

PENERAPAN METODE *K-MEANS CLUSTERING* UNTUK PENGELOMPOKAN POTENSI PADI DI KOTA PAGAR ALAM

Della Tri Cahaya, Desi Puspita, Riduan Syahri

Teknik Informatika, Institut Teknologi Pagar Alam

Jln. Masik Siagim No.75 Simpang Mbacang, Kota Pagar Alam, Sumatera Selatan, Indonesia

dellatricahaya572@gmail.com

ABSTRAK

Tujuan dari penelitian ini adalah untuk mengklasifikasikan potensi padi menggunakan Algoritma *K-Means* Di Kota Pagar Alam. Penelitian ini dilatar belakangi dengan proses pengelolaan data padi yang dilakukan masih hanya sebatas pengarsipan saja, data disimpan dalam bentuk file microsoft *excel* dan *Microsoft word*. Belum adanya pengelolaan yang lebih lanjut sehingga sulit untuk menentukan daerah mana saja yang memiliki potensi padi diantara 35 kelurahan yang ada di Kota Pagar Alam. Metode pengembangan sistem menggunakan metode *CRISP-DM*. Pengelompokan ini dilakukan dengan menggunakan metode *K-Means Clustering* yang akan menghasilkan 3 *cluster* kelurahan berdasarkan potensi padi. *Cluster 0* merupakan *cluster* dengan potensi rendah, *cluster 1* dengan potensi sedang, dan *cluster 2* dengan potensi tinggi. Pada tahun 2020 didapatkan hasil *cluster-0* total 25 kelurahan, *cluster 1* total 4 kelurahan dan *cluster 2* total 6 kelurahan. Pada tahun 2021 didapatkan hasil *cluster 0* total 23 kelurahan, *cluster 1* total 7 kelurahan dan *cluster 2* total 5 kelurahan. Pada tahun 2022 didapatkan hasil *cluster 0* total 23 kelurahan, *cluster 1* total 6 kelurahan dan *cluster 2* total 6 kelurahan. Saran untuk penelitian selanjutnya agar dapat menjadi pembanding dengan algoritma lain. serta data hasil klasterisasi dapat dimanfaatkan untuk mengklasifikasi tingkat potensi padi di kota pagar alam.

Kata kunci : *Clustering, K-Means, Padi, Rapid Miner.*

1. PENDAHULUAN

Perkembangan Teknologi Informasi dan Komunikasi yang sudah sangat pesat telah digunakan dalam berbagai bidang salah-satunya ialah bidang pertanian. Sektor pertanian merupakan salah satu sektor utama sebagai penghasil pangan[1].

Seiring dengan pertumbuhan penduduk yang semakin meningkat pada setiap tahunnya, membuat kebutuhan pangan juga ikut meningkat[2] Tanaman pangan memiliki banyak jenis, salah satunya adalah tanaman padi. Tanaman padi dianggap sebagai tanaman pangan utama di Indonesia.

Upaya untuk meningkatkan hasil produksi padi secara teknis terus dilakukan oleh Pemerintah khususnya Kementerian Pertanian dengan melakukan pengamanan pertanaman dengan pengendalian OPT dan meminimalkan kehilangan hasil dengan penanganan pasca panen yang baik dan benar[2]

Melihat begitu pentingnya sektor pertanian khususnya padi sebagai pemenuhan kebutuhan pangan penduduk Indonesia khususnya di Kota Pagar Alam, maka perencanaan berbasis wilayah sangat diperlukan guna melihat potensi padi Di Kota Pagar Alam berdasarkan wilayah kelurahan. Perencanaan ini dapat dilakukan dengan memanfaatkan teknologi pada sector pertanian salah satunya ialah berbasis *data mining*.

Data mining adalah hasil *hibridisasi statistik*, ilmu komputer, kecerdasan buatan dan pembelajaran mesin[3] Salah satu metode *data mining* yang dapat digunakan untuk memetakan atau mengelompokkan data yang mirip adalah *clustering*[4].

