematang bandar. Diharapkan penelitian ini dapat menjadi Gambaran pihak BBP dalam mengambil kebijakan pengembangan sumber daya masing-masing kelompok tani guna kedepannya semakin banyak kelompok tani yang masuk dalam *cluster* tertinggi dan diharapkan berimbas pada peningkatan produktivitas kelompok tani tersebut.

Kata Kunci: K-Medoids, Klastering, Produktivitas, Pertanian, Kelompok Tani

Abstract

Farmer groups have a role in driving agricultural development in rural areas and are the main actors in agriculture-based economic empowerment. This study aims to create a clustering model using the K-Medoids algorithm. The K-Medoids algorithm or also known as PAM (Partitioning Around Medoids) uses the clustering partition method to group a set of n objects into a number of k clusters, by grouping farmer groups so that the level of development and farmer groups according to superior commodities is obtained. This research was conducted in Pematang Bandar Simalungun. The data in this study were sourced from BBP (Agricultural Extension Center). The data collected is calculated according to the data so that farmer groups are obtained according to the leading commodity. The grouping is done based on the number of farmer group data from 4 villages in Pematang Bandar. It is hoped that this research can be an illustration of the BBP in taking resource development policies for each farmer group so that in the future more farmer groups are included in the highest cluster and are expected to have an impact on increasing the productivity of these farmer groups.

Keywords: K-Medoids, Clustering, Productivity, Agriculture, Farmer Groups

1. PENDAHULUAN

Data mining merupakan proses pengumpulan informasi-informasi penting dari suatu data yang besar [1]. Data mining memiliki peranan penting dalam beberapa bidang yaitu, bidang keuangan, perindustrian, ilmu, teknologi dan cuaca. Pada umumnya kajian dalam data mining membahas beberapa metode seperti, *classification*, *association*, *clustering* [2]. *Association* mencari pola hubungan yang terdapat pada data atau basis data. Data Mining adalah suatu proses ataupun kegiatan untuk mengumpulkan data yang berukuran besar kemudian mengekstraksi data tersebut menjadi informasi-informasi yang nantinya dapat digunakan [3], [4]. Untuk itu pada penelitian ini menggunakan algoritma K-Medoids. Algoritma K-Medoids ini juga sering disebut algoritma *Partitioning Around Medoids* (PAM) [5]. K-Medoids menggunakan metode pengelompokan partisi untuk mengelompokkan sekumpulan n objek menjadi sejumlah *cluster*. Algoritma ini menggunakan objek yang mewakili sebuah *cluster* [6]. Algoritma *partitioning* merupakan algoritma pengelompokan data kedalam sejumlah *cluster* tanpa ada dilakukan suatu struktur hirarki antara satu dengan yang lainnya [7]. Menurut [8] Kelompok tani merupakan wadah tempat bernaungnya beberapa petani/peternak/pekebun sebagai tempat belajar, bekerjasama dan unit produksi yang dibentuk atas dasar kesamaan domisili dan hamparan lahan pertanian. Tujuan dibentuknya kelompok tani supaya petani dapat menjalankan usaha taninya secara bersama-sama sehingga dapat meningkatkan dan mengembangkan usaha tani yang dijalankan oleh anggota dan kelompok dengan beberapa prinsip kehidupan berkelompok diantaranya adalah prinsip partisipatif [9]. Kelompok tani juga membutuhkan penyuluh untuk menumbuhkembangkan usaha taninya serta membantu menggali potensi, memecahkan masalah usaha tani anggotanya secara lebih efektif dan memudahkan dalam mengakses informasi pasar, teknologi, permodalan dan sumber daya lainnya. Kelompok tani di Pematang Bandar memiliki komoditi unggulan yaitu padi sawah dan perkebunan, peternakan, perikanan. maka dari itu penelitian ini dilakukan untuk mengelompokkan kelompok tani sesuai komoditi unggulan [10]. Dan perlu juga dikelompokkan penyuluh sesuai komoditi unggulan, dengan mengelompokkan terlebih dahulu kelompok tani yang memiliki komoditi unggulan yang sama diberbagai desa agar mempermudah penyuluh dalam membina kelompok tani sesuai komoditi unggulan sehingga datanya jelas penyuluh dengan komoditi unggulannya. Metode yang digunakan dalam penelitian ini adalah Metode K-Medoids dan data diambil di kantor BBP (Balai Penyuluhan Pertanian) di Pematang Bandar Simalungun [11].

Beberapa penelitian pernah membahas masalah pengelompokan dan metode yang digunakan adalah algoritma K-Medoids *Clustering*. Seperti pada Penelitian yang dilakukan oleh (Pramesiti et al., 2017) bahwa hasil pengelompokan

Potensi Kebakaran Hutan/ Lahan Berdasarkan Persebaran Titik Panas menggunakan algoritma K-Medoids terbentuk 2 *cluster*, *cluster* tinggi dengan hasil rata-rata brightness sebesar 344.47 OK dengan rata-rata confidence 87.08% dan *cluster* 2 masuk dalam potensi sedang dengan hasil rata-rata brightness sebesar 318.80 OK dengan rata-rata confidence sebesar 58,73%. Pada Penelitian yang dilakukan oleh [12] bahwa hasil pengelompokan Produktivitas Padi Menurut Provinsi menggunakan algoritma K-Medoids terbentuk *cluster* 1 sebanyak 17 Provinsi, *cluster* 2 sebanyak 7 Provinsi dan *cluster* 3 sebanyak 10 Provinsi. Berdasarkan permasalahan yang telah diuraikan di atas, perlu dilakukan pengelompokan untuk menentukan keragaman kelompok tani dengan menggunakan teknik data mining dengan metode K-Medoids. Berdasarkan penelitian sebelumnya Algoritma K-Medoids dinilai sangat cocok digunakan untuk menyelesaikan masalah pengelompokan oleh sebab itu penulis mengangkat penelitian yang berjudul “Implementasi Algoritma K-Medoids Untuk Pengelompokan Keragaman Kelompok Tani”.

