

Clustering Petani Penerima Pupuk Berdasarkan Luas Lahan Dengan Menggunakan Algoritma *K-Means* Dan *K-Medoids*

Eri Sasmita Susanto^{1*}, Novitasari², Yudi Mulyanto³, Fahri Hamdani⁴

Fakultas Rekayasa Sistem, Program Studi Informatika, Universitas Teknologi Sumbawa, Sumbawa, Indonesia

¹Eri.sasmita.susanto@uts.ac.id, ²nopitas850@gmail.com, ³yudi.mulyanto@uts.ac.id, ⁴fahri.hamdani@uts.ac.id



Histori Artikel:

Diajukan: 10 Juli 2024

Disetujui: 7 September 2024

Dipublikasi: 30 September 2024

Kata Kunci:

Clustering; K-Means; K-Medoids; Penerima Pupuk; Pertanian; Davies-Bouldin Index.

Digital Transformation Technology (Digitech) is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).

Abstrak

Di Kecamatan Moyo Hulu, banyak kelompok tani yang tidak mendapatkan pupuk sesuai dengan luas lahan yang dimiliki, meskipun pupuk selalu didistribusikan setiap tahunnya. Hal ini menyebabkan ketidakpuasan di kalangan petani, yang kemudian membeli pupuk dari pihak ketiga dengan harga yang sangat tinggi. Masalah ini sering terjadi karena kurangnya transparansi antara ketua kelompok tani dan anggotanya. Penelitian ini bertujuan untuk mengevaluasi efektivitas penyaluran pemerataan pupuk berdasarkan luas lahan dengan menggunakan algoritma *K-Means* dan *K-Medoids* pada kelompok tani di wilayah Kecamatan Moyo Hulu. Penelitian ini menggunakan metode pengumpulan data kuantitatif dan metode pengolahan data mining dengan algoritma K-Means dan K-Medoids dalam proses clustering dengan menggunakan indeks Davies-Bouldin untuk pengujian validitas. Algoritma K-Means menghasilkan nilai Davies-Bouldin sebesar 0,554, sementara algoritma K-Medoids menghasilkan nilai Davies-Bouldin sebesar -1,580 dan -0,878. Berdasarkan nilai Davies-Bouldin, algoritma K-Means menunjukkan performa yang lebih baik dalam melakukan pengelompokan data petani penerima pupuk di Kecamatan Moyo Hulu dibandingkan dengan algoritma K-Medoids.

PENDAHULUAN

Pertanian merupakan kegiatan pemanfaatan sumber daya hayati yang dilakukan manusia untuk menghasilkan bahan pangan, bahan baku industri, sumber energi, serta untuk mengelola lingkungan hidupnya. Kegiatan ini termasuk budidaya tanaman (crop cultivation) dan pembesaran hewan ternak (raising). (Purba et al., 2020). Menurut Badan Pusat Statistik (BPS), jumlah petani pada tahun 2021 mencapai 38,8 juta orang dari total angkatan kerja yang ada di Indonesia yang sekitar 140,15 juta orang. Jumlah petani yang besar ini menunjukkan bahwa sektor pertanian, khususnya di Kabupaten Sumbawa, berkontribusi signifikan terhadap perekonomian masyarakat setempat. Mata pencaharian utama masyarakat Sumbawa adalah bertani dan beternak, dengan padi dan palawija sebagai bahan pokok utama yang tidak tergantikan. (Sumbawa, 2020).

Di Kecamatan Moyo Hulu, banyak kelompok tani yang tidak mendapatkan pupuk sesuai dengan luas lahan yang dimiliki, meskipun pupuk selalu didistribusikan setiap tahunnya. Hal ini menyebabkan ketidakpuasan di kalangan petani, yang kemudian membeli pupuk dari pihak ketiga dengan harga yang sangat tinggi.

Salah satu metode yang digunakan pada data mining adalah clustering yang berfungsi untuk mengelompokkan data. Teknik clustering ini digunakan untuk melakukan pengelompokan data yang menghasilkan data petani penerima pupuk yang memiliki hasil rendah, sedang dan tinggi. Algoritma K-Means dan K-Medoids adalah algoritma clustering yang dapat digunakan untuk melakukan pengelompokan data kedalam cluster dengan karakteristik yang mirip atau sama. Dalam konteks ini, teknik data mining dengan algoritma clustering K-Means dan K-Medoids dapat digunakan untuk menyelesaikan masalah distribusi pupuk kepada petani berdasarkan luas lahan yang dimiliki. (Marlina et al., 2018). Berdasarkan permasalahan yang telah diuraikan, penulis tertarik untuk mengangkat materi skripsi dengan judul “Clustering Petani Penerima Pupuk Berdasarkan Luas Lahan Dengan Menggunakan Algoritma K-Means Dan K-Medoids (Studi Kasus: Kecamatan Moyo Hulu)”.