Clustering dapat diartikan proses untuk mengorganisasikan sekelompok data ke dalam

berbagai kelompok-kelompok sedemikian rupa sehingga objek-objek yang serupa akan menjadi satu *cluster* sedangkan objek-objek yang tidak serupa menjadi anggota *cluster* yang lain[5]. Ada beberapa algoritma yang dapat digunakan pada metode *clustering* seperti *algoritma k-means*.

Algoritma *K-Means clustering* merupakan salah satu metode *clustering non hirarki* yang mempartisi data menjadi beberapa *cluster/kelompok* sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu *cluster* yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain[6] Berdasarkan hal tersebut perlu dilakukan pengelompokan potensi padi untuk dapat mengetahui daerah kelurahan mana yang memiliki potensi padi di Kota Pagar Alam. Pengelompokan ini dilakukan dengan menggunakan metode *K-Means Clustering* yang akan menghasilkan 3 *cluster* kelurahan berdasarkan potensi padi. *Cluster 0* merupakan *cluster* dengan potensi rendah, *cluster 1* dengan potensi sedang, dan *cluster 2* dengan potensi tinggi. Pengelompokan ini dilakukan berdasarkan data yang didapat dari dinas pertanian yang terdiri dari beberapa atribut seperti kelurahan, luas panen(ha), produksi(ton/kg), produktivitas, lahan sementara yang tidak diusahakan, luas pengairan, letak geografis dan jumlah pupuk.

Berdasarkan hal tersebut perlu dilakukan pengelompokan potensi padi untuk dapat mengetahui daerah kelurahan mana yang memiliki potensi padi di Kota Pagar Alam. Penggunaan algoritma *K-Means Clustering* dalam pengelompokan potensi padi dapat menjadi pendukung keputusan pada Dinas Pertanian

Kota Pagar Alam dalam menentukan daerah kelurahan mana yang memiliki potensi padi yang rendah, sedang, dan tinggi. Untuk bisa menentukan prioritas daerah-daerah mana untuk dijadikan sentra padi di Kota Pagar Alam.

2. TINJAUAN PUSTAKA

2.1. Data Mining

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. *Data mining* adalah proses yang menggunakan teknik statistic, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar[7].

Data mining dibagi menjadi beberapa metode diantaranya estimasi, prediksi, klasifikasi, pengklusteran, dan asosiasi.

- a. Estimasi
Evaluasi mirip dengan klasifikasi kecuali bahwa variabel target dievaluasi secara numeric dan bukan secara kategoris. Model dibangun menggunakan kumpulan data lengkap, dimana nilai variabel target adalah nilai prediksi. Selain itu, selama pengujian berikutnya, estimasi nilai variabel target dibuat berdasarkan nilai variabel prediksi.
- b. Prediksi
Prediksi pada dasarnya mirip dengan klasifikasi dan evaluasi, hanya saja memprediksi nilai hasil di masa mendatang. Beberapa metode dan teknik yang digunakan dalam klasifikasi dan evaluasi (dalam kondisi yang sesuai) juga dapat digunakan untuk peramalan.
- c. Klasifikasi
Di klasifikasi ada variabel kategori target. Misalnya, klasifikasi pendapatan dapat dibagi menjadi tiga kategori, yaitu pendapatan tinggi, pendapatan menengah, dan pendapatan rendah.
- d. *Clustering*
Clustering adalah pengelompokan atau fokus dari *record*, observasi, dan pembentukan kategori dari item-item yang serupa. Pengelompokan mengelompokkan data ke dalam kelompok-kelompok tertentu (*cluster*) berdasarkan kesamaan karakteristik dari masing-masing data dalam kelompok yang ada. Ada banyak metode pengelompokan termasuk: *fuzzy c-means*, *k-means*, *SOM* dan *k-medoids*. Algoritma *clustering* yang digunakan dalam penelitian ini adalah algoritma *K-Means*.
- e. Asosiasi
Tugas asosiasi dalam *data mining* adalah menemukan atribut yang muncul bersamaan. Dalam dunia bisnis, hal ini lebih dikenal dengan istilah *shopping cart analysis*.