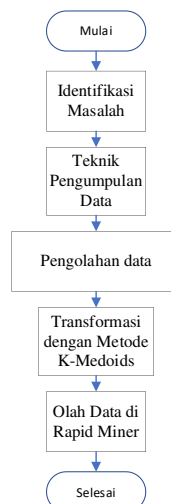
2. METODOLOGI PENELITIAN

2.1. Metode Penelitian

Langkah yang dilakukan peneliti adalah Metodologi penelitian untuk mengumpulkan data serta melakukan investigasi pada data yang didapatkan. Beberapa Gambaran rancangan metodologi penelitian antara lain adalah prosedur pengumpulan data, analisa data, lokasi dan waktu peneliti, rancangan penelitian [13].

2.1.1. Rancangan Penelitian

Berikut ini rancangan penelitian yang berisi diagram blok yang menjelaskan alur kerja penelitian sebagai berikut [14]:



Gambar 1. Rancangan Penelitian

Berikut penjelasan dari alur rancangan penelitian pada Gambar 1

- Identifikasi Masalah Merupakan proses awal dari pengenalan suatu masalah yang terjadi dimana suatu objek tertentu dalam situasi yang terpilih dapat dikenali sebagai suatu masalah. Permasalahan yang ada dalam penelitian adalah Kelompok Tani
- Teknik Pengumpulan Data Pengumpulan data penelitian ini diperoleh dari langsung dari Kantor BPP (Balai Penyuluhan Pertanian)
- Pengolahan Data Pada tahap pengolahan data, data diolah untuk mendapatkan hasil yang kemudian dapat berguna dalam pengambilan keputusan. Data diolah sehingga menjadi informasi.
- Transformasi dengan Metode K-Medoids Transformasi data pada data yang berjenis alfabet seperti nama Kelompok Tani berdasarkan Komoditi Unggulan, harus dilakukan proses inisialisasi data terlebih dahulu ke dalam bentuk angka/numerikal. Kemudian mengelompokkan data yang sudah ada dengan metode K-Medoids.
- Olah Data di RapidMiner. RapidMiner merupakan suatu aplikasi pilihan yang tujuannya untuk mengekstrak data dengan metode yang ada di dalam data mining. Pengujian disini kita menggunakan sebuah *software* RapidMiner, dengan pengujian data menggunakan *software* kita akan membandingkan bagaimana hasil pengolahan data secara manual dengan hasil pengolahan data menggunakan sebuah *software*.

2.1.2. Prosedur Pengumpulan Data

Dalam melakukan penelitian ada beberapa prosedur dalam penumpulan data yaitu:

- Sumber data penelitian didapat dari Kantor BPP (Balai Penyuluhan Pertanian) yang digunakan dalam penelitian ini adalah data Kelompok Tani yang terdapat pada Desa Pardomuan Nauli Pematang Bandar.
- Kepustakaan dengan memanfaatkan buku, prosiding atau jurnal sebagai media untuk bahan referensi dalam menentukan parameter yang digunakan dalam penelitian.

2.2. Analisis Data

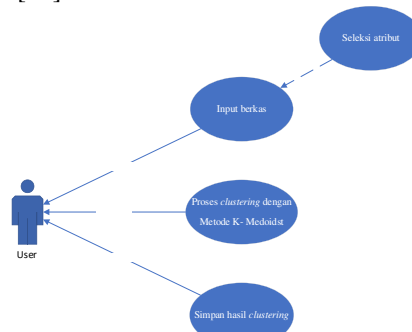
Proses analisis data bisa dilakukan setelah adanya pengumpulan data. Pengumpulan data yang digunakan penulis untuk mendukung terlaksananya penelitian ini adalah dengan menggunakan Data Sekunder. Data Sekunder adalah sumber data penelitian yang dihasilkan melalui media perantara seperti buku, bukti yang telah ada, catatan, atau arsip baik yang tidak dipublikasikan maupun yang dipublikasikan secara umum. Data yang diperoleh dari Kantor BPP (Balai Penyuluhan Pertanian).

2.2.1. Alat Analisis Data

Dalam penelitian ini peneliti menggunakan Microsoft Excel dan RapidMiner, Microsoft Excel digunakan untuk memudahkan dalam proses perhitungan yang akan dilakukan. Selanjutnya diaplikasikan ke RapidMiner untuk melihat keakuratan hasil yang didapat dari perhitungan di Excel dan akan dicocokkan dengan hasil yang sudah diaplikasikan ke RapidMiner [15].