STUDI LITERATUR

Setelah dilakukan tinjauan pustaka dengan penelitian yang dilakukan, berikut adalah beberapa penelitian terdahulu yang memiliki keterkaitan sebagai pendukung penelitian yang diteliti: Penelitian pertama yang dilakukan oleh (Kholila et al., 2023) berjudul “Pemetaan Kondisi Lingkungan Tanam Menggunakan K-Means Clustering” memiliki tujuan untuk memberikan rekomendasi tindakan pertanian yang tepat kepada petani berdasarkan kondisi lingkungan tanam. Penelitian kedua yang dilakukan oleh (Gustrianda & Mulyana, 2022)

berjudul “Penerapan Data Mining Dalam Pemilihan Produk Unggulan Dengan Metode Algoritma K-Means Dan K-Medoids”. Bertujuan untuk membantu perusahaan dalam menentukan produk unggulan menggunakan data mining. .

Penelitian ketiga yang dilakukan oleh (Rohmah & Sembiring, 2021) berjudul “Implementasi Algoritma K-Means Clustering Analysis Untuk Menentukan Hambatan Pembelajaran Daring (Studi Kasus: SMK Yaspim Gegerbitung)” bertujuan untuk membantu mengidentifikasi hambatan-hambatan dalam proses pembelajaran daring di SMK Yaspim Gegerbitung.

Penelitian keempat yang dilakukan oleh (Khakim et al., 2020) berjudul “Analisa Jaringan Dokumentasi dan Informasi Hukum Kementerian Tentang Teknologi Informasi Menggunakan Metode K-Means Clustering” bertujuan untuk melakukan clustering terhadap Jaringan Dokumentasi dan Informasi Hukum (JDIH) dari masing-masing kementerian berdasarkan kata kunci peraturan yang terkait dengan teknologi informasi.

Penelitian kelima dilakukan oleh (Ningsih et al., 2019) berjudul “Pengklastran Dokumen Tentang Dispensasi Nikah Menggunakan Algoritma K-Medoids” memiliki tujuan untuk mengelompokkan dokumen dispensasi nikah menggunakan algoritma K-Medoids.

Pertanian merupakan kegiatan pemanfaatan sumber daya hayati yang dilakukan manusia untuk menghasilkan bahan pangan, bahan mentah untuk industri, sumber energi, serta untuk mengelola lingkungan hidupnya. (Purba et al., 2020).

Lahan pertanian adalah penentu dari komoditas pertanian. Secara umum, semakin luasnya lahan yang dimiliki maka hasil produksi yang di dapatkan juga semakin besar. Luas lahan yang dimiliki oleh petani di setiap daerah berbeda-beda. (Nurmala, 2012).

Pupuk digunakan untuk meningkatkan hasil produksi karena memiliki kandungan unsur hara esensial yang dibutuhkan oleh tanaman, seperti nitrogen, kalium, fosfor dan juga magnesium (Mansyur et al., 2021).

Data merupakan suatu bahan mentah yang memerlukan proses pengolahan untuk mendapatkan suatu informasi atau keterangan baik itu kualitatif maupun kuantitatif sehingga menghasilkan suatu fakta yang memberikan manfaat kepada peneliti atau memberikan gambaran tentang suatu kondisi ataupun keadaan (Wahono & Ali, 2021).

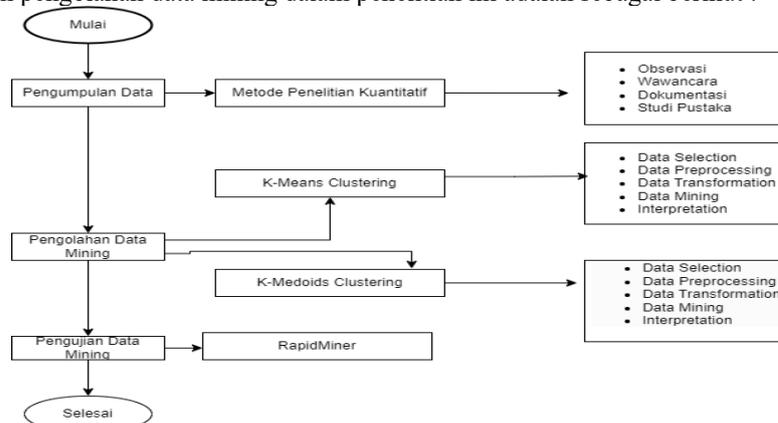
K-Means adalah algoritma yang digunakan dalam proses clustering non-hierarki untuk memartisi data ke dalam bentuk cluster. Algoritma ini memartisi data ke cluster agar data yang memiliki karakteristik yang sama di kelompokkan ke cluster yang sama dan data yang memiliki karakteristik yang berbeda dikelompokkan ke dalam cluster yang berbeda. Bentuk esensial dari k-means banyak ditemukan oleh beberapa ahli ternama seperti Lloyd (1982), Forgey (1965), Friedman dan Rubin (1967) dan McQueen (1967). Algoritma k-means berkembang menjadi lebih besar sehingga dinamakan dengan algoritma hill climbing, sesuai dengan pendapat Gray dan Nuhoff (1988) (Jollyta et al., 2020).

Davies Bouldin Indeks atau DBI juga disebut sebagai reliabilitas klasifikasi indeks. Davies Boulden Index pertama kali diperkenalkan oleh David L. Davies dan Donald W. Bouldin pada tahun 1979 sebagai rasio dari rata-rata jarak dalam antar kluster untuk setiap kluster terdekat. Davies Bouldin Index merupakan evaluasi dari kinerja pengelompokan (Sopyan et al., 2022).