2.2. Clustering

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[8]

2.3. K-Means

K-Means merupakan metode yang digunakan untuk mengelompokkan data menjadi beberapa bagian, yang mana metode *K-Means* yang mana berbasis jarak yang membagi data menjadi beberapa *cluster* yang hanya dapat bekerja pada atribut *numeric*, *K-Means* dapat digolongkan dalam *partitioning clustering* yang memisahkan data ke k daerah bagian yang berbeda[9]

Langkah-langkah dalam Algoritma *K-Means* adalah :

- a. Dataset disiapkan
- b. Menentukan jumlah *cluster*
- c. Tentukan titik *centroid* secara acak
- d. Menghitung jarak data ke *centroid*
- e. Kelompokkan data hingga terbentuk K dengan titik *centroid* dari setiap *cluster*
- f. Ulangi langkah 3 sampai 5 sampai nilai dari titik *centeroid* tidak lagi berubah.

Adapun persamaan untuk Algoritma *K- Means*

$$d(x,y) = \sum_{i=1}^n (x_i - y_i)^2 \dots \dots \dots (1)$$

$d(x,y)$ adalah jarak antara data x ke data y, x_i adalah data testing ke-i dan y_i adalah data training ke-i.

2.4. Tanaman Pangan

Tanaman pangan adalah segala jenis tanaman yang didalamnya terdapat karbohidrat dan protein sebagai sumber energy manusia. Tanaman pangan juga dapat dikatakan sebagai tanaman utama yang dikonsumsi manusia sebagai makanan untuk memberikan asupan energi bagi tubuh[10].

2.5. Padi

Padi merupakan tanaman pangan penting karena menghasilkan beras yang menjadi sumber bahan makanan pokok. Padi merupakan komoditas utama di Indonesia dalam menyokong pangan masyarakat[11]

2.6. RapidMiner

RapidMiner merupakan perangkat lunak yang bersifat terbuka (*opensource*). *RapidMiner* adalah sebuah solusi untuk untuk melakukan analisis terhadap *data mining*, *text mining* dan analisis prediksi. *RapidMiner* menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik[12]

2.7. CRISP-DM

CRISP-DM adalah standarisasi pengolahan data yang dikembangkan dimana data yang ada melewati setiap langkah terstruktur dan jelas serta efisien. Selain penerapan model dalam proses data mining, pemilihan algoritma sangat mempengaruhi perbandingan metode data mining. Metodologi ini terdiri dari enam tahapan yaitu *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modelling*, *Evaluation*, dan *Deployment*[13].

3. METODE PENELITIAN

3.1. Metode Pengumpulan Data

3.1.1. Observasi

Observasi adalah metode pengumpulan data yang dilakukan dengan pengamatan secara langsung pada objek penelitian, yang dalam hal ini adalah staf kepegawaian Dinas Pertanian Kota Pagar Alam.

3.1.2. Wawancara

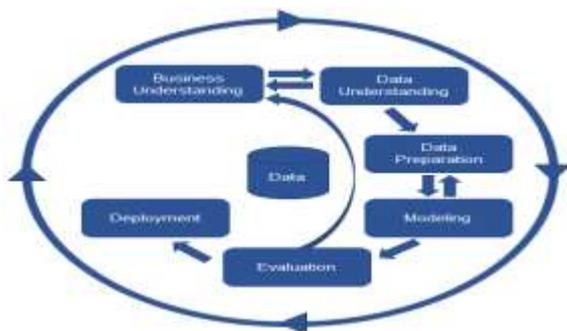
Wawancara adalah metode pengumpulan data yang melakukan tanya jawab secara langsung dari sumber informasi yang tepat dan diteliti yaitu staff pada Dinas Pertanian Kota Pagar Alam.

3.1.3. Studi Pustaka

Studi Pustaka yaitu mempelajari teori-teori baik dari Jurnal, buku cetak maupun non cetak yang berhubungan dengan objek penelitian.