2.2.2. Instrumen Penelitian

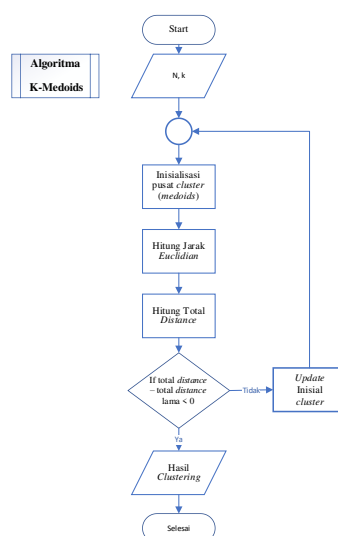
Dalam penelitian ini, penulis melakukan beberapa proses pengumpulan data dan pengolahan data yaitu dengan cara memasukkan data diolah, kemudian diproses dengan menggunakan metode K-Medoids sehingga ditemukan Jumlah *cluster*. Sebelum diproses data terlebih dahulu diseleksi menggunakan *Software* RapidMiner kemudian disimpan hasil *cluster* yang telah didapat [16].



Gambar 2. Instrumen Penelitian

2.2.3. Penyelesaian Metode Kmedoids

Proses pengklasteran akan dilakukan dengan menggunakan algoritma K-Medoids. Dimulai dari memasukkan jumlah data yang akan dibutuhkan dan jumlah *cluster* yang ditentukan. Selanjutnya dilakukan proses dari seluruh perhitungan K-Medoids. Hasil *cluster* akan ditampilkan pada akhir proses ketika proses perhitungan telah selesai. Diagram alir algoritma K-Medoids dapat dilihat pada Gambar.



Gambar 3. Flowchart Metode K-Medoids

3. HASIL DAN PEMBAHASAN

3.1. Hasil

Pada bagian ini akan menguraikan bagaimana proses perhitungan data secara manual dan uji data secara sistematis. Untuk mengetahui adanya proses analisa data, terlebih dahulu data yang digunakan harus bersifat valid. Data yang digunakan dalam penelitian ini sudah bersifat valid karena didapatkan dari BPP (Balai Penyuluhan Pertanian), sehingga data dapat langsung diproses dan diuji kevalidannya untuk sebuah penelitian. Dalam melakukan penelitian ini penulis melakukan sebuah analisa data statistik yang bersifat sekunder dimana data berkaitan dengan sebuah permasalahan yang diteliti.

3.2. Pengelolaan Data

Untuk mendapatkan hasil dari penelitian yang dilakukan, berikut uraian perhitungan manual proses *clustering* Pengelompokan Keragaman Kelompok Tani sesuai Komoditi Unggulan dengan menggunakan sebuah algoritma K-Medoids. Proses *clustering* dimulai dengan menentukan sebuah data yang akan di *cluster*. Dalam hal ini variable data yang akan digunakan adalah data Kelompok Tani Pematang Bandar, dan dari penelitian ini penulis mengambil nilai rata-rata dari data setiap desa sebagai data yang akan ditentukan *cluster* nya. Berikut adalah langkah-langkah penyelesaian yang dilakukan penulis dalam mengelompokkan keragaman kelompok tani dengan menggunakan algoritma K-Medoids. Tahapan *Knowledge Discovery in Database* (KKD).

a) Seleksi Data

Tahap data selection merupakan pemilihan (seleksi) data yang akan digunakan nantinya pada tahapan data mining. Tahapan ini perlu dilakukan dalam penggalian informasi pada KKD. Seleksi yang dilakukan adalah dengan menghapus data yang tidak diperlukan dalam proses data mining. Pada penelitian ini data yang digunakan adalah data kelompok tani masing-masing desa. Sedangkan data kelompok tani dari keseluruhan masing-masing desa tidak digunakan dalam proses data mining. Berikut data hasil selection pada penelitian ini dapat dilihat pada Tabel dibawah ini.

Tabel 1. Data Selection yang Digunakan

No	Nama Kelompok Tani	Penyuluh	Komoditi Unggulan	Luas Darat Ha	Luas sawah Ha
1	Indah	1	1	0,5	9,97
2	Bersih	1	1	0,5	7,76
3	Suka Cita	1	2	0	10,61
4	Suka Dame	1	1	0,5	10,83
5	Swasembada	1	2	0	8,78
6	Swakarya	1	3	0	18,93
7	Tani Mulyo	1	2	0	12,73
8	Tani Maju	1	1	0,5	11,57
9	Melati	1	1	0,5	11,99
10	Mawar	1	1	0,5	11,47
....
96	Subur Karya	3	2	0	20
97	Maju Karya	3	2	0	12,5
98	Hidup Karya	3	2	0	23
99	Harapan karya	3	2	0	9
100	Serasi	3	2	0	15
101	Saroha	3	2	0	15,5
102	Maju Bersama	3	5	7,75	0
103	Sumber Makmur	3	4	0	0

b) Data Preprocessing

Tahapan ini biasa dikenal sebagai tahapan data cleaning sesuai dengan artinya yaitu melakukan pembersihan data. Data dibersihkan dari missing values dan data duplikasi. Solusi yang digunakan pada penelitian ini untuk masalah duplikasi data dengan menghapusnya dan missing values. Karena data yang diambil sudah tidak memiliki missing. Maka data tidak perlu diberlakukan perubahan.

c) Transformasi Data

Pada tahap ini akan dilakukan transformasi data berdasarkan data pada Tabel 4.1 yaitu dengan mengubah bentuk data awal menjadi bentuk yang baru agar bisa diproses sesuai dengan algoritma yang digunakan yaitu K-Medoids. Atribut merupakan bagian data yang mewakili karakteristik atau feature dari objek data. Metode yang digunakan pada tahap transformasi ini ialah melakukan normalisasi data, dengan perhitungan Min-Max. *Normalization* pada rumus :

$$v' = \frac{v - \min a}{\max a - \min a} \quad (1)$$

Dimana:

v' : Hasil normalisasi;

v : Nilai data yang akan di normalisasi;

maxa : Nilai maksimum pada kolom data;
mina : Nilai minimum pada kolom data.

