PROSISI: Prosiding Seminar Nasional Saintek
Vol.1 No.1, 2026, Hal. 369- 380
ISSN XXXX-XXXX

Analisis *Clustering* Wilayah Kerentanan Produksi Padi Berdasarkan Iklim Dan Luas Panen Di Provinsi Sumatera Selatan Menggunakan Algoritma *K-Means*

Reza Koswara[1*], M Dimas Ulinuha[1], Febri Hadi Noto[1], Haris Al Mufaridun[1], Ferry Muhammad Ramadhan[1] , Annisa Nur Azizah[2]

[1]Prodi Informatika, Universitas Nurul Huda
[2]Prodi Matematika, Universitas Nurul Huda
[1*]E-mail: koswarareza07@gmail.com

Abstrak

Kerentanan produksi padi di berbagai wilayah menjadi tantangan utama dalam mewujudkan ketahanan pangan daerah. Penelitian ini bertujuan untuk mengelompokkan kabupaten atau kota di Provinsi Sumatera Selatan berdasarkan tingkat kerentanan produksi padi menggunakan algoritma *K-Means*. Data sekunder dari 17 kabupaten atau kota periode 2018–2025 diperoleh dari BPS dan NASA POWER API, mencakup variabel luas panen, curah hujan, kelembapan udara, suhu rata-rata, dan produksi padi. Tahapan penelitian meliputi preprocessing, normalisasi *Min-Max*, *clustering K-Means*, serta evaluasi menggunakan *Elbow Method* dan *Silhouette Score*. Hasil clustering menghasilkan tiga kelompok wilayah yaitu kerentanan rendah (Kabupaten Banyuasin, OKI, dan OKU Timur), kerentanan sedang (sebagian besar kabupaten atau kota), dan kerentanan tinggi (wilayah dengan luas panen dan produksi rendah). Jumlah cluster optimal adalah tiga dengan *Silhouette Score* 0,31 yang mengindikasikan pemisahan antar *cluster* yang dapat diterima. Penelitian ini berkontribusi dalam pemanfaatan data iklim berbasis API sebagai pendekatan analisis kerentanan produksi padi untuk mendukung perencanaan pertanian dan kebijakan ketahanan pangan di Provinsi Sumatera Selatan.

Kata Kunci : *K-Means, Clustering,* Kerentanan Produksi Padi, Ketahanan Pangan, NASA POWER, Sumatera Selatan.

Abstract

The vulnerability of rice production across various regions poses a major challenge in achieving regional food security. This study aims to cluster districts and cities in South Sumatra Province based on the level of rice production vulnerability using the K-Means algorithm. Secondary data from 17 districts and cities for the period 2018–2025 were obtained from BPS and NASA POWER API, covering variables of harvested area, rainfall, air humidity, average temperature, and rice production. The research stages included preprocessing, Min-Max normalization, K-Means clustering, and evaluation using the Elbow Method and Silhouette Score. The clustering results produced three regional groups: low vulnerability (Banyuasin Regency, OKI, and OKU Timur), moderate vulnerability (most districts and cities), and high vulnerability (regions with relatively low harvested area and production). The optimal number of clusters is three with a Silhouette Score of 0.31, indicating an acceptable separation between clusters. This study contributes to the utilization of API-based climate data as an analytical approach to rice production vulnerability in support of agricultural planning and food security policy in South Sumatra Province.

Keywords: *K-Means, Clustering, Rice Production Vulnerability, Food Security, NASA POWER, South Sumatra.*

PENDAHULUAN

Produksi padi merupakan salah satu komponen utama dalam menjaga ketahanan pangan nasional karena beras masih menjadi makanan pokok yang dikonsumsi oleh sebagian besar masyarakat Indonesia. Tingginya tingkat konsumsi beras menyebabkan ketersediaan produksi padi harus terus dijaga agar kebutuhan pangan nasional dapat terpenuhi. Selain berperan sebagai penyedia pangan, sektor pertanian juga menjadi salah satu sektor strategis yang mendukung pertumbuhan ekonomi nasional. Oleh karena itu, peningkatan dan stabilitas produksi padi perlu mendapat perhatian khusus, terutama pada daerah yang memiliki potensi pertanian yang besar seperti Provinsi Sumatera Selatan.

Provinsi Sumatera Selatan merupakan salah satu daerah penghasil padi yang memiliki kontribusi penting terhadap ketahanan pangan nasional. Namun demikian, tingkat produksi padi pada setiap kabupaten/kota tidak

selalu menunjukkan kondisi yang sama. Perbedaan luas panen, kondisi geografis, serta karakteristik iklim menyebabkan kemampuan produksi setiap wilayah menjadi berbeda. Akibatnya, terdapat daerah yang mampu menghasilkan produksi padi tinggi secara konsisten, sementara daerah lainnya mengalami fluktuasi produksi dari tahun ke tahun. Kondisi tersebut menunjukkan bahwa setiap wilayah memiliki karakteristik produksi yang berbeda sehingga diperlukan analisis yang mampu menggambarkan pola dan karakteristik wilayah secara lebih terstruktur.

Produktivitas padi dipengaruhi oleh berbagai faktor yang saling berkaitan, baik dari aspek produksi maupun kondisi lingkungan. Luas panen merupakan faktor yang berhubungan langsung dengan jumlah produksi yang dihasilkan, sedangkan faktor iklim seperti curah hujan, kelembapan udara, dan suhu rata-rata berperan dalam mendukung pertumbuhan tanaman padi. Perbedaan kondisi tersebut dapat menyebabkan variasi hasil produksi antarwilayah sehingga menghasilkan karakteristik produksi yang berbeda pada setiap daerah. [1], menjelaskan bahwa faktor produksi dan kondisi lingkungan memiliki keterkaitan terhadap perbedaan produktivitas padi pada berbagai wilayah.

Selain faktor produksi, perubahan iklim juga menjadi tantangan yang perlu diperhatikan dalam sektor pertanian. Perubahan pola curah hujan, peningkatan suhu udara, serta meningkatnya frekuensi kejadian banjir maupun kekeringan dapat memengaruhi pertumbuhan tanaman dan berpotensi menurunkan hasil panen. Apabila kondisi tersebut terjadi secara berkelanjutan, maka stabilitas produksi pangan dapat terganggu dan berdampak pada ketahanan pangan daerah maupun nasional. [2], perubahan kondisi iklim memiliki hubungan dengan produktivitas tanaman padi sehingga faktor iklim perlu dipertimbangkan dalam perencanaan dan pengelolaan sektor pertanian.