3.2. Metode Pengembangan CRISP-DM

Metode pengembangan Sistem yang digunakan pada penelitian ini adalah metode pengembangan *Cross Industry Standard Process for Data Mining* (CRISP-DM). CRISP-DM sebagai pemecah masalah yang umum untuk bisnis dan penelitian. Metodologi ini terdiri dari enam tahapan yaitu *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modelling*, *Evaluation*, dan *Deployment*. Berikut ini adalah gambar model CRIPS-DM .



Gambar 1. Metode Pengembangan CRISP-DM

Pada gambar 1. Menunjukkan alur proses yang dilakukan pada pengembangan CRISP-DM.

- a. Pemahaman Bisnis (*Business Understanding*)
Tahap pertama subjek pada penelitian ini adalah mengacu pada data padi. Pada tahapan ini diperlukan pemahaman tentang pentingnya pemanfaatan data padi, agar dapat digunakan untuk menentukan daerah kelurahan yang memiliki potensi padi. Oleh karena itu perlu adanya strategi dengan *clustering* atau pengelompokan padi.
- b. Pemahaman data (*Data Understanding*)
Peneliti melakukan pemahaman terhadap kebutuhan data terkait dengan Padi di Kota Pagar Alam. Data diambil dari bagian staff Dinas Pertanian, pemahaman data mengacu pada klasterisasi padi dengan variabel seperti kelurahan, luas panen(ha), produksi(ton/kg), teknik penanaman, luas pengairan, kondisi geografis lahan dan jumlah pupuk. Setelah data didapatkan dilakukan eksplorasi data sejumlah 35 record kelurahan. Data yang dikumpulkan yaitu data padi selama 3 Tahun pada tahun 2020 sampai dengan 2022
- c. Persiapan Data (*Data Preparation*)
Setelah tahap pemahaman data partisipatif pegawai selesai, tahap pengolahan data ini menghasilkan kumpulan data akhir yang diterapkan pada pemodelan data. Dalam tahapan ini yaitu membangun *dataset* akhir dari berupa data mentah. Ada beberapa hal yang akan dilakukan mencakup melakukan pembersihan data (*Data Cleaning*), *dataset* yang sudah dilakukan proses *cleaning* kemudian akan dilakukan pemilihan data (*Data Selection*), *record* dan atribut-atribut.
- d. Pemodelan (*Modelling*)
Pada tahapan ini secara langsung melibatkan *Machine Learning* untuk penentuan teknik *data mining*, alat bantu serta algoritma yang ada dalam *data mining*. Adapun pemodelan yang dilakukan pada penelitian ini adalah metode *clustering* menggunakan algoritma *k-means*. Berikut langkah-langkah algoritma *K-Means* yang digunakan dalam pemodelan ini :

- a. Siapkan *Dataset*
Dataset berasal dari data-data mentah padi dari tahun 2020 sampai tahun 2022 yang telah di *cleaning* yang terdiri dari beberapa atribut seperti kelurahan, luas panen(ha), produksi(ton/kg), produktivitas, lahan sementara yang tidak diusahakan, luas pengairan, letak geografis dan jumlah pupuk.
- b. Tentukan Jumlah *Cluster*
Berdasarkan *dataset* yang disiapkan, digunakan *cluster* yang dibentuk berdasarkan 3 klaster berdasarkan potensi lahan . Klaster-klaster tersebut yaitu klaster dengan potensi rendah, potensi sedang dan potensi tinggi.
C0 : Rendah
C1 : Sedang
C2 : Tinggi

- c. Tentukan *centroid* secara acak
 Karena ada 3 *cluster* yang diharapkan, diperlukan 3 *centroid* sebagai pusat *cluster*, yang ditentukan secara acak.
- d. Hitung jarak data ke *centroid*
 Untuk menghitung jarak ke *centroid* digunakan rumus *Euclidean Distance*:

$$d(x, y) = \sum_{k=1}^n (x_i - y_i)^2$$

Dendan keterangan $d(x,y)$ adalah jarak antara data x ke data y , x_i adalah data testing ke- i dan y_i adalah data training ke- i .