Dari data yang ada diperoleh diketahui terdapat 4 kolom. Serta diperoleh nilai max a dan min a sebagai berikut:

Tabel 2. Data Min dan Max

Kolom	MIN	MAX
Kolom ke 1	1	3
Kolom ke 2	1	5
Kolom ke 3	0	33,2
Kolom ke 4	0	56

Sehingga diperoleh perhitungan normalisasi sebagai berikut:

$$v^1 = \frac{1-1}{3-1} = 0,000$$

Tabel 3. Data Min a dan Max a Masing-Masing Kolom

No	Penyuluh	Komoditi Unggulan	Luas Darat Ha	Luas Sawah Ha
1	0,000	0,000	0,015	0,178
2	0,000	0,000	0,015	0,139
3	0,000	0,250	0,000	0,189
4	0,000	0,000	0,015	0,193
5	0,000	0,250	0,000	0,157
6	0,000	0,500	0,000	0,338
7	0,000	0,250	0,000	0,227
8	0,000	0,000	0,015	0,207
9	0,000	0,000	0,015	0,214
10	0,000	0,000	0,015	0,205
11	0,000	0,500	0,000	0,198
12	0,000	0,000	0,015	0,128
13	0,000	0,000	0,030	0,088
14	0,000	0,500	0,000	0,095
15	0,000	0,500	0,000	0,261
16	0,000	0,250	0,000	0,121
17	0,000	0,250	0,000	0,149
18	0,000	0,000	0,060	0,133
19	0,000	0,250	0,000	0,162
20	0,000	0,250	0,000	0,247
21	0,000	0,500	0,000	0,393
22	0,000	0,000	0,030	0,192
23	0,000	0,000	0,030	0,336
24	0,000	0,250	0,000	0,158
25	0,000	0,500	0,000	0,253
26	0,000	0,750	0,000	0,000
27	0,000	0,750	0,000	0,000
28	0,000	0,750	0,000	0,000
29	0,000	0,750	0,000	0,000
30	0,000	0,750	0,000	0,000
31	0,000	0,750	0,000	0,000
32	0,000	0,750	0,000	0,000
33	0,000	0,750	0,000	0,000
34	0,000	0,750	0,000	0,000
.....
103	0,500	0,750	0,000	0,000

3.2.1. Proses Perhitungan Algoritma Kmedoids

Teknik partisi klasik *Clustering* yang mengelompokkan data set dari ni objek ke dalam kelompok k yang dikenal apriori. Dibandingkan dengan K-Mens, K-Medoid lebih kuat untuk mengatasi kebisingan (noise) dan pencilan (outlier) karena meminimalkan sejumlah dissimilarities berpasangan, bukan jumlah kuadrat jarak Euclidean. Sebuah medoid dapat didefinisikan sebagai objek *cluster* yang rata-rata perbedaan untuk semua objek dalam *cluster* minimal yaitu titik paling berlokasi di *cluster*.

Langkah-langkah algoritma K-Medoid:

a) Inisialisasi pusat *cluster* sebanyak 3 *cluster* dari data sampel.

Untuk pemilihan setiap medoid dipilih secara acak (random) di asumsikan Indah, Mulia tani dan Nauli godang sebagai medoid awal. Berikut merupakan Tabel medoid awal.

Tabel 4. Medoids Awal

No	Nama Kelompok Tani	Penyuluh	Komoditi Unggulan	Luas Darat Ha	Luas sawah Ha
1	Indah	1	1	0,5	9,97
37	Mulia tani	2	1	0	7
67	Nauli godang	3	2	0	25,5

b) Menghitung nilai jarak terdekat (*cost*) dengan persamaan *Euclidian Distance*.

Untuk menghitung jarak antara titik centroid dengan titik tiap objek menggunakan *Euclidian Distance*. Rumus untuk menghitung jarak menggunakan persamaan diatas. Maka perhitungan untuk menghitung jarak setiap objek dengan medoid awal pada Cost 1 adalah sebagai berikut :

$$\text{Indah (C1)} = \sqrt{\frac{((0,000 - 0,000)^2 + (0,000 - 0,000)^2 + (0,015 - 0,015)^2 + (0,178 - 0,178)^2)}{4}} = 0$$

$$\text{Bersih (C2)} = \sqrt{\frac{((0,000 - 0,000)^2 + (0,000 - 0,000)^2 + (0,015 - 0,015)^2 + (0,178 - 0,139)^2)}{4}} = 0,039464286$$

$$\text{Suka Cita (C3)} = \sqrt{\frac{((0,000 - 0,000)^2 + (0,000 - 0,250)^2 + (0,015 - 0,000)^2 + (0,178 - 0,189)^2)}{4}} = 0,250713827$$

Seterusnya hingga data ke 103 (Sumber makmur).