Pada Provinsi Sumatera Selatan, kondisi iklim antarwilayah menunjukkan karakteristik yang beragam. Curah hujan, suhu udara, dan kelembapan relatif pada setiap kabupaten/kota tidak selalu berada pada kondisi yang sama. Perbedaan tersebut dapat memengaruhi kondisi pertumbuhan tanaman padi dan berkontribusi terhadap variasi hasil produksi yang dihasilkan. Oleh karena itu, analisis yang mempertimbangkan faktor produksi dan faktor iklim secara bersamaan diperlukan untuk memperoleh gambaran karakteristik wilayah produksi padi yang lebih komprehensif.

Perkembangan teknologi informasi memungkinkan pemanfaatan teknik data mining untuk mengolah data dalam jumlah besar dan menemukan pola yang tersembunyi di dalam data. Salah satu teknik yang banyak digunakan adalah *clustering*, yaitu proses pengelompokan data berdasarkan tingkat kemiripan karakteristik tertentu. [3], teknik *clustering* dapat membantu mengidentifikasi kelompok data yang memiliki karakteristik serupa sehingga mempermudah proses analisis dan interpretasi data. Melalui proses tersebut, pola hubungan antar data dapat diketahui secara lebih sistematis dan terstruktur.

Salah satu metode *clustering* yang banyak digunakan adalah algoritma *K-Means*. Metode ini bekerja dengan mengelompokkan data berdasarkan tingkat kedekatan karakteristik menggunakan perhitungan jarak antar objek data. Kemampuan algoritma *K-Means* dalam mengolah data numerik tanpa memerlukan label kelas menjadikannya sesuai untuk menganalisis data produksi padi yang terdiri atas variabel kuantitatif seperti luas panen, curah hujan, kelembapan udara, suhu rata-rata, dan produksi padi. [4], menjelaskan bahwa algoritma *K-Means* mampu menghasilkan kelompok data yang memiliki tingkat homogenitas tinggi dalam satu *cluster* dan heterogenitas yang tinggi antarcluster sehingga sesuai digunakan untuk mengidentifikasi kemiripan karakteristik wilayah.

Penerapan algoritma *K-Means* pada sektor pertanian telah banyak dilakukan dalam berbagai penelitian.[1], menerapkan algoritma *K-Means* untuk mengelompokkan produktivitas padi di Pulau Sumatera dan menunjukkan bahwa metode tersebut mampu mengidentifikasi pola karakteristik produksi padi berdasarkan kesamaan data antarwilayah. [4], memanfaatkan algoritma *K-Means* untuk memetakan potensi tanaman padi berdasarkan karakteristik produksi wilayah dan menghasilkan kelompok wilayah yang memiliki karakteristik serupa. [5], juga menerapkan metode *K-Means* pada data pertanian dan menunjukkan bahwa proses *clustering* dapat membantu pengambilan keputusan melalui pengelompokan data berdasarkan tingkat kemiripan karakteristik yang dimiliki.

Hasil penelitian-penelitian tersebut menunjukkan bahwa algoritma *K-Means* efektif digunakan untuk mengidentifikasi pola dan karakteristik data pertanian. Namun demikian, sebagian besar penelitian terdahulu masih berfokus pada variabel produksi atau produktivitas pertanian. Padahal, karakteristik produksi padi pada suatu wilayah tidak hanya dipengaruhi oleh luas panen dan hasil produksi, tetapi juga berkaitan dengan kondisi iklim yang berbeda pada setiap daerah. Pengelompokan yang hanya menggunakan variabel produksi berpotensi belum mampu memberikan gambaran karakteristik wilayah secara menyeluruh. Selain itu, penelitian yang mengombinasikan variabel produksi padi, luas panen, curah hujan, kelembapan udara, dan suhu rata-rata untuk mengelompokkan karakteristik wilayah produksi padi di Provinsi Sumatera Selatan masih relatif terbatas. Oleh karena itu, diperlukan penelitian yang mampu mengelompokkan wilayah berdasarkan kombinasi faktor produksi dan faktor iklim guna memperoleh gambaran pola kemiripan antarwilayah yang lebih komprehensif.

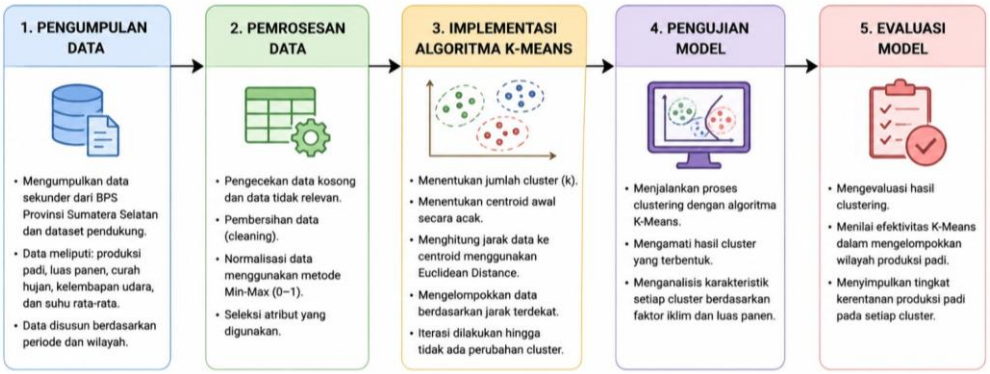
Berdasarkan permasalahan tersebut, penelitian ini menggunakan data produksi padi dan luas panen pada 17 kabupaten/kota di Provinsi Sumatera Selatan selama periode 2018–2025 yang diperoleh dari Badan Pusat Statistik (BPS). Data iklim yang digunakan meliputi suhu udara rata-rata (T2M), kelembapan relatif (RH2M), dan curah hujan (*PRECTOTCORR*) yang diperoleh dari *NASA POWER* melalui layanan *Application*

Programming Interface (API) berdasarkan koordinat geografis masing-masing kabupaten/kota. Seluruh data yang diperoleh kemudian melalui tahapan preprocessing dan normalisasi menggunakan metode *Min-Max Normalization* sebelum dilakukan proses clustering menggunakan algoritma *K-Means*.