METODE

Metode Penelitian

Pada penulisan ini penulis menggunakan dua metode yaitu metode pengumpulan data dan pengolahan data mining. Jenis penulisan ini menggunakan metode kuantitatif. Adapun tahapan yang digunakan dalam pengumpulan data dan pengolahan data mining dalam penelitian ini adalah sebagai berikut :



Gambar 1. Alur Metode Penulisan

a) Metode Pengumpulan Data

Metode atau teknik pengumpulan data adalah teknik atau cara yang digunakan dalam mengumpulkan data. Dalam penulisan ini, penulis menggunakan data primer dan data sekunder. Adapun tahapan yang digunakan dalam metode pengumpulan data yaitu observasi, wawancara, dokumentasi dan studi pustaka.

1. Observasi

Pada metode pertama penulis melakukan pengamatan terhadap alur dan pengolahan data petani sesuai dengan luas lahan di Kantor Dinas Pertanian Sumbawa.

2. Wawancara

Wawancara merupakan kegiatan yang digunakan untuk mendapatkan informasi dengan melakukan proses tanya jawab dengan narasumber. Dalam hal ini penulis melakukan wawancara bersama ibu Nuraeni, S.P. selaku bagian yang menangani pupuk di Dinas Pertanian dan langsung diarahkan ke BPP Kecamatan Moyo Hulu untuk mendapatkan data yang lebih terperinci sesuai dengan data yang dibutuhkan.

3. Dokumentasi

Metode dokumentasi adalah teknik yang digunakan dalam penulisan yang menggunakan dokumen sebagai sumber data dari sebuah instansi. Pengumpulan informasi dan dokumen dalam penulisan ini penulis menggunakan data khususnya di Kecamatan Moyo Hulu yaitu data di RDKK dengan menggunakan Erdkk.

4. Studi Pustaka

Untuk mendapatkan data atau informasi yang terkait dengan masalah yang dihadapi penulis mencari sumber data atau informasi berupa jurnal, buku dan laporan sebelumnya untuk dijadikan referensi dalam penulisan selanjutnya.

b) Metode Pengolahan Data Mining

Pada proses pengolahan data mining penulis menggunakan tahapan dalam *Knowledge Discovery in Database* (KDD) yaitu untuk menghasilkan sebuah informasi. Adapun tahapan yang digunakan dalam proses KDD yaitu data, selection, preprocessing, transformation dan data mining.

1. Data

Penulis mengumpulkan data yang akan digunakan terlebih dahulu sebelum data itu diolah dan dilakukannya pemilihan data.

2. Selection

Penulis melakukan pemilihan data yang sesuai dari data kelompok tani berdasarkan atribut yang terdiri dari tahun, nama kelompok tani, jumlah luas lahan yang dimiliki, jumlah pupuk yang diterima dan hasil yang didapatkan untuk dilakukan seleksi atau segmentasi data.

3. Preprocessing

Penulis melakukan pembersihan data, dimana *field* yang dianggap tidak sesuai dengan kebutuhan penulisan dapat dihapus. Selain itu, format data diubah kembali untuk memastikan keseragaman format yang konsisten. Seperti mengubah data nama kelompok tani menjadi bilangan angka sehingga mempermudah perhitungannya.

4. Transformation

Proses yang mengubah data agar dapat digunakan dan ditelusuri oleh *tools* rapidminer. Dengan kata lain, transformasi data melibatkan pemetaan data yang rumit.

5. Data Mining

Penulis melakukan proses pengambilan pola dari data yang tersedia dengan bantuan *tools rapidminer* menggunakan algoritma *K-Means* dan *K-Medoids Clustering*.

HASIL

Data yang digunakan pada penelitian ini diperoleh dengan menggunakan metode dokumentasi. Data yang digunakan berasal dari staff Badan Penyuluhan Pertanian (BPP) Kecamatan Moyo Hulu, yang mencakup data kelompok tani dari tahun 2020-2023. Dataset petani penerima pupuk di Kecamatan Moyo Hulu berjumlah sebanyak 363 kelompok tani yang diperlihatkan pada Tabel 1.

Tabel 1 Data Petani Penerima Pupuk Kec. Moyo Hulu Tahun 2020-2023

| Nama Kelompok Tani | Luas Lahan (Ha) | | | | Jumlah Pupuk Yang Diterima (Kg) | | | | Hasil Produk |
|--------------------|-----------------|----|----|------|---------------------------------|-------|--------|------------|--------------|
| Ai Keban | 22 | 22 | 20 | 26,5 | 12.100 | 8.250 | 9.500 | 7.574 ... | 159 |
| Asar Edas | 23 | 25 | 26 | 37,5 | 12.650 | 9.225 | 12.350 | 10.718 ... | 150 |

| | | | | | | | | | | |
|--------------|-------|-------|------|-------|--------|--------|--------|--------|-----|-----|
| Buin Bage | 26 | 26 | 26 | 35,5 | 14.300 | 10.450 | 12.350 | 11.538 | ... | 142 |
| Buya Rib | 26 | 26 | 62 | 58 | 14.300 | 10.450 | 34.100 | 18.718 | ... | 232 |
| Lenang Suma | 23 | 23,5 | 23,5 | 24 | 12.925 | 8.810 | 10.710 | 6.860 | ... | 168 |
| Lutuk Peti | 23 | 23 | 33 | 25 | 12.650 | 8.625 | 15.675 | 7.146 | ... | 100 |
| | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| Unter Salolo | 17,75 | 19,85 | 22,2 | 21,45 | 9.758 | 7.432 | 10.526 | 6.126 | | 86 |