Sedangkan perhitungan untuk menghitung jarak setiap objek dengan medoid awal pada Cost 2 adalah sebagai berikut :

$$\text{Indah (C1)} = \sqrt{\frac{((0,000 - 0,000)^2 + (0,500 - 0,000)^2 + (0,000 - 0,015)^2 + (0,253 - 0,178)^2)}{4}} = 0,505871041$$

$$\text{Bersih (C2)} = \sqrt{\frac{((0,000 - 0,000)^2 + (0,500 - 0,000)^2 + (0,000 - 0,015)^2 + (0,253 - 0,139)^2)}{4}} = 0,51323559$$

$$\text{Suka Cita (C3)} = \sqrt{\frac{((0,000 - 0,000)^2 + (0,500 - 0,250)^2 + (0,000 - 0,000)^2 + (0,253 - 0,189)^2)}{4}} = 0,258044303$$

Seterusnya hingga data ke 103 (Sumber makmur).

Sedangkan perhitungan untuk menghitung jarak setiap objek dengan medoid awal pada Cost 3 adalah sebagai berikut :

$$\text{Indah (C1)} = \sqrt{\frac{((0,500 - 0,000)^2 + (0,750 - 0,000)^2 + (0,000 - 0,015)^2 + (0,000 - 0,178)^2)}{4}} = 0,918925202$$

$$\text{Bersih (C2)} = \sqrt{\frac{((0,500 - 0,000)^2 + (0,750 - 0,000)^2 + (0,000 - 0,015)^2 + (0,000 - 0,139)^2)}{4}} = 0,139387416$$

$$\text{Suka Cita (C3)} = \sqrt{\frac{((0,500 - 0,000)^2 + (0,750 - 0,250)^2 + (0,000 - 0,000)^2 + (0,000 - 0,189)^2)}{4}} = 0,313682508$$

Seterusnya hingga data ke 103 (Sumber makmur). Hasil dari keseluruhan dapat dilihat pada Tabel 5 sebagai berikut :

Tabel 5. Nilai Jarak Terdekat (*Cost*) Pada Iterasi Pertama

No	Cost 1	Cost 2	Cost 3	Kedekatan	Cluster
1	0	0,505871041	0,918925202	0	1
2	0,039464286	0,51323559	0,139387416	0,039464286	1
3	0,250713827	0,258044303	0,313682508	0,250713827	1
4	0,015357143	0,503812277	0,922023648	0,015357143	1
5	0,251353085	0,268016679	0,724280167	0,251353085	1
6	0,525192166	0,084642857	0,653274938	0,084642857	2
7	0,255256523	0,251355763	0,337897961	0,251355763	2
8	0,028571429	0,502409906	0,20715531	0,028571429	1
9	0,036071429	0,501767056	0,214636156	0,036071429	1
10	0,026785714	0,502579342	0,205374362	0,026785714	1
...
96	1,046329862	1,035984586	1,090894596	1,035984586	2
97	1,03187592	1,031218089	1,054668013	1,031218089	2
98	1,056818872	1,042712823	1,109588313	1,042712823	2
99	1,031031931	1,034934451	1,043230119	1,031031931	1
100	1,034792105	1,030877886	1,065010539	1,030877886	2

101	1,035605317	1,031041816	1,067291119	1,031041816	2
102	1,44200683	1,169914206	1,433349681	1,169914206	2
103	1,034792105	1,030877886	1,065010539	1,030877886	2
Total Kedekatan				63,94580235	

Setelah didapatkan hasil jarak dari setiap objek (*cost*) pada iterasi ke-1 maka lanjut ke iterasi ke-2. Inisialisasi kandidat medoid baru (*non-medoid*) pada iterasi ke-2 dapat dilihat pada Tabel berikut:

Tabel 6. Medoids Awal

No	Nama Kelompok Tani	Penyuluh	Komoditi Unggulan	Luas Darat Ha	Luas sawah Ha
4	Suka Dame	1	1	0,5	10,83
14	Sumber Makmur	1	3	0	5,34
91	Karya Indah	3	4	0	0

Maka perhitungan untuk menghitung jarak setiap objek dengan medoid awal Indah (C1) =

$$\sqrt{((0,000 - 0,000)^2 + (0,000 - 0,000)^2 + (0,015 - 0,015)^2 + (0,193 - 0,178)^2)} = 0,015357143$$

$$\text{Bersih (C2)} = \sqrt{((0,000 - 0,000)^2 + (0,000 - 0,000)^2 + (0,015 - 0,015)^2 + (0,193 - 0,139)^2)} = 0,054821429$$

$$\text{Suka Cita (C3)} = \sqrt{((0,000 - 0,000)^2 + (0,000 - 0,250)^2 + (0,015 - 0,000)^2 + (0,193 - 0,189)^2)} = 0,250484021$$

Seterusnya hingga data ke 103 (Sumber Makmur).

Sedangkan perhitungan untuk menghitung jarak setiap objek dengan medoid awal pada Cost 2 adalah sebagai berikut :

$$\text{Indah (C1)} = \sqrt{((0,000 - 0,000)^2 + (0,500 - 0,000)^2 + (0,000 - 0,015)^2 + (0,095 - 0,178)^2)} = 0,50701337$$

$$\text{Bersih (C2)} = \sqrt{((0,000 - 0,000)^2 + (0,500 - 0,000)^2 + (0,000 - 0,015)^2 + (0,095 - 0,139)^2)} = 0,052089918$$

$$\text{Suka Cita (C3)} = \sqrt{((0,000 - 0,000)^2 + (0,500 - 0,250)^2 + (0,000 - 0,000)^2 + (0,095 - 0,189)^2)} = 0,267125728$$

Seterusnya hingga data ke 103 (Sumber Makmur).