Penelitian ini bertujuan untuk mengelompokkan kabupaten/kota di Provinsi Sumatera Selatan berdasarkan karakteristik produksi padi, luas panen, curah hujan, kelembapan udara, dan suhu rata-rata menggunakan algoritma *K-Means*. Hasil pengelompokan diharapkan mampu memberikan gambaran mengenai pola kemiripan karakteristik wilayah produksi padi sehingga dapat menjadi bahan pertimbangan dalam perencanaan dan pengelolaan sektor pertanian yang lebih efektif serta mendukung upaya peningkatan ketahanan pangan di Provinsi Sumatera Selatan.

METODE PENELITIAN

Pada Penelitian ini menggunakan pendekatan *data mining* dengan teknik *clustering* menggunakan algoritma *K-Means* untuk mengelompokkan wilayah berdasarkan karakteristik produksi padi, luas panen, dan faktor iklim di Provinsi Sumatera Selatan. Algoritma *K-Means* dipilih karena mampu mengelompokkan data berdasarkan tingkat kemiripan karakteristik serta memiliki proses perhitungan yang relatif sederhana. Penerapan metode ini pada data pertanian telah menunjukkan kemampuan yang baik dalam mengidentifikasi pola dan karakteristik data yang serupa [6]. Oleh karena itu, algoritma *K-Means* dinilai sesuai untuk digunakan dalam proses pengelompokan wilayah berdasarkan variabel produksi padi, luas panen, curah hujan, kelembapan udara, dan suhu rata-rata.



Gambar 1 Tahapan Metode Penelitian

Tahapan penelitian diawali dengan proses pengumpulan data. Data yang digunakan merupakan data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) Provinsi Sumatera Selatan dan Data iklim berupa suhu udara rata-rata (T2M), kelembapan relatif (RH2M), dan curah hujan (*PRECTOTCORR*) diperoleh dari *NASA POWER* (*Prediction Of Worldwide Energy Resources*) melalui layanan *API NASA POWER*. Data diambil berdasarkan koordinat geografis masing-masing kabupaten/kota di Provinsi Sumatera Selatan untuk periode 2018–2025.

Tahap berikutnya adalah *preprocessing* atau pembersihan data (*data cleaning*). Pada tahap ini dilakukan pengecekan terhadap data kosong, penghapusan data yang tidak relevan, serta pemilihan atribut yang digunakan dalam penelitian. Variabel yang digunakan terdiri dari produksi padi, luas panen, curah hujan, kelembapan udara, dan suhu rata-rata.

Selanjutnya dilakukan proses analisis menggunakan algoritma *K-Means*. Tahapan analisis meliputi penentuan jumlah *cluster,* penentuan *centroid* awal, perhitungan jarak data menggunakan metode *Euclidean Distance*, serta proses pengelompokan data berdasarkan *centroid* terdekat. Metode *K-Means Clustering* mampu membantu proses analisis data pertanian melalui pengelompokan data berdasarkan karakteristik tertentu sehingga hasil analisis menjadi lebih efektif dan terstruktur [5]. Proses iterasi dilakukan secara berulang hingga tidak terjadi perubahan anggota pada masing-masing *cluster*. Hasil pengelompokan kemudian dianalisis untuk mengetahui tingkat kerentanan produksi padi pada setiap wilayah di Provinsi Sumatera Selatan. *Cluster* yang terbentuk dibagi menjadi beberapa kategori, yaitu tingkat kerentanan rendah, sedang, dan tinggi berdasarkan karakteristik produksi padi serta faktor iklim yang digunakan dalam penelitian.

Penelitian ini menghasilkan informasi mengenai pola pengelompokan wilayah produksi padi berdasarkan faktor iklim dan luas panen. Hasil penelitian diharapkan dapat menjadi bahan pertimbangan bagi pemerintah maupun pihak terkait dalam menentukan kebijakan pertanian guna meningkatkan produktivitas padi serta memperkuat ketahanan pangan di Provinsi Sumatera Selatan.

Pengumpulan Data

Tahap pengumpulan data dilakukan dengan mengumpulkan data yang berkaitan dengan produksi padi dan faktor-faktor yang memengaruhinya dari sumber yang relevan. Data produksi padi dan luas panen diperoleh dari Badan Pusat Statistik (BPS) Provinsi Sumatera Selatan, sedangkan data pendukung lainnya digunakan untuk melengkapi kebutuhan penelitian [7]. Data yang digunakan meliputi produksi padi, luas panen, curah hujan, kelembapan udara, dan suhu rata-rata pada wilayah kabupaten/kota di Provinsi Sumatera Selatan selama periode 2018–2025. Seluruh data yang diperoleh kemudian disusun dalam format *spreadsheet* sebagai bahan untuk proses pengolahan dan analisis lebih lanjut.

Pemanfaatan data produksi padi yang dipadukan dengan faktor-faktor pendukung pertanian memungkinkan proses analisis dilakukan secara lebih komprehensif. Kombinasi variabel tersebut dapat membantu mengidentifikasi karakteristik wilayah berdasarkan kondisi produksi dan faktor lingkungan yang memengaruhinya [6]. Oleh karena itu, data yang digunakan dalam penelitian ini tidak hanya mencakup aspek produksi, tetapi juga mempertimbangkan faktor iklim yang berpotensi memengaruhi hasil produksi padi.

Dataset penelitian terdiri atas lima variabel utama, yaitu luas panen, curah hujan, kelembapan udara, suhu rata-rata, dan produksi padi. Variabel-variabel tersebut dipilih karena memiliki keterkaitan dengan kondisi produksi padi pada masing-masing wilayah. Selain itu, faktor iklim dan kondisi lingkungan diketahui memiliki pengaruh terhadap produktivitas sektor pertanian sehingga dapat dimanfaatkan dalam proses *clustering* untuk mengidentifikasi pola kemiripan antar wilayah [5]. Melalui penggunaan variabel produksi dan faktor iklim secara bersamaan, hasil pengelompokan diharapkan mampu memberikan gambaran yang lebih representatif mengenai karakteristik wilayah produksi padi di Provinsi Sumatera Selatan.