- Pada perhitungan ini adalah melakukan menentukan nilai initial *cluster centre* untuk masing-masing *cluster* pada setiap variabelnya. Nilai initial *cluster centre* pada iterasi yang pertama (perhitungan pertama kali) diberikan secara acak. Pada literasi selanjutnya, nilai initial *cluster centre* (pengulangan ke-1 sampai dengan posisi normal/maksimal iterasi) diberikan dengan menghitung nilai rata-rata data pada setiap *cluster*nya. Jika nilai initial *cluster centre* yang baru sama dengan nilai initial *cluster centre* yang baru maka proses literasi dilanjutkan hingga nilai sama atau sampai dengan nilai maksimal iterasi yang telah ditetapkan sebelumnya. Namun jika nilai initial *cluster centre* yang baru sama dengan initial *cluster centre* yang lama, proses pengelompokan berhenti.
- e. Perbarui titik *Centroid*
 Perhitungan titik *centroid* baru dengan berdasarkan pengelompokan yang pertama untuk melakukan iterasi selanjutnya.
- f. Ulangi langkah 3 sampai 5 sampai nilai dari titik *centroid* tidak lagi berubah.
 Perhitungan dilakukan seterusnya sampai dengan jumlah data yang telah ditentukan. Setelah menghitung jarak data dari pusat, langkah selanjutnya adalah mengelompokkan data berdasarkan jarak terdekat. Kemudian Anda dapat melakukan lebih banyak pengulangan untuk membandingkan hasilnya. Jika *centroid* berubah, ulangi langkah tersebut sampai *centroid* tidak berubah.
- g. Evaluasi (*Evaluation*)
 Pada tahap ini lebih difokuskan pada model yang dihasilkan sudah sesuai standar *K-Means Clustering* dan tidak ada yang dilewatkan saat melakukan tahap awal hingga tahap pemodelan selesai dengan pengujian menggunakan *Elbow Method*.
- h. Penyebaran (*Deployment*)
 Pada tahap terakhir ini *CRIPS-DM* adalah penyebaran pengetahuan atau informasi yang telah diperoleh selanjutnya akan diimplementasikan pada laporan dan melakukan analisis tiap *cluster* yang diperoleh

supaya dapat mudah dimengerti oleh pihak Dinas Pertanian Kota Pagar Alam.

4. HASIL DAN PEMBAHASAN

Hasil dari penelitian ini menggunakan teknik CRISP-DM:

4.1. Business Understanding

Pada pemahaman bisnis dilakukan langkah menentukan tujuan pencarian dan lingkungan pencarian dilakukan saat maksud pencarian ini adalah untuk mengelompokkan potensi padi, proses pengelolaan data padi untuk menentukan kelurahan mana yang memiliki potensi masih dilakukan secara manual, yaitu melihat dan menghitung di *Excel*. Dari proses pengelompokan tersebut belum adanya pengelolaan lebih lanjut sehingga menyebabkan Dinas Pertanian Kota Pagar Alam sulit untuk menentukan daerah mana saja yang memiliki potensi padi diantara 35 kelurahan yang ada di Kota Pagar Alam. Maka dari itu peneliti melakukan pengelompokan potensi padi menggunakan metode *k-means* untuk dapat mengetahui daerah kelurahan mana yang memiliki potensi padi di Kota Pagar Alam. Sehingga dapat berguna bagi Dinas Pertanian Kota Pagar Alam untuk bisa menentukan prioritas daerah-daerah mana untuk dijadikan sentra padi di Kota Pagar Alam.

4.2. Data Understanding

Pada fase pemahaman data ini, data didapat dari bagian tanaman pangan Dinas Pertanian Kota Pagar Alam yaitu data potensi padi selama 3 tahun pada tahun 2020 sampai dengan tahun 2022 dengan jumlah 105 *record* dengan 8 atribut yaitu kelurahan, luas panen, produksi, lahan sementara yang tidak diusahakan, luas lahan pengairan, produktivitas, letak geografis, jumlah pupuk data yang diterima dalam bentuk *excel* dan data yang didapat perlu di *cleaning* dan akan dilakukan pemilihan data dengan atributnya.