PEMBAHASAN

Data yang digunakan dalam melakukan penelitian ini yaitu gabungan data petani dari tahun 2020-2023 yang diperoleh berdasarkan dokumentasi dari staff BPP (Badan Penyuluhan Pertanian) Kec. Moyo Hulu. Jumlah seluruh kelompok tani terdapat sebanyak 363 kelompok yang terdiri dari 12 desa. Adapun atribut yang digunakan yaitu nama kelompok tani, luas lahan yang dimiliki, jumlah pupuk yang diterima dan hasil produksi.

Data Selection

Data yang digunakan dalam melakukan penelitian ini yaitu gabungan data petani dari tahun 2020-2023 yang diperoleh berdasarkan dokumentasi dari staff BPP (Badan Penyuluhan Pertanian) Kec. Moyo Hulu. Jumlah seluruh kelompok tani terdapat sebanyak 363 kelompok yang terdiri dari 12 desa.

Adapun atribut yang digunakan yaitu nama kelompok tani, luas lahan yang dimiliki, jumlah pupuk yang diterima dan hasil produksi.

Tabel 2 Data Selection

| Nama Kelompok Tani | Luas Lahan (Ha) | | | | Jumlah Pupuk Yang Diterima (Kg) | | | | Hasil Produksi | |
|--------------------|-----------------|-------|------|-------|---------------------------------|--------|--------|--------|----------------|-----|
| | | | | | | | | | | |
| Ai Keban | 22 | 22 | 20 | 26,5 | 12.100 | 8.250 | 9.500 | 7.574 | ... | 159 |
| Asar Edas | 23 | 25 | 26 | 37,5 | 12.650 | 9.225 | 12.350 | 10.718 | ... | 150 |
| Buin Bage | 26 | 26 | 26 | 35,5 | 14.300 | 10.450 | 12.350 | 11.538 | ... | 142 |
| Buya Rib | 26 | 26 | 62 | 58 | 14.300 | 10.450 | 34.100 | 18.718 | ... | 232 |
| Lenang Suma | 23 | 23,5 | 23,5 | 24 | 12.925 | 8.810 | 10.710 | 6.860 | ... | 168 |
| Lutuk Peti | 23 | 23 | 33 | 25 | 12.650 | 8.625 | 15.675 | 7.146 | ... | 100 |
| | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| Unter Salolo | 17,75 | 19,85 | 22,2 | 21,45 | 9.758 | 7.432 | 10.526 | 6.126 | | 86 |

Data Preprocessing

Dalam proses preprocessing, terdapat langkah-langkah yang digunakan untuk membersihkan data. Data yang telah dipilih akan divalidasi dan dikonfigurasi ulang agar formatnya konsisten dan tidak ada missing. Pada tabel 4 diperlihatkan data preprocessing.

Tabel 3 Data Preprocessing

| Nama Kelompok Tani | Luas Lahan (Ha) | | | | Jumlah Pupuk Yang Diterima (Kg) | | | | Hasil Produksi | |
|--------------------|-----------------|----|----|------|---------------------------------|-------|--------|--------|----------------|-----|
| | | | | | | | | | | |
| Ai Keban | 22 | 22 | 20 | 26,5 | 12.100 | 8.250 | 9.500 | 7.574 | ... | 159 |
| Asar Edas | 23 | 25 | 26 | 37,5 | 12.650 | 9.225 | 12.350 | 10.718 | ... | 150 |

| | | | | | | | | | | |
|--------------|-------|-------|------|-------|--------|--------|--------|--------|-----|-----|
| Buin Bage | 26 | 26 | 26 | 35,5 | 14.300 | 10.450 | 12.350 | 11.538 | ... | 142 |
| Buya Rib | 26 | 26 | 62 | 58 | 14.300 | 10.450 | 34.100 | 18.718 | ... | 232 |
| Lenang Suma | 23 | 23,5 | 23,5 | 24 | 12.925 | 8.810 | 10.710 | 6.860 | ... | 168 |
| Lutuk Peti | 23 | 23 | 33 | 25 | 12.650 | 8.625 | 15.675 | 7.146 | ... | 100 |
| | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| Unter Salolo | 17,75 | 19,85 | 22,2 | 21,45 | 9.758 | 7.432 | 10.526 | 6.126 | | 86 |

Data Transformasi

Hasil transformasi data yang dihasilkan sebanyak 363 dataset yang akan diproses menggunakan algoritma K-Means dan K-Medoids Clustering. Transformasi dilakukan dengan melakukan tahap seleksi atribut, dimana pada tabel 4 diketahui bahwa data luas lahan (Ha), jumlah pupuk yang diterima (Kg) dan hasil produksi masih dalam bentuk data terpisah. Ketiga data tersebut akan diintegrasikan menjadi satu data sehingga menghasilkan data baru yang selanjutnya disebut data luas lahan total, data jumlah pupuk total dan data hasil total. Pada tahap transformasi data akan dihasilkan empat variable.