Tabel Data Produksi Padi dan Faktor Iklim

Kabupaten/Kota	Tahun	Produksi (Ton)	Luas Panen (Ha)	Curah Hujan (mm)	Kelembapan (%)	Suhu Rata-Rata (°C)
Ogan Komering Ulu	2018	14,124.16	3,039.42	2,562.33	87.11	24.36
Ogan Komering Ilir	2018	484,123.06	95,573.80	2,208.73	88.03	26.17
Muara Enim	2018	84,206.44	18,082.82	2,511.98	86.76	25.86
Lahat	2018	75,360.65	13,966.04	2,562.33	87.11	24.36
…	…	…	…	…	...	…
Musi Rawas Utara	2025	10,873.00	2,435.00	2,905.20	87.00	26.46
Kota Palembang	2025	21,729.00	4,100.00	3,051.93	89.06	26.50
Kota Prabumulih	2025	862.00	193.00	3,076.90	87.63	26.49
Kota Pagar Alam	2025	19,633.00	3,636.00	4,032.19	88.73	23.52

Data produksi padi yang digunakan pada penelitian ini dipertahankan dalam bentuk nilai asli tanpa dilakukan pembulatan angka. Hal tersebut dilakukan agar informasi pada setiap variabel tetap terjaga dan tidak mengurangi tingkat akurasi dalam proses pengelompokan data. Penggunaan data asli juga membantu algoritma *K-Means* dalam mengenali pola serta tingkat kemiripan antar data secara lebih optimal.

Pemrosesan Data

Tahap pemrosesan data dilakukan sebagai langkah awal untuk menyiapkan data sebelum digunakan dalam proses *clustering* menggunakan algoritma *K-Means.* Kualitas data menjadi salah satu faktor yang berpengaruh terhadap hasil pengelompokan, sehingga data perlu dipastikan berada dalam kondisi yang baik sebelum dilakukan analisis. Dalam penelitian terkait metode *clustering* [8], kualitas data yang baik diketahui dapat membantu menghasilkan kelompok data yang lebih representatif dan mendukung proses analisis secara lebih efektif.

Pada penelitian ini, tahapan pemrosesan data meliputi pemeriksaan data, normalisasi data, dan seleksi atribut. Pemeriksaan data bertujuan untuk memastikan bahwa data yang digunakan tidak mengandung *missing*

value maupun ketidaksesuaian data yang dapat memengaruhi hasil analisis. Setelah itu dilakukan normalisasi data untuk menyamakan rentang nilai antar variabel sehingga setiap atribut memiliki kontribusi yang seimbang dalam proses perhitungan jarak. Penelitian mengenai *preprocessing* data [9] menunjukkan bahwa normalisasi merupakan tahapan penting yang dapat membantu meningkatkan kualitas data sebelum dilakukan proses *clustering*.

Selanjutnya dilakukan seleksi atribut untuk menentukan variabel yang sesuai dengan tujuan penelitian. Variabel yang digunakan terdiri atas luas panen, curah hujan, kelembapan udara, suhu rata-rata, dan produksi padi. Pemilihan variabel tersebut didasarkan pada keterkaitannya dengan kondisi produksi padi serta pengaruh faktor iklim terhadap sektor pertanian. Dengan menggunakan atribut yang relevan, proses *clustering* diharapkan mampu menghasilkan kelompok wilayah yang memiliki karakteristik serupa sehingga pola yang terbentuk dapat menggambarkan kondisi produksi padi di Provinsi Sumatera Selatan secara lebih baik.

Pengecekan Data Kosong

Tahapan pengecekan data kosong dilakukan untuk memastikan bahwa data yang digunakan dalam penelitian telah lengkap dan siap untuk dianalisis. Keberadaan data yang hilang atau tidak lengkap berpotensi memengaruhi hasil pengelompokan karena dapat menyebabkan ketidaktepatan dalam proses perhitungan pada algoritma *K-Means*. Berdasarkan penelitian terkait *K-Means Clustering* [10]. tahapan data *cleaning* dan validasi data merupakan bagian penting dalam proses persiapan data karena berpengaruh terhadap kualitas hasil pengelompokan yang dihasilkan. Oleh karena itu, pengecekan data dilakukan sebelum memasuki tahapan analisis.

Pada tahap ini, seluruh data diperiksa untuk mengidentifikasi adanya *missing value* maupun ketidak sesuaian data pada setiap variabel penelitian. Jika ditemukan data yang tidak lengkap, maka dilakukan penyesuaian atau perbaikan berdasarkan informasi yang tersedia agar data tetap dapat digunakan dalam proses analisis. Penelitian mengenai preprocessing data [11], menjelaskan bahwa kualitas data perlu dipastikan terlebih dahulu sebelum dilakukan proses *clustering* agar hasil analisis yang diperoleh lebih optimal. Dengan demikian, data yang digunakan dalam penelitian ini diharapkan memiliki tingkat kelengkapan dan konsistensi yang baik sehingga proses pengelompokan menggunakan algoritma *K-Means* dapat menghasilkan kelompok data yang lebih representatif.

Normalisasi Data

Data yang digunakan dalam penelitian ini memiliki rentang nilai yang berbeda pada setiap variabel. Sebagai contoh, variabel luas panen memiliki nilai yang jauh lebih besar dibandingkan variabel suhu rata-rata yang berada pada rentang nilai yang lebih kecil. Perbedaan skala data tersebut dapat menyebabkan variabel dengan nilai yang lebih besar mendominasi proses perhitungan jarak pada algoritma *K-Means.* Dalam penelitian mengenai normalisasi data pada metode *K-Means* [12]. dijelaskan bahwa proses normalisasi berperan dalam menyamakan rentang nilai antarvariabel sehingga setiap atribut dapat memberikan kontribusi yang lebih seimbang pada proses pengelompokan.

Pada penelitian ini, normalisasi diterapkan pada seluruh variabel yang digunakan, yaitu luas panen, curah hujan, kelembapan udara, suhu rata-rata, dan produksi padi. Tahapan ini dilakukan untuk mengubah data ke dalam skala yang seragam tanpa menghilangkan karakteristik asli dari masing-masing variabel. Dengan rentang nilai yang lebih seimbang, proses *clustering* menggunakan algoritma *K-Means* dapat dilakukan dengan lebih optimal sehingga hasil pengelompokan yang diperoleh mampu menggambarkan kemiripan karakteristik antarwilayah secara lebih akurat.

Proses normalisasi dilakukan untuk menyamakan rentang nilai setiap variabel agar proses perhitungan jarak pada algoritma *K-Means* menjadi lebih seimbang. Metode normalisasi yang digunakan adalah *Min-Max Normalization* dengan rentang nilai 0 sampai 1.

Rumus normalisasi yang digunakan adalah sebagai berikut:

$$y = \frac{x - x_{min}}{x_{max} - x_{min}} \qquad (1)$$

Keterangan:

- x = nilai data
- x_{min} = nilai minimum data
- x_{max} = nilai maksimum data

Seleksi Atribut

Pemilihan atribut dilakukan sebagai langkah untuk menentukan variabel yang akan digunakan dalam proses *clustering*. Tahapan ini penting karena atribut yang dipilih akan memengaruhi hasil pengelompokan yang dihasilkan. Dalam penelitian yang membahas penerapan algoritma *K-Means* [13]. dijelaskan bahwa pemilihan variabel yang sesuai dapat mendukung proses analisis serta menghasilkan kelompok data yang lebih terstruktur dan representatif.