4.3. Data Preparation

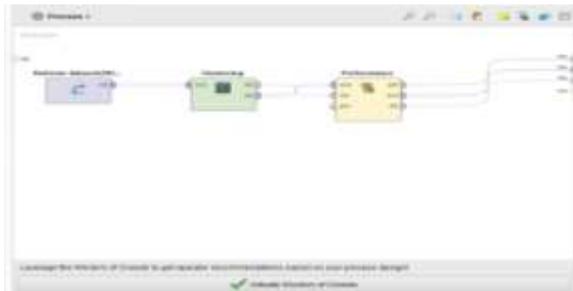
Pengolahan data langsung diimplementasikan pada *Rapid Miner* dengan dua tahapan yang ditetapkan pada pengolahan data yaitu :

4.3.1. Data Selection

Pemilihan (seleksi) berarti pengumpulan informasi tentang data pekerjaan harus dilakukan sebelum dimulainya penambangan data KDD. Dilakukan pemilihan atribut dan sebagian data dalam atribut akan dikonversikan agar proses data *mining* menjadi lebih mudah. Jika data yang diperoleh masih terdapat banyak *inkonsistensi data*, maka perlu adanya penyeleksian data dengan cara filter data. Data yang difilter adalah data yang memiliki atribut yang lengkap sesuai dengan yang telah ditentukan. Peneliti hanya memakai data yang lengkap dengan atribut yang dipilih data sudah bersih.

Datasheet padi pada tahun 2020 sebanyak 35 *recod* yang kemudian di proses menjadi *datasheet* dan selanjutnya dilakukan proses *modeling* menggunakan algoritma *k-means clustering* dan *Performance vector* proses dapat dilihat pada gambar 8

Datasheet padi pada tahun 2021 sebanyak 35 *recod* yang kemudian di proses menjadi *datasheet* dan selanjutnya dilakukan proses *modeling* menggunakan algoritma *k-means clustering* dan *Performance vector* proses dapat dilihat pada gambar 9.



Gambar 9. Model Algoritma K-Means 2021

Datasheet padi pada tahun 2022 sebanyak 35 *recod* yang kemudian di proses menjadi *datasheet* dan selanjutnya dilakukan proses *modeling* menggunakan algoritma *k-means clustering* dan *Performance vector* proses dapat dilihat pada gambar 10.

Tipe	Uraian	Bidang	Status	Nilai (1)	Nilai (2)	Nilai (3)
kelurahan	Pasir	1	1	10000	10000	10000
kelurahan	Pasir	1	1	10000	10000	10000
kelurahan	Pasir	1	1	10000	10000	10000
kelurahan	Pasir	1	1	10000	10000	10000
kelurahan	Pasir	1	1	10000	10000	10000
kelurahan	Pasir	1	1	10000	10000	10000
kelurahan	Pasir	1	1	10000	10000	10000
kelurahan	Pasir	1	1	10000	10000	10000
kelurahan	Pasir	1	1	10000	10000	10000
kelurahan	Pasir	1	1	10000	10000	10000

Gambar 10. Model Algoritma K-Means 2022

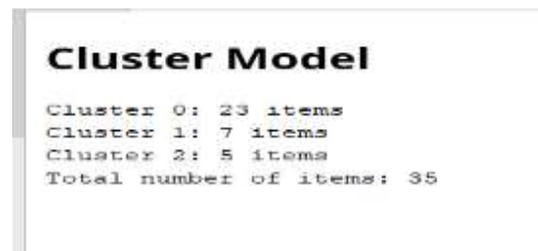
4.3.4. Evaluation

Pada tahap *evaluation* model sudah membentuk cluster yang terdiri dari cluster 0 merupakan cluster dengan potensi rendah, cluster 1 merupakan cluster dengan potensi sedang dan cluster 2 dengan potensi tinggi. Pada tahun 2020 terdapat 25 kelurahan yang termasuk ke *cluster 0*, 4 kelurahan pada *cluster 1* dan 6 kelurahan pada *cluster 2* dapat dilihat pada gambar 11.