Sedangkan perhitungan untuk menghitung jarak setiap objek dengan medoid awal pada Cost 3 adalah sebagai berikut :

$$\text{Indah (C1)} = \sqrt{((1,000 - 0,000)^2 + (0,750 - 0,000)^2 + (0,000 - 0,015)^2 + (0,000 - 0,178)^2)} = 1,262704845$$

$$\text{Bersih (C2)} = \sqrt{((1,000 - 0,000)^2 + (0,750 - 0,000)^2 + (0,000 - 0,015)^2 + (0,000 - 0,139)^2)} = 1,257747531$$

$$\text{Suka Cita (C3)} = \sqrt{((1,000 - 0,000)^2 + (0,750 - 0,250)^2 + (0,000 - 0,000)^2 + (0,000 - 0,189)^2)} = 1,13397386$$

Seterusnya hingga data ke 103 (Sumber makmur). Hasil dari keseluruhan dapat dilihat pada Tabel 6 sebagai berikut :

Tabel 6. Nilai Jarak Terdekat (*Cost*) Pada Iterasi Kedua

No	Cost 1	Cost 2	Cost 3	Kedekatan	Cluster
1	0,015357143	0,50701337	1,262704845	0,50701337	2
2	0,054821429	0,502089918	1,257747531	0,502089918	2
3	0,250484021	0,267125728	1,13397386	1,026712573	2
4	0	0,509742888	1,264961505	1,509742888	2
5	0,253114389	0,25743634	1,128973764	1,25743638	2
6	0,520719086	0,242678571	1,084789447	0,242678571	2
7	0,252740893	0,282691657	1,140909739	1,282691657	2
8	0,013214286	0,51244841	1,267049061	1,51244841	2
9	0,020714286	0,514128752	1,268293609	1,514128752	2

No	Cost 1	Cost 2	Cost 3	Kedekatan	Cluster
10	0,011428571	0,512063708	1,266759104	1,512063708	2
..
96	1,043810746	1,06349977	0,614451805	1,043810746	1
97	1,031317666	1,038675815	0,547562432	1,031317666	1
98	1,053544216	1,077937905	0,647059676	1,053544216	1
99	1,031404237	1,032846337	0,525194327	1,031404237	1
100	1,03357232	1,04511064	0,567227863	1,03357232	1
101	0,19397837	1,046621386	0,571498322	0,19397837	1
102	1,034253924	1,14611705	0,342039923	0,342039923	3
103	1,034253924	1,14611705	0,342039923	0,342039923	3
Total Kedekatan				1075516140	

c) Hitung Total Simpangan (S).

Setelah didapatkan nilai jarak antara iterasi ke-1 dan iterasi ke-2, hitung total simpangan (S) dengan mencari selisih dari nilai total cost baru- nilai total cost lama. Dengan ketentuan jika $S < 0$, maka tukar nilai objek dengan menentukan medoid baru.

$S = \text{Total cost baru} - \text{Total cost lama}$

$= 63,94580235 - 1075516140 = 1075516076$

Karena nilai $S > 0$ maka iterasi dihentikan.

3.2.2. Proses Pengujian dengan RapidMiner

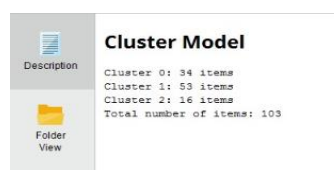
Pada tahap akhir penerapan algoritma K-Medoids dilakukan penyesuaian hasil perhitungan manual melalui pengujian menggunakan *software* RapidMiner 8.1. Pengujian terhadap hasil perhitungan manual menggunakan *Software* RapidMiner. Pada bagian ini akan menjelaskan tahapan-tahapan dalam k-Medoids pada data yang telah di import dalam rapidminer. Tahapan pertama pilih pada pojok atas bagian kiri tab operators lalu pilih “k-Medoids” dan ketika muncul hubungkan “Read excel dan k-Medoids” dan ubah nilai K sebanyak 2 *cluster*. Kemudian akan muncul tampilan seperti Gambar dibawah ini :



Gambar 4. Pemrosesan Data

3.3. Pembahasan

Setelah melakukan pemrosesan data selanjutnya untuk mendapatkan sebuah hasil pengelompokkan maka pada tahap selanjutnya yaitu meng-klik tanda *play* panah biru pada bagian *toolbar* atas untuk menjalankan pemrosesan data sehingga mendapatkan hasil akhir dalam sebuah keluaran sistem. Berikut tampilan hasil akhir dari pemrosesan data diatas :