Variabel yang digunakan dalam penelitian ini terdiri atas luas panen, curah hujan, kelembapan udara, suhu rata-rata, dan produksi padi. Kelima variabel tersebut dipilih karena mampu menggambarkan kondisi produksi padi serta karakteristik iklim pada masing-masing wilayah di Provinsi Sumatera Selatan. Penggunaan atribut yang relevan juga membantu proses pengelompokan dilakukan berdasarkan karakteristik yang sesuai dengan tujuan penelitian sehingga hasil yang diperoleh dapat memberikan informasi yang lebih bermakna.

Selain menentukan kualitas hasil pengelompokan, pemilihan atribut yang tepat juga berperan dalam membantu identifikasi pola kemiripan antar wilayah. Penelitian mengenai penerapan metode *K-Means Clustering* pada sektor pertanian [14], menunjukkan bahwa penggunaan variabel yang relevan dapat menghasilkan kelompok data yang lebih representatif dan mudah diinterpretasikan. Oleh karena itu, atribut yang digunakan dalam penelitian ini dipilih berdasarkan keterkaitannya dengan kondisi produksi padi dan faktor iklim agar hasil *clustering* dapat menggambarkan karakteristik wilayah di Provinsi Sumatera Selatan secara lebih optimal.

Implementasi algoritma K-means

Tahap implementasi merupakan proses penerapan algoritma *K-Means* untuk melakukan pengelompokan wilayah berdasarkan karakteristik produksi padi, luas panen, dan faktor iklim di Provinsi Sumatera Selatan. Tahapan ini diawali dengan menentukan jumlah *cluster* yang akan digunakan dalam proses pengelompokan data. Jumlah *cluster* yang ditetapkan akan memengaruhi pola kelompok yang terbentuk sehingga perlu disesuaikan dengan tujuan penelitian. Pada penelitian ini digunakan tiga *cluster* yang selanjutnya diinterpretasikan sebagai kelompok wilayah dengan tingkat kerentanan produksi padi rendah, sedang, dan tinggi.

Setelah jumlah *cluster* ditentukan, langkah berikutnya adalah menetapkan titik pusat awal (centroid) secara acak. Selanjutnya dilakukan penghitungan jarak antara setiap data dengan centroid menggunakan metode *Euclidean Distance.* Menurut penelitian yang membahas penerapan algoritma *K-Means* [15], pengukuran jarak digunakan untuk menentukan kedekatan suatu data terhadap pusat *cluster* sehingga data dapat ditempatkan pada kelompok yang paling sesuai dengan karakteristiknya. Melalui proses tersebut, data yang memiliki tingkat kemiripan yang tinggi akan tergabung ke dalam cluster yang sama.

Proses pengelompokan dilakukan secara berulang melalui beberapa iterasi dengan memperbarui posisi centroid berdasarkan anggota *cluster* yang terbentuk. Iterasi akan terus dilakukan hingga tidak terdapat perubahan keanggotaan *cluster* atau telah mencapai kondisi konvergen. Dalam penelitian terkait metode *K-Means* [16], proses *clustering* dijelaskan mampu mengidentifikasi pola kemiripan antar data sehingga menghasilkan kelompok yang lebih terorganisir dan mudah diinterpretasikan. Hasil akhir dari tahapan ini berupa pengelompokan wilayah yang memiliki karakteristik serupa berdasarkan data produksi padi, luas panen, curah hujan, kelembapan udara, dan suhu rata-rata.

Rumus *Euclidean Distance* yang digunakan adalah sebagai berikut:

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

Keterangan:

- $d(x,y) =$ jarak antara data dan centroid
- $x_i =$ nilai data ke i
- $y_i =$ nilai centroid ke i
- $n =$ jumlah atribut data

Pengujian Model

Tahap pengujian dilakukan untuk mengevaluasi hasil pengelompokan yang telah diperoleh dari penerapan algoritma *K-Means*. Evaluasi dilakukan dengan mengamati karakteristik data pada setiap *cluster* yang terbentuk sehingga perbedaan antar kelompok dapat diidentifikasi dengan lebih jelas. Menurut penelitian mengenai metode *K-Means Clustering* [17], proses pengujian diperlukan untuk mengetahui kualitas hasil pengelompokan berdasarkan tingkat kemiripan karakteristik data yang digunakan.

Pada penelitian ini, hasil *clustering* dianalisis berdasarkan variabel produksi padi, luas panen, curah hujan, kelembapan udara, dan suhu rata-rata pada setiap *cluster*. Analisis tersebut bertujuan untuk mengetahui karakteristik masing-masing kelompok wilayah yang terbentuk serta melihat perbedaan karakteristik antar *cluster*. Dengan membandingkan nilai rata-rata atau karakteristik setiap variabel pada masing-masing kelompok, pola yang terbentuk dari hasil pengelompokan dapat diamati secara lebih jelas.

Penelitian terkait penerapan algoritma *K-Means* [18], menjelaskan bahwa hasil pengelompokan dapat dianalisis melalui tingkat kedekatan karakteristik data pada setiap *cluster.* Oleh karena itu, tahap pengujian dilakukan untuk memastikan bahwa kelompok yang terbentuk memiliki karakteristik yang relatif serupa dalam satu *cluster* dan berbeda dengan *cluster* lainnya. Hasil analisis tersebut kemudian digunakan sebagai dasar untuk menginterpretasikan karakteristik wilayah berdasarkan data produksi padi dan faktor iklim yang digunakan dalam penelitian.

Elbow Method

Elbow Method merupakan salah satu teknik yang sering digunakan untuk menentukan jumlah *cluster* optimal pada algoritma *K-Means*. Metode ini bekerja dengan menghitung nilai *Within Cluster Sum of Squares* (WCSS) atau *Sum of Squared Error* (SSE) pada beberapa alternatif jumlah *cluster.* Nilai WCSS menggambarkan tingkat kedekatan data terhadap centroid pada masing-masing *cluster*, sehingga semakin kecil nilai yang diperoleh maka semakin tinggi tingkat kesamaan data dalam cluster tersebut. Selanjutnya, nilai WCSS divisualisasikan dalam bentuk grafik yang menunjukkan hubungan antara jumlah *cluster* (k) dan nilai WCSS. Jumlah *cluster* yang optimal ditentukan berdasarkan titik yang membentuk pola menyerupai siku (*elbow point*), yaitu ketika penurunan nilai WCSS mulai melambat dan tidak lagi menunjukkan perubahan yang signifikan. [19], *Elbow Method* dapat digunakan untuk membantu menentukan jumlah *cluster* yang paling sesuai karena mampu menunjukkan keseimbangan antara banyaknya *cluster* yang dibentuk dan kualitas pengelompokan data.