Tabel 4 Data Transformation

| Nama Kelompok Tani | LL Total (Ha) | Jml Ppk Total (Kg) | Hasil (Ton) |
|--------------------|---------------|--------------------|-------------|
| Ai Keban | 90,5 | 37.424 | 687 |
| Asar Edas | 111,5 | 44.943 | 556 |
| Buin Bage | 113,5 | 48.638 | 568 |
| Buya Rib | 172 | 77.568 | 859 |
| Lenang Suma | 94,5 | 39.305 | 757,5 |
| Lutuk Peti | 104 | 44.096 | 535 |
| ... | ... | ... | ... |
| Unter Salolo | 81,25 | 33.842 | 414 |

Model clustering yang digunakan pada segmentasi ini adalah metode k-means dan k-medoids, yaitu membagi data petani menjadi beberapa bagian berdasarkan dari 3 variabel yaitu luas lahan (X1), jumlah pupuk (X2), dan hasil (X3). Adapun langkah-langkah yang digunakan dalam proses k-means yaitu:

- a. Menentukan jumlah cluster yang diinginkan. Penulis dalam penelitian ini akan mengelompokkan data kedalam tiga cluster yaitu tinggi, sedang, dan rendah.
- b. Dalam proses penentuan titik pusat awal cluster yang diinginkan dari data tabel 5 yaitu akan dibentuk 3 cluster (**k=3**). Penulis dalam penelitian ini menentukan titik centroid dengan cara acak atau manual.

Tabel 5 Data Pusat Cluster Awal

| Clustering | LL (Ha) | Total | Jml Ppk Total (Kg) | Hasil (Ton) |
|------------|---------|-------|--------------------|-------------|
| C1 | 90,5 | | 37.424 | 687 |
| C2 | 111,5 | | 44.943 | 556 |
| C3 | 172 | | 77.568 | 859 |

- c. Setelah titik pusat awal cluster sudah ditentukan selanjutnya dilakukan perhitungan jarak antara setiap data dengan titik pusat setiap cluster dengan menggunakan rumus Euclidean Distance, adalah sebagai berikut:

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

Perhitungan jarak data pertama dengan pusat cluster pertama adalah:

$$D(1, 1) = \sqrt{\frac{(90,5 - 90,5)^2 + (37.424 - 37.424)^2}{(687 - 687)^2}} = 0$$

Perhitungan jarak data pertama dengan pusat kluster kedua adalah:

$$D(1, 2) = \sqrt{\frac{(90,5 - 111,5)^2 + (37.424 - 44.943)^2}{(687 - 556)^2}} = 7520,17$$

Perhitungan jarak data pertama dengan pusat kluster ketiga adalah:

$$D(1,3) = \sqrt{\frac{(90,5 - 172)^2 + (37.424 - 77.568)^2}{(687 - 859)^2}} = 40144,45$$

- d. Langkah berikutnya adalah menghitung kembali titik pusat dari cluster berdasarkan keanggotaan cluster.
 e. Langkah berikutnya adalah menugaskan kembali setiap objek dengan menggunakan titik pusat cluster baru. Jika tidak terjadi perubahan pada rasio maka clustering dianggap selesai.

Setelah proses K-Means selesai langkah selanjutnya yaitu melakukan proses pengelompokan data dengan algoritma K-Medoids. Proses penentuan cluster pada algoritma K-Medoids sama dengan algoritma K-Means yaitu dengan menggunakan rumus Euclidien Distance. Adapun langkah-langkah yang digunakan pada proses K-Medoids adalah sebagai berikut:

- a. Menentukan jumlah cluster yang diinginkan. Penulis dalam penelitian ini akan mengelompokkan data kedalam tiga cluster yaitu tinggi, sedang, dan rendah.
- b. Dalam proses penentuan titik pusat awal cluster yang diinginkan dari data tabel 4.5 yaitu akan dibentuk 3 cluster (**k=3**).
- b. Dalam proses penentuan titik pusat awal cluster yang diinginkan dari data tabel 4.5 yaitu akan dibentuk 3 cluster (**k=3**).