Gambar 11. Cluster Model 2020

Pada tahun 2021 terdapat 23 kelurahan yang termasuk ke *cluster 0*, 7 kelurahan pada *cluster 1* dan 5 kelurahan pada *cluster 2* dapat dilihat pada gambar 12.



Gambar 12. Cluster Model 2021

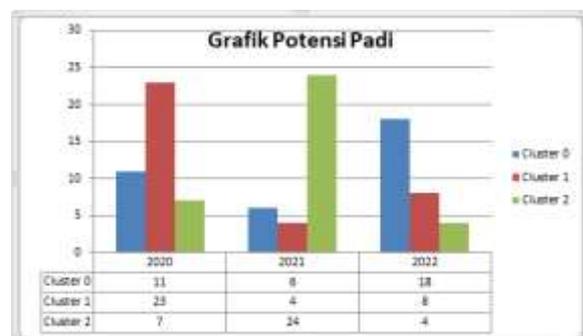
Pada tahun 2022 terdapat 23 kelurahan yang termasuk ke *cluster 0*, 6 kelurahan pada *cluster 1* dan 6 kelurahan pada *cluster 2* dapat dilihat pada gambar 13.



Gambar 13. Cluster Model 2022

4.3.5. Deployment

Pada tahap ini *deployment* merupakan tahapan terakhir berupa pengetahuan atau informasi yaitu mengenai pola potensi padi di Dinas Pertanian Kota Pagar Alam selama 3 tahun.



Gambar 14. Grafik Potensi Padi 2020-2022

Dari gambar 14. dapat diketahui ada 3 *cluster*, yang pertama *cluster 0* adalah rendah dengan total 25

kelurahan pada tahun 2020, 23 kelurahan di tahun 2021 dan 23 kelurahan pada tahun 2022, kemudian *cluster 1* memiliki keterangan sedang dengan total 6 kelurahan pada tahun 2020, 7 kelurahan ditahun 2021 dan 6 kelurahan ditahun 2022, dan *cluster 2* dengan keterangan tinggi dengan total 6 kelurahan di tahun 2020, 5 kelurahan pada tahun 2021 dan 6 kelurahan ditahun 2022 .

5. KESIMPULAN DAN SARAN

Berdasarkan penelitian yang telah dilakukan dengan menggunakan algoritma K-Means, data potensi padi dari tahun 2020-2022 di kota pagar alam dapat dikelompokkan menjadi 3 *cluster*. *Cluster 0* dengan potensi rendah, *cluster 1* dengan potensi sedang dan *cluster 2* dengan potensi tinggi. Pada tahun 2020 didapatkan hasil *cluster-0* total 25 kelurahan, *cluster 1* total 4 kelurahan dan *cluster 2* total 6 kelurahan. Pada tahun 2021 didapatkan hasil *cluster 0* total 23 kelurahan, *cluster 1* total 7 kelurahan dan *cluster 2* total 5 kelurahan. Pada tahun 2022 didapatkan hasil *cluster-0* total 23 kelurahan, *cluster 1* total 6 kelurahan dan *cluster 2* total 6 kelurahan. Saran untuk penelitian selanjutnya agar dapat menjadi pembanding dengan algoritma lain.serta data hasil klasterisasi dapat dimanfaatkan untuk mengklasifikasi tingkat potensi padi di kota pagar alam.