Pada penelitian ini, *Elbow Method* digunakan untuk memperoleh jumlah *cluster* yang paling optimal pada data produksi padi di Provinsi Sumatera Selatan yang terdiri dari variabel luas panen, curah hujan, kelembapan udara, suhu rata-rata, dan produksi padi. Penerapan metode ini bertujuan untuk memastikan bahwa jumlah *cluster* yang digunakan mampu merepresentasikan karakteristik data secara baik sebelum dilakukan proses pengelompokan menggunakan algoritma *K-Means*. Hasil yang diperoleh dari *Elbow Method* kemudian dijadikan dasar dalam penentuan nilai (k) sehingga proses *clustering* dapat menghasilkan kelompok wilayah yang memiliki karakteristik serupa berdasarkan tingkat kerentanan produksi padi.

Evaluasi Model

Tahap evaluasi dilakukan untuk menilai sejauh mana algoritma *K-Means* mampu mengelompokkan data wilayah produksi padi di Provinsi Sumatera Selatan. Evaluasi ini dilaksanakan dengan mengamati hasil pengelompokan pada setiap *cluster* serta tingkat kemiripan karakteristik antar data dalam satu kelompok. Dalam metode *K-Means Clustering*, proses evaluasi umumnya menggunakan beberapa ukuran seperti *Davies-Bouldin Index* dan *Silhouette Score* [20], yang berfungsi untuk menilai kualitas cluster yang terbentuk.

Selain itu, tahap evaluasi juga digunakan untuk mengukur kualitas hasil *clustering* yang dihasilkan oleh algoritma *K-Means*. Tujuan evaluasi ini adalah untuk mengetahui tingkat kesamaan data dalam satu *cluster* serta perbedaan karakteristik antar *cluster*. Pada penelitian ini, kualitas *cluster* diukur menggunakan *Davies-Bouldin Index* (DBI) dan *Silhouette Score* [21], yang dapat menggambarkan tingkat kinerja hasil pengelompokan yang diperoleh. Hasil evaluasi tersebut kemudian menjadi dasar dalam menilai efektivitas algoritma *K-Means* dalam mengelompokkan data produksi padi berdasarkan variabel luas panen dan kondisi iklim di Provinsi Sumatera Selatan.

Silhouette Score

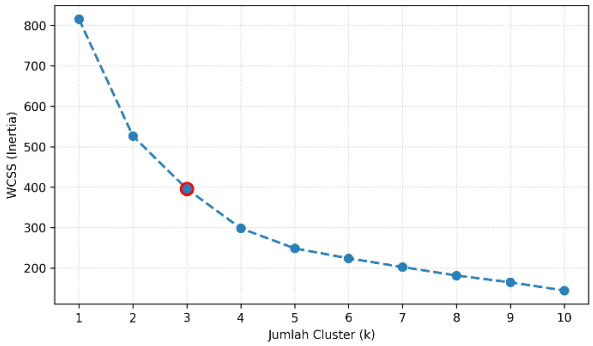
Silhouette Score merupakan salah satu metode evaluasi internal yang digunakan untuk mengukur kualitas hasil *clustering* berdasarkan tingkat kesamaan data dalam suatu *cluster* dan tingkat perbedaannya terhadap *cluster* lain. Nilai *Silhouette Score* berada pada rentang -1 hingga 1, di mana nilai yang semakin mendekati 1 menunjukkan bahwa data dalam suatu *cluster* memiliki tingkat kemiripan yang tinggi dan terpisah dengan baik dari *cluster* lainnya. Sebaliknya, nilai yang mendekati 0 mengindikasikan adanya tumpang tindih antar *cluster*, sedangkan nilai negatif menunjukkan bahwa suatu data kemungkinan ditempatkan pada *cluster* yang kurang tepat. [22], *Silhouette Score* merupakan metrik evaluasi yang efektif untuk menilai kualitas hasil *clustering* karena mampu mengukur tingkat kohesi *cohesion* dan separasi *separation* antar *cluster* secara bersamaan.

Pada penelitian ini, *Silhouette Score* digunakan untuk mengevaluasi kualitas hasil pengelompokan wilayah berdasarkan variabel luas panen, curah hujan, kelembapan udara, suhu rata-rata, dan produksi padi di Provinsi Sumatera Selatan. Metode ini diterapkan untuk memastikan bahwa *cluster* yang terbentuk memiliki tingkat homogenitas yang tinggi di dalam kelompok yang sama serta memiliki perbedaan yang jelas dengan

kelompok lainnya. Hasil pengujian *Silhouette Score* digunakan sebagai pendukung dalam menentukan jumlah *cluster* yang paling optimal setelah dilakukan pengujian menggunakan *Elbow Method*. Dengan demikian, penggunaan kedua metode tersebut diharapkan dapat menghasilkan pengelompokan wilayah yang lebih representatif berdasarkan karakteristik produksi padi dan faktor iklim yang digunakan dalam penelitian.

HASIL DAN PEMBAHASAN

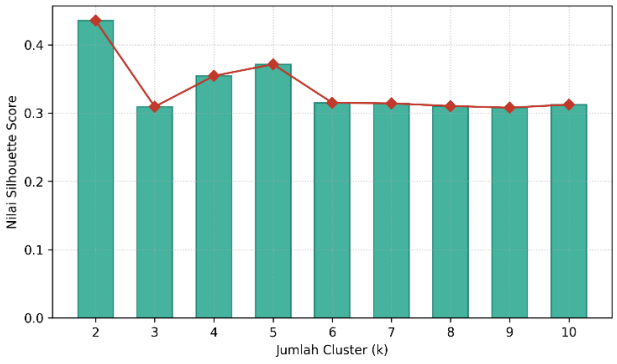
Hasil Penelitian ini menggunakan data produksi padi Provinsi Sumatera Selatan periode 2018–2025 yang terdiri dari variabel luas panen, curah hujan, kelembapan udara, suhu rata-rata, dan produksi padi. Data yang telah dikumpulkan kemudian melalui tahap preprocessing yang meliputi pengecekan data kosong, seleksi atribut, dan normalisasi data sebelum dilakukan proses *clustering* menggunakan algoritma *K-Means.*



Gambar 2 Elbow Method

Elbow Method digunakan untuk menentukan jumlah kluster (k) terbaik dengan melihat penurunan nilai WCSS (*Inertia*). Nilai ini mengukur total jarak kuadrat antara titik data dengan pusat kluster *(centroid)* masing-masing. Semakin kecil nilainya, semakin rapat data tersebut di dalam kelompoknya.