Tabel 6 Data Pusat Cluster Awal

| Clustering | LL Total (Ha) | Jml Ppk Total (Kg) | Hasil (Ton) |
|------------|---------------|--------------------|-------------|
| C1 | 253,4 | 106.000 | 1896 |
| C2 | 104,1 | 44.197 | 537 |
| C3 | 81,25 | 33.842 | 414 |

- c. Setelah titik pusat awal cluster sudah ditentukan selanjutnya dilakukan perhitungan jarak antara setiap data dengan titik pusat setiap cluster. Data yang memiliki jarak terdekat akan menentukan apakah data tersebut masuk ke dalam cluster tersebut. Untuk menghitung jarak antar semua data dengan titik pusat cluster, maka rumus yang digunakan yaitu Euclidean Distance, adalah sebagai berikut:

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

Perhitungan jarak data pertama dengan pusat cluster pertama adalah:

$$D(1, 1) = \sqrt{\frac{(90,5 - 253,4)^2 + (37.424 - 106.000)^2}{(687 - 1896)^2}} = 68586,85$$

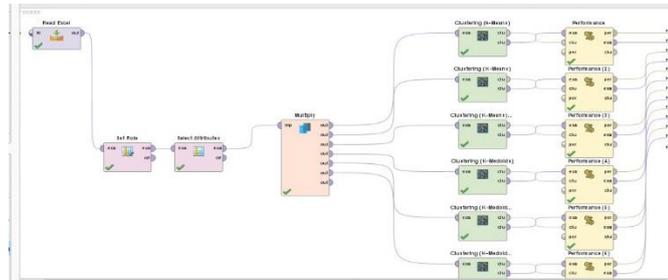
Perhitungan jarak data pertama dengan pusat kluster kedua adalah:

$$D(1, 2) = \sqrt{\frac{(90,5 - 104,1)^2 + (37.424 - 44.197)^2}{(687 - 537)^2}} = 6774,67$$

Perhitungan jarak data pertama dengan pusat kluster ketiga adalah:

$$D(1,3) = \sqrt{(90,5 - 81,25)^2 + (37.424 - 33.842)^2 + (687 - 414)^2} = 3592,40$$

Dalam proses pengelompokan data petani, penulis menggunakan algoritma K-Means dan K-Medoids untuk mengelompokkan data berdasarkan dengan atribut dan jarak pusat cluster, seperti yang ada pada tabel 4.6. Iterasi adalah proses eksekusi yang digunakan pada algoritma K-Means dan K-Medoids untuk mengelompokkan data berdasarkan pada jarak pusatnya. Nilai jarak pusat cluster akan terus berubah sesuai dengan iterasi yang dilakukan.



Gambar 3 Desain penggunaan algoritma K-Means dan K-Medoids

Pada gambar 3 dalam penjelasan model excel pertama, data excel kemudian masuk ke tahapan pelebelan data (Set Role) yang akan digunakan, setelah proses pelebelan data kemudian melalui proses penyeleksian atribut sesuai dengan atribut yang ingin diolah. Semua parameter yang diperlukan disimpan dalam objek multiply karena data yang diolah menggunakan dua algoritma sehingga model multiply sangat cocok untuk digunakan. Dalam penelitian ini, penulis menggunakan sampel data kelompok tani sebanyak 363 kelompok dengan 3 atribut yaitu Luas Lahan Total (LLT), Jumlah Pupuk Total (JPT), Hasil Produksi (HP).

Tabel 7 Nilai Cluster K-Means Pada Iterasi Terakhir

| Attribut | C1 | C2 | C3 |
|----------|----------|----------|----------|
| NKT1 | 6318,162 | 10291,68 | 38853,47 |
| NKT2 | 13831,49 | 2772,633 | 31336,49 |
| NKT3 | 17526,4 | 923,2511 | 27641,77 |
| NKT4 | 46457,93 | 29853,9 | 1302,976 |
| NKT5 | 8200,437 | 8411,791 | 8200,437 |
| NKT6 | 12984,29 | 3619,874 | 3619,874 |
| ... | ... | ... | ... |
| NKT363 | 2729,593 | 13874,56 | 42438,42 |

Hasil proses k-means clustering menghasilkan tiga model segmentasi seperti yang terdapat pada tabel 9 hasil dari segmentasi tersebut memberikan informasi bahwa segmen pertama atau cluster 0 terdiri dari 188 kelompok tani, segmen kedua atau cluster 1 terdiri dari 26 kelompok tani, segmen ketiga atau cluster 2 terdiri dari 149 kelompok tani. Sedangkan untuk nilai Davies Bouldin dari cluster 1-3 adalah -0.554 dan tidak mengalami perubahan.

Tabel 8 Hasil Segmentasi Data Kelompok Tani Tahun 2020-2023 Menggunakan K-Means Clustering

| Cluster | Jumlah |
|-----------|--------|
| Cluster 0 | 188 |
| Cluster 1 | 26 |
| Cluster 2 | 149 |

Tabel 9 Nilai Cluster K-Medoids Pada Iterasi Pertama

| Attribut | C1 | C2 | C3 |
|----------|------------|------------|-------------|
| NKT1 | 68586,85 | 6774,67445 | 3592,400112 |
| NKT2 | 61071,8674 | 746,278607 | 11101,94938 |
| NKT3 | 57377,5409 | 4441,11814 | 14796,83656 |
| NKT4 | 28451,0214 | 33372,6225 | 43728,35849 |