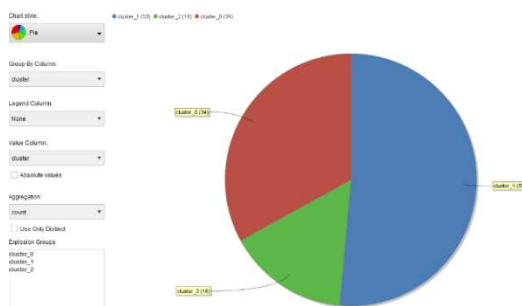


Gambar 5. Tampilan Cluster Model

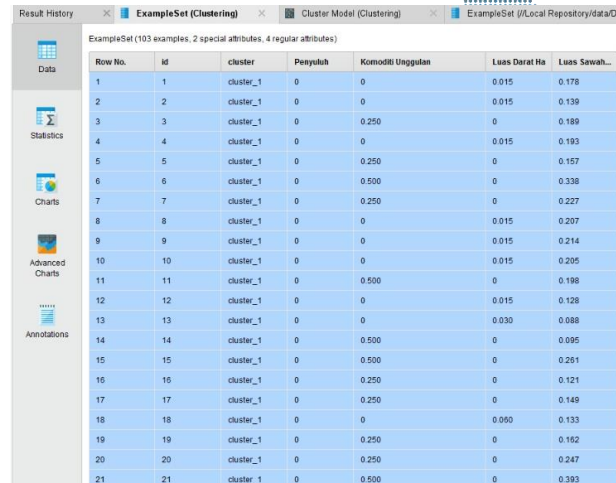
Keterangan :

1. Jumlah Cluster 0 berjumlah 34 *items*
2. Jumlah Cluster 1 berjumlah 53 *items*
3. Jumlah Cluster 2 berjumlah 16 *items*
4. Jumlah keseluruhan *items* adalah 103

Sehingga dapat diketahui hasil pengelompokkan dari *Rapidminer* 8.1 dapat dilihat dari Gambar dibawah ini :



Gambar 6. Grafik Hasil Akhir Pengelompokkan Rapidminer



Row No.	Id	cluster	Penyuluh	Komoditi Unggulan	Luas Darat Ha	Luas Sawah...
1	1	cluster_1	0	0	0.015	0.178
2	2	cluster_1	0	0	0.015	0.139
3	3	cluster_1	0	0.250	0	0.189
4	4	cluster_1	0	0	0.015	0.193
5	5	cluster_1	0	0.250	0	0.157
6	6	cluster_1	0	0.500	0	0.338
7	7	cluster_1	0	0.250	0	0.227
8	8	cluster_1	0	0	0.015	0.207
9	9	cluster_1	0	0	0.015	0.214
10	10	cluster_1	0	0	0.015	0.205
11	11	cluster_1	0	0.500	0	0.198
12	12	cluster_1	0	0	0.015	0.128
13	13	cluster_1	0	0	0.030	0.088
14	14	cluster_1	0	0.500	0	0.095
15	15	cluster_1	0	0.500	0	0.251
16	16	cluster_1	0	0.250	0	0.121
17	17	cluster_1	0	0.250	0	0.149
18	18	cluster_1	0	0	0.090	0.133
19	19	cluster_1	0	0.250	0	0.162
20	20	cluster_1	0	0.250	0	0.247
21	21	cluster_1	0	0.500	0	0.393

Gambar 6. Hasil Akhir Pengelompokkan RapidMiner

Pembahasan posisi *cluster* menggunakan *software* rapidminer dapat di nilai bahwa hasil akurasi nilai yang terdapat di perhitungan manual excel dan aplikasi *software* rapidminer dihasilkan dengan tingkat kesamaannya adalah 100% selain itu nilai tersebut tetap sama pada proses perhitungan manual maupun aplikasi.

4. KESIMPULAN

Hasil akhir dari penelitian yang menggunakan data sebanyak 103 data kelompok tani ini, dapat disimpulkan bahwa telah didapatkan masing-masing nilai *cluster* yakni *Cluster* tertinggi (C1) dengan jumlah sebanyak 34 Kelompok Tani, dengan rata-rata penyuluh Sahnun, SP, rata-rata komoditi unggulan padi sawah, serta rata-rata luas lahan darat perkebunan 0 Ha, dan rata-rata luas sawah 22,76 Ha. *Cluster* menengah (C2) dengan jumlah sebanyak 53 Kelompok Tani, dengan rata-rata penyuluh Agus Edi Santoso, SP, rata-rata komoditi unggulan padi sawah, serta rata-rata luas lahan darat perkebunan 0,1603 Ha, dan rata-rata luas sawah 8,47 Ha. *Cluster* Rendah (C3) dengan jumlah sebanyak 16 Kelompok Tani, dengan rata-rata penyuluh Mesdi, SP, rata-rata komoditi unggulan peternakan, serta rata-rata luas lahan darat perkebunan 3,52Ha, dan rata-rata luas sawah 4,09 Ha. Proses pemberentihan iterasi pada pengujian yang dilakukan pada penelitian ini yaitu terjadi pada iterasi ke 3. Nilai hasil akurasi yang dilakukan dengan perhitungan manual dan dengan aplikasi rapidminer bernilai sama.