Berdasarkan grafik Gambar 2 Elbow Method di atas, hasil iterasi dari $(k) = 1$ hingga $(k) = 10$ terhadap 136 data padi dan iklim Sumatera Selatan menunjukkan penurunan nilai WCSS yang tajam pada awalnya. Namun, tepat pada titik $(k) = 3$, grafik mengalami tekukan tajam menyerupai siku tangan (*elbow point*) dan mulai berjalan melandai secara konstan pada titik setelahnya.Patahan sikut yang jelas pada $(k) = 3$ ini membuktikan secara matematis bahwa pembagian dataset ke dalam 3 kluster adalah pilihan yang paling optimal. Berdasarkan indikator luas lahan dan produksi, ketiga kelompok ini kemudian ditetapkan menjadi kategori Kerentanan Tinggi, Kerentanan Sedang, dan Kerentanan Rendah.



Gambar 3 Grafik Evaluasi Nilai Silhouette Score

Berdasarkan Gambar 3, hasil pengujian terhadap 136 data menunjukkan nilai tertinggi pada $(k) = 2$ (di atas 0.4), diikuti $(k) = 5$ (0.37), dan $(k) = 3$ sebesar 0.31. Meskipun secara matematis $(k) = 2$ lebih tinggi, jumlah 3 kluster $(k) = 3$ dipilih sebagai keputusan final yang paling optimal. Pembagian menjadi 2 kluster dinilai terlalu global sehingga kurang mampu menggambarkan variasi tingkat kerentanan secara mendalam. Pilihan $(k) = 3$ ini juga konsisten selaras dengan titik tekukan tajam pada pengujian *Elbow Method* sebelumnya. Nilai *Silhouette Score* sebesar 0.31 membuktikan bahwa kelompok yang terbentuk telah memenuhi syarat pemisahan struktur yang cukup (*fair structure*) untuk mengategorikan tingkat Kerentanan Tinggi, Sedang, dan Rendah.

Hasil Preprocessing Data

Tahap preprocessing data dilakukan untuk memastikan kualitas data yang digunakan sebelum memasuki proses *clustering* menggunakan algoritma *K-Means*. Pada tahap ini dilakukan pengecekan terhadap data yang kosong (*missing value)* pada setiap variabel penelitian. Variabel yang digunakan meliputi produksi padi, luas panen, curah hujan, kelembapan udara, dan suhu rata-rata.

Tabel 1. Hasil Pengecekan Missing Value

Nama Kolom	Jumlah Missing Value	Persentase (%)	Status Data
Kabupaten/Kota	0	0	Lengkap / Bersih
Tahun	0	0	Lengkap / Bersih
Luas Panen (Ha)	0	0	Lengkap / Bersih
Produktivitas (Ku/Ha)	0	0	Lengkap / Bersih
Produksi (Ton)	0	0	Lengkap / Bersih
Suhu Rata-Rata (°C)	0	0	Lengkap / Bersih
Kelembapan Udara (%)	0	0	Lengkap / Bersih
Curah Hujan (mm)	0	0	Lengkap / Bersih

Berdasarkan Tabel 1, seluruh variabel penelitian memiliki nilai missing value sebesar 0. Hasil tersebut menunjukkan bahwa tidak terdapat data yang kosong atau tidak lengkap pada dataset yang digunakan. Dengan demikian, seluruh data dapat langsung digunakan pada tahap analisis tanpa memerlukan proses penanganan data hilang (*missing value handling*) seperti imputasi atau penghapusan data.

Kondisi data yang lengkap menjadi salah satu faktor penting dalam proses *clustering* karena dapat membantu menghasilkan pengelompokan yang lebih akurat dan representatif. Oleh karena itu, data produksi padi, luas panen, curah hujan, kelembapan udara, dan suhu rata-rata yang telah memenuhi kriteria kelengkapan data selanjutnya digunakan pada tahap normalisasi dan proses *K-Means Clustering*.

Tabel 2. Hasil Pengelompokan Data Produksi Padi

Kabupaten Kota	Tahun	Luas Panen (Ha)	Produktivitas (Ku/H)	Produksi (Ton)	Suhu Rata-Rata(c)	Kelembapan Udara (%)	Curah Hujan (mm)	Cluster	Kategori
Ogan Komering Ulu	2018	3039,42	46,47	14124,16	24,36	87,11	2562,33	0	Kerentanan Tinggi
Ogan Komering Ilir	2018	95573,8	50,65	484123,06	26,17	88,03	2208,73	1	Kerentanan Rendah
Muara Enim	2018	18082,82	46,57	84206,44	25,86	86,76	2511,98	2	Kerentanan Sedang
Lahat	2018	13966,04	53,96	75360,65	24,36	87,11	2562,33	0	Kerentanan Tinggi
…	…	…	…	…	…	…	…	…	…
Musi Rawas Utara	2025	2435	44,66	10873	26,46	87	2905,2	2	Kerentanan Sedang
Kota Palembang	2025	4100	53	21729	26,5	89,06	3051,93	2	Kerentanan Sedang
Kota Prabumulih	2025	193	44,79	862	26,49	87,63	3076,9	2	Kerentanan Sedang
Kota Pagar Alam	2025	3636	54	19633	23,52	88,73	4032,19	0	Kerentanan Tinggi

Kota Lubuklinggau	2025	1356	54,48	7390	25,09	87,04	3616,27	0	Kerentanan Tinggi

Berdasarkan hasil pengelompokan tersebut, setiap cluster memiliki karakteristik yang berbeda berdasarkan kondisi produksi padi, luas panen, dan faktor iklim

Analisis Hasil *Clustering*

Cluster 0 – Tingkat Kerentanan Rendah

Cluster pertama merupakan kelompok wilayah dengan tingkat kerentanan rendah yang terdiri dari 3 kabupaten, yaitu Kabupaten Banyuasin, Kabupaten Ogan Komering Ilir, dan Kabupaten Ogan Komering Ulu Timur. Kelompok ini memiliki luas panen yang sangat besar dengan rata-rata mencapai 133.448 hektare (berkisar antara 85.002 hingga 230.280 hektare) serta jumlah produksi padi yang sangat tinggi dan stabil dengan rata-rata sebesar 727.258 ton per data pengamatan. Selain itu, kondisi iklim pada kelompok ini cukup mendukung pertumbuhan tanaman padi dengan rata-rata curah hujan 2.581 mm, kelembapan udara 86,67%, dan suhu rata-rata 26,66°C sehingga produktivitas yang dihasilkan jauh lebih baik dibandingkan cluster lainnya.
Ketiga wilayah tersebut secara konsisten masuk dalam kelompok kerentanan rendah selama seluruh periode pengamatan 2018–2025. Hal ini menunjukkan bahwa Kabupaten Banyuasin, Kabupaten Ogan Komering Ilir, dan Kabupaten Ogan Komering Ulu Timur memiliki potensi pertanian yang sangat baik dan menjadi penopang utama ketahanan pangan di Provinsi Sumatera Selatan.

Cluster 1 – Tingkat Kerentanan Sedang

Cluster kedua merupakan kelompok wilayah dengan tingkat kerentanan sedang yang mencakup 13 kabupaten/kota, yaitu Kabupaten Muara Enim, Kabupaten Musi Banyuasin, Kabupaten Ogan Ilir, Kabupaten Penukal Abab Lematang Ilir, Kabupaten Musi Rawas Utara, Kota Palembang, Kota Prabumulih, Kabupaten Ogan Komering Ulu Selatan, Kabupaten Ogan Komering Ulu, Kabupaten Lahat, Kabupaten Musi Rawas, Kabupaten Empat Lawang, dan Kota Lubuklinggau. Pada kelompok ini, produksi padi berada pada kondisi yang cukup stabil dengan rata-rata sebesar 51.936 ton dan luas panen rata-rata 10.913 hektare (berkisar antara 33 hingga 39.039 hektare).
Kondisi iklim pada kelompok ini memiliki rata-rata curah hujan 2.609 mm, kelembapan udara 86,88%, dan suhu rata-rata 26,21°C. Meskipun produksi padi pada kelompok ini tidak setinggi cluster pertama, wilayah-wilayah pada kelompok ini masih memiliki peluang untuk meningkatkan produktivitas melalui pengelolaan pertanian yang lebih optimal, mengingat kondisi iklim yang relatif mendukung namun luas lahan yang belum dimanfaatkan secara maksimal.

Cluster 2 – Tingkat Kerentanan Tinggi

Cluster ketiga merupakan kelompok wilayah dengan tingkat kerentanan tinggi yang mencakup 7 kabupaten/kota, yaitu Kota Pagar Alam, Kabupaten Ogan Komering Ulu, Kabupaten Lahat, Kabupaten Musi Rawas, Kabupaten Empat Lawang, Kota Lubuklinggau, dan Kabupaten Ogan Komering Ulu Selatan. Kelompok ini memiliki luas panen yang paling rendah dengan rata-rata hanya 8.872 hektare (berkisar antara 1.225 hingga 24.368 hektare) dan hasil produksi padi yang cenderung rendah dengan rata-rata sebesar 46.094 ton per data pengamatan.

Kondisi iklim pada kelompok ini ditandai dengan rata-rata curah hujan yang lebih tinggi sebesar 3.373 mm, kelembapan udara 87,27%, dan suhu rata-rata yang lebih rendah yaitu 24,68°C dibandingkan cluster lainnya. Tingginya curah hujan namun rendahnya produksi mengindikasikan bahwa faktor luas lahan yang terbatas menjadi kendala utama di wilayah-wilayah ini. Kota Pagar Alam merupakan satu-satunya wilayah yang secara konsisten berada pada kelompok kerentanan tinggi selama seluruh periode 2018–2025, sehingga memerlukan perhatian dan intervensi kebijakan pertanian yang lebih intensif.

Hasil clustering menunjukkan bahwa luas panen memiliki pengaruh dominan terhadap tingkat produksi padi di Provinsi Sumatera Selatan. Wilayah dengan luas panen yang besar seperti Kabupaten Banyuasin, Kabupaten Ogan Komering Ilir, dan Kabupaten Ogan Komering Ulu Timur cenderung memiliki tingkat produksi padi yang jauh lebih tinggi, sedangkan wilayah dengan luas panen kecil seperti Kota Pagar Alam dan Kota Lubuklinggau memiliki tingkat kerentanan produksi yang lebih tinggi meskipun didukung oleh kondisi iklim yang cukup baik.

SIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, algoritma *K-Means* berhasil diterapkan untuk mengelompokkan wilayah kerentanan produksi padi di Provinsi Sumatera Selatan berdasarkan variabel luas

panen, curah hujan, kelembapan udara, suhu rata-rata, dan produksi padi pada 17 kabupaten/kota periode tahun 2018–2025. Proses *clustering* memetakan 136 data ke dalam 3 kelompok tingkat kerentanan dengan hasil konkret sebagai berikut:

Tingkat Kerentanan Rendah Terdiri dari 3 kabupaten yang secara konsisten masuk dalam kelompok ini selama seluruh periode pengamatan, yaitu Kabupaten Banyuasin, Kabupaten Ogan Komering Ilir, dan Kabupaten Ogan Komering Ulu Timur. Ketiga wilayah ini dicirikan oleh luas panen yang sangat besar (rata-rata 133.448 hektare), produksi padi yang tinggi (rata-rata 727.258 ton), serta kondisi iklim yang relatif stabil, sehingga memiliki risiko penurunan produksi yang paling rendah dibandingkan wilayah lainnya.

Tingkat Kerentanan Sedang Didominasi oleh 7 kabupaten/kota yang selalu berada pada kelompok ini sepanjang 2018–2025, yaitu Kabupaten Muara Enim, Kabupaten Musi Banyuasin, Kabupaten Ogan Ilir, Kabupaten Penukal Abab Lematang Ilir, Kabupaten Musi Rawas Utara, Kota Palembang, dan Kota Prabumulih. Selain itu, Kabupaten Ogan Komering Ulu Selatan, Kabupaten Ogan Komering Ulu, Kabupaten Lahat, Kabupaten Musi Rawas, Kabupaten Empat Lawang, dan Kota Lubuklinggau juga masuk dalam kelompok ini pada sebagian tahun pengamatan. Wilayah-wilayah tersebut memiliki rata-rata produksi sebesar