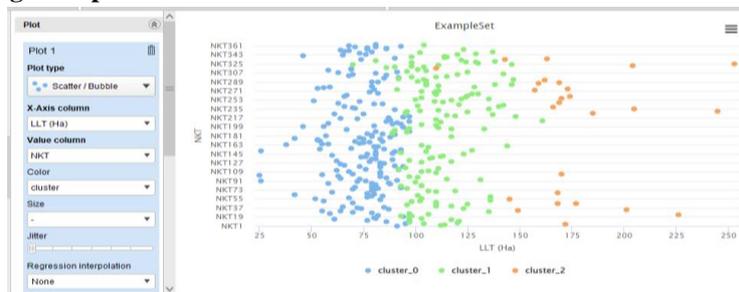
| | | | |
|--------|------------|------------|-------------|
| NKT5 | 66704,9058 | 4896,97625 | 5473,804601 |
| NKT6 | 61919,1397 | 101,01985 | 10254,73913 |
| ... | ... | ... | ... |
| NKT363 | 72173,4226 | 10355,7557 | 0 |

Hasil proses k-medoids clustering menghasilkan tiga model segmentasi seperti yang terdapat pada tabel 11 hasil dari segmentasi tersebut memberikan informasi bahwa segmen pertama atau cluster 0 terdiri dari 8 kelompok tani, segmen kedua atau cluster 1 terdiri dari 185 kelompok tani, segmen ketiga atau cluster 2 terdiri dari 170 kelompok tani. Sedangkan untuk nilai Davies Bouldin dari cluster 1 dan 2 adalah -1.580, dan pada cluster ketiga nilai davies bouldin yang didapat adalah -0.878.

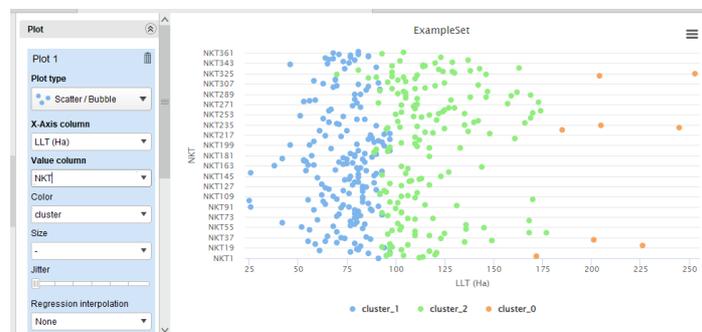
Tabel 10 Hasil Segmentasi Data Kelompok Tani Tahun 2020-2023 Menggunakan K-Medoids Clustering

| Cluster | Jumlah |
|-----------|--------|
| Cluster 0 | 8 |
| Cluster 1 | 185 |
| Cluster 2 | 170 |

Hasil Clustering dengan Rapidminer

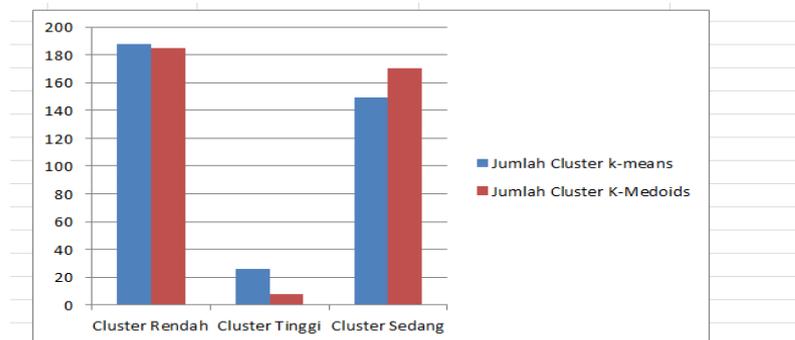


Gambar 4 Plot Hasil Segmentasi Petani Dengan Clustering K-Means



Gambar 5 Plot Hasil Segmentasi Petani Dengan Clustering K-Medoids

Diagram Hasil Clustering Dengan Menggunakan Excel



Gambar 6 Diagram Hasil Clustering K-Means Dan K-Medoids Dengan Menggunakan Excel

Pada gambar 6 menunjukkan diagram dari hasil clustering menggunakan algoritma k-means dan k-medoids dimana cluster yang digunakan yaitu cluster rendah, sedang dan tinggi.. Pada cluster rendah untuk algoritma k-means dan k-medoids terdapat sebanyak 188 dan 185 kelompok tani, untuk cluster tinggi terdapat sebanyak 26 dan 8 kelompok tani, dan untuk cluster sedang terdapat sebanyak 149 dan 170 kelompok tani.