DAFTAR PUSTAKA

- [1] H. Husdi and Y. Lasena, "Real Time Analisis Berbasis Internet Of Things Untuk Prediksi Iklim Lahan Pertanian," *J. Media Inform. Budidarma*, vol. 4, no. 3, pp. 834–840, 2020, [Online]. Available: <http://ejurnal.stmik-budidarma.ac.id/index.php/mib/article/view/2165>
- [2] E. Suprpto, "Pengelompokan Potensi Padi di Indonesia Menggunakan K-Means Cluster," *J. Ilm. Pop.*, vol. 5, 2021.
- [3] C. Armayani, A. Fauzi, and H. Sembiring, "Implementasi Data Mining Pengelompokan Jumlah Data Produktivitas Ubinan Tanaman Pangan Berdasarkan Jenis Ubinan Dengan Metode Clustering Dikab Langkat," *J. Inform. Kaputama*, vol. 5, no. 1, pp. 185–196, 2021.
- [4] S. Mulyono and W. Utami, "Pemetaan Potensi Lahan Pertanian Pangan Berkelanjutan Guna Mendukung Ketahanan Pangan," *Bhumi J. Agrar. dan Pertanah.*, vol. 6 Nomor 2, pp. 201–218, 2020.
- [5] T. Tendean and W. Purba, "Analisis Cluster Provinsi Indonesia Berdasarkan Produksi Bahan Pangan Menggunakan Algoritma K-Means," *J. Sains dan Teknol.*, vol. 1, no. 2, pp. 5–11, 2020.
- [6] D. W. Sitohang and A. Rikki, "Implementasi Algoritma K- Means Clustering untuk Mengelompokkan Data Gizi Balita pada Kecamatan Garoga Tapanuli Utara," *KAKIFIKOM (Kumpulan Artik. Karya Ilm. Fak. Ilmu Komputer)*, vol. 02, pp. 80–92, 2019, doi: 10.54367/kakifikom.v1i2.642.
- [7] L. Maulida, "Penerapan Datamining Dalam Mengelompokkan Kunjungan Wisatawan Ke Objek Wisata Unggulan Di Prov. Dki Jakarta Dengan K-Means," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 2, no. 3, p. 167, 2018, doi: 10.14421/jiska.2018.23-06.
- [8] W. Ginting, "Pengelompokan Data Pasien Test Urine Dengan Metode Clustering Pada Kantor Badan Narkotika Nasional," *J. Tek. Inform. Kaputama*, vol. 5, no. 2, pp. 327–338, 2021.
- [9] A. Febriyanto, S. Achmadi, and A. P. Sasmito, "Penerapan Metode K-Means Untuk Clustering Pengunjung Perpustakaan Itn Malang," *J. Mhs. Tek. Inform.*, vol. 5, no. 1, pp. 61–70, 2021.
- [10] I. Nuriati, B. S. Ginting, and Y. Maulita, "Sistem Pendukung Keputusan Pemilihan Jenis Tanaman Pangan Berdasarkan Kondisi Tanah dengan Metode Moora," *Semin. Nas. Inform.*, pp. 285–294, 2021.
- [11] A. Rambu, E. Umbu, and K. Retang, "Faktor-Faktor Yang Mempengaruhi Produksi Padi Sawah Tadah Hujan Di Desa Umbu Pabal Kecamatan Umbu Ratu Nggay Barat Kabupaten Sumba Tengah Factors Affecting Rice Production of Rainfed Rice in the Village of Umbu Pabal, Umbu Ratu Nggay Barat District, Centr," *J. Pertan. Agros*, vol. 25, no. 1, pp. 714–723, 2023.
- [12] N. Iriadi, L. Setioningias, and P. Priatno, "Implementasi Data Mining Pada Klasifikasi Ketidakhadiran Pegawai Menggunakan Metode C4.5," *Comput. Sci.*, vol. 1, no. 1, pp. 53–61, 2021, doi: 10.31294/coscience.v1i1.198.
- [13] M. A. Hasanah, S. Soim, and A. S. Handayani, "Implementasi CRISP-DM Model Menggunakan Metode Decision Tree dengan Algoritma CART untuk Prediksi Curah Hujan Berpotensi Banjir," *J. Appl. Informatics Comput.*, vol. 5, no. 2, pp. 103–108, 2021, doi: 10.30871/jaic.v5i2.3200.