REFERENCES

- [1] S. Darma And G. W. Nurcahyo, "Klasterisasi Teknik Promosi Dalam Meningkatkan Mutu Kampus Menggunakan Algoritma K-Medoids," *J. Inform. Ekon. Bisnis*, Vol. 3, Pp. 89–94, 2021, Doi: 10.37034/InfEb.V3i3.87.
- [2] P. Alkhairi And Z. Situmorang, "Penerapan Data Mining Untuk Menganalisis Kepuasan Pegawai Terhadap Pelayanan Bidang Sdm Dengan Algoritma C4.5," *Jurasik (Jurnal Ris. Sist. Inf. Dan Tek. Inform.)*, Vol. 7, No. 1, P. 40, 2022, Doi: 10.30645/Jurasik.V7i1.414.
- [3] R. U. B. Barus, I. Gunawan, B. E. Damanik, I. Parlina, And W. Saputra, "Pengelompokan Data Penjualan Mie Berdasarkan Bulan Dengan Menggunakan Algoritma K-Medoids," *J. Ilmu Komput. Dan Inform.*, Vol. 1, No. 2, Pp. 141–156, 2021, Doi: 10.54082/Jiki.15.
- [4] P. Alkhairi, P. P. P. A. N. . F. I. R.H.Zer, E. R. Batubara, F. N. Tambunan, And R. Rosnelly, "Pengenalan Pola Kemampuan Pelanggan Dalam Membayar Air Pdam Menggunakan Algoritma Naive Bayes," *Jurnaltimes*, Vol. X, No. 2, Pp. 29–38, 2022.
- [5] G. Dwilestari, Mulyawan, Martanto, And I. Ali, "Analisis *Clustering* Menggunakan K-Medoid Pada Data Penduduk Miskin Indonesia," *Jursima J. Sist. Inf. Dan Manaj.*, Vol. 9, No. 3, Pp. 282–290, 2021.
- [6] W. Utomo, "The Comparison Of K-Means And K-Medoids Algorithms For *Clustering* The Spread Of The Covid-19 Outbreak In Indonesia," *Ilk. J. Ilm.*, Vol. 13, No. 1, Pp. 31–35, 2021, Doi: 10.33096/Ilkom.V13i1.763.31-35.
- [7] S. Harlina, "Data Mining Pada Penentuan Kelayakan Kredit Menggunakan Algoritma K-Nn Berbasis Forward Selection Data Mining On Credit Feasibility Determination Using K-Nn Algorithm Based On Forward Selection," *Ccit J.*, Vol. 11, No. 2, Pp. 236–244, 2018, Doi: 10.33050/Ccit.V11i2.591.
- [8] D. R. Agustian And B. A. Darmawan, "Analisis *Clustering* Demam Berdarah Dengue Dengan Algoritma K-Medoids (Studi Kasus Kabupaten Karawang)," *Jiko (Jurnal Inform. Dan Komputer)*, Vol. 6, No. 1, P. 18, 2022, Doi: 10.26798/Jiko.V6i1.504.
- [9] A. Supriyadi, A. Triayudi, And I. D. Sholihati, "Perbandingan Algoritma K-Means Dengan K-Medoids Pada Pengelompokan Armada Kendaraan Truk Berdasarkan Produktivitas," *Jipi (Jurnal Ilm. Penelit. Dan Pembelajaran Inform.)*, Vol. 6, No. 2, Pp. 229–240, 2021, Doi: 10.29100/Jipi.V6i2.2008.
- [10] A. Fira, C. Rozikin, And G. Garo, "Komparasi Algoritma K-Means Dan K-Medoids Untuk Pengelompokan Penyebaran Covid-19 Di Indonesia," *J. Appl. Informatics Comput.*, Vol. 5, No. 2, Pp. 133–138, 2021, Doi: 10.30871/Jaic.V5i2.3286.
- [11] F. D. R. Sari And S. P. Ediwijojo, "Pemetaan Tingkat Kemiskinan Di Provinsi Jawa Tengah Berdasarkan Kabupaten/Kota Dengan Metode K-Medoids," *J. Media Inform. Budidarma*, Vol. 5, No. 4, P. 1528, 2021, Doi: 10.30865/Mib.V5i4.3278.

- [12] M. N. P. Pamulang, M. N. Aini, And U. Enri3, “Komparasi Distance Measure Pada K-Medoids *Clustering* Untuk Pengelompokkan Penyakit Ispa,” *Edumatic J. Pendidik. Inform.*, Vol. 5, No. 1, Pp. 99–107, 2021, Doi: 10.29408/Edumatic.V5i1.3359.
- [13] C. Sukmayadi, A. Primajaya, And I. Maulana, “Penerapan Algoritma K-Medoids Dalam Menentukan Daerah Rawan Banjir Di Kabupaten Karawang,” *Informal Informatics J.*, Vol. 6, No. 3, P. 187, 2021, Doi: 10.19184/Isj.V6i3.25423.
- [14] F. Alfiah, A. Almadayani, D. Al Farizi, And E. Widodo, “Analisis *Clustering* K-Medoids Berdasarkan Indikator Kemiskinan Di Jawa Timur Tahun 2020,” *J. Ilm. Sains*, Vol. 22, No. 1, P. 1, 2021, Doi: 10.35799/Jis.V22i1.35911.
- [15] F. Fajriana, “Analisis Algoritma K-Medoids Pada Sistem Klasterisasi Produksi Perikanan Tangkap Kabupaten Aceh Utara,” *J. Edukasi Dan Penelit. Inform.*, Vol. 7, No. 2, P. 263, 2021, Doi: 10.26418/Jp.V7i2.47795.
- [16] A. Aditya, B. N. Sari, And T. N. Padilah, “Comparison Analysis Of Euclidean And Gower Distance Measures On K-Medoids *Cluster*,” *J. Teknol. Dan Sist. Komput.*, Vol. 9, No. 1, Pp. 1–7, 2021, Doi: 10.14710/Jtsiskom.2020.13747.