KESIMPULAN

Kesimpulan yang didapat setelah melakukan penelitian yaitu bahwa pada sebelumnya data yang didapat penulis dari pihak Badan Penyuluhan Pertanian (BPP) Kecamatan Moyo Hulu masih bersifat mentah dan tercampur dengan semua kelompok tani. Sehingga penulis melakukan proses pengolahan data petani dengan clusterisasi menggunakan algoritma K-Means dan K-Medoids dengan mengidentifikasi faktor-faktor penting yang mempengaruhi penyaluran pemerataan pupuk yang ada di Kecamatan Moyo Hulu. Pada proses cluster dengan menggunakan algoritma *K-Means* menghasilkan tiga buah cluster yaitu cluster 0 (rendah) terdiri dari 188 kelompok tani, cluster 1 (tinggi) terdiri dari 26 kelompok tani, dan cluster 2 (sedang) terdiri dari 149 kelompok tani. Algoritma *K-Medoids* menghasilkan tiga buah cluster yaitu cluster 0 (tinggi) terdiri dari 8 kelompok tani, cluster 1 (rendah) terdiri dari 185 kelompok tani, dan cluster 2 (sedang) terdiri dari 170 kelompok tani. Berdasarkan dari kedua hasil algoritma tersebut untuk validitas cluster diukur menggunakan indeks davies bouldin. Algoritma K-Means menghasilkan nilai Davies-Bouldin sebesar -0,554 dan algoritma K-Medoids menghasilkan nilai Davies-Bouldin sebesar -1,580 dan -0,878. Berdasarkan dari nilai Davies-Bouldin algoritma K-Means menunjukkan performa yang lebih baik dalam melakukan pengelompokan data petani penerima pupuk di Kecamatan Moyo Hulu dibandingkan dengan algoritma K-Medoids.

REFERENSI

- Purba, D. W., Thohiron, M., Surjaningsih, D. R., Sagala, D., Ramdhini, R. N., Gandasari, D., Wati, C., Purba, T., Herawati, J., Aristia, I., Purba, B., Wisnujati, N. S., & Manullang, S. O. (2020). *Buku 3—Pengantar Ilmu Pertanian*. removed. 201.
- Sumbawa, B. K. (2020). *Kecamatan Moyo Hulu Dalam Angka 2020*.
- Marlina, D., Lina, N., Fernando, A., & Ramadhan, A. (2018). Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokan Wilayah Sebaran Cacat pada Anak. *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer dan Teknologi Informasi*, 4(2), 64. <https://doi.org/10.24014/coreit.v4i2.4498>
- Kholila, N., Mujiono, M., & Wahyudi, D. (2023). Pemetaan Kondisi Lingkungan Tanam menggunakan K-Means Clustering. *JSITIK: Jurnal Sistem Informasi Dan Teknologi Informasi Komputer*, 1(2), 137–147. <https://doi.org/10.53624/jsitik.v1i2.182>
- Gustrianda, R., & Mulyana, D. I. (2022). Penerapan Data Mining Dalam Pemilihan Produk Unggulan dengan Metode Algoritma K-Means Dan K-Medoids. *Jurnal Media Informatika Budidarma*, 6(1), 27. <https://doi.org/10.30865/mib.v6i1.3294>
- Rohmah, A., & Sembiring, F. (2021). Implementasi Algoritma K-Means Clustering Analysis Untuk Menentukan Hambatan Pembelajaran Daring (Studi Kasus: Smk Yaspim *Sistem Informasi Dan ...*, 290–298.
- Khakim, M. A., Rahmadhani, L., Budi Purnomo, E. S., Idayani, R. W., & Rakhmawati, N. A. (2020). Analisa Jaringan Dokumentasi dan Informasi Hukum Kementerian Tentang Teknologi Informasi Menggunakan Metode K-Means Clustering. *Fountain of Informatics Journal*, 5(1), 27. <https://doi.org/10.21111/fij.v5i1.4039>
- Ningsih, S., Suhada, S., Dewi, R., & Windarto, A. P. (2019). Pengklasteran Dokumen Tentang Dispensasi Nikah Menggunakan Algoritma K-Medoids. *Prosiding Seminar Nasional Riset Information Science (SENARIS)*, 1(September), 677. <https://doi.org/10.30645/senaris.v1i0.74>
- Nurmala, T. dkk. (2012). *Pengantar Ilmu Pertanian*. Yogyakarta. GRAHA ILMU.
- Mansyur, N. I., Pudjiwati, E. H., & Murti Laksono, A. (2021). *Book_cf73c831a14995ff5a01536a5944d259*.
- Wahono, S., & Ali, H. (2021). Peranan Data Warehouse, Software Dan Brainware Terhadap Pengambilan Keputusan (Literature Review Executive Support Sistem for Business). *Jurnal Ekonomi Manajemen Sistem Informasi*, 3(2), 225–239. <https://doi.org/10.31933/jemsi.v3i2.781>
- Jollyta, D., Ramdhan, W., & Zarlis, M. (2020). *Konsep Data Mining Dan Penerapan*. Deepublish.
- Sopyan, Y., Lesmana, A. D., & Juliane, C. (2022). Analisis Algoritma K-Means dan Davies Bouldin Index dalam Mencari Cluster Terbaik Kasus Perceraian di Kabupaten Kuningan. *Building of Informatics, Technology and Science (BITS)*, 4(3). <https://doi.org/10.47065/bits.v4i3.2697>
- Yuli Mardi. (2019). Data Mining: Klasifikasi Menggunakan Algoritma C4. 5 Data mining merupakan bagian dari tahapan proses Knowledge Discovery in Database (KDD). *Jurnal Edik Informatika. Jurnal Edik Informatika*, 2(2), 213–219.
- Tarigan, V. (2023). Pembuatan Aplikasi Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma Naive Bayes. *Informatika*, 11(1), 54–62. <https://doi.org/10.36987/informatika.v11i1.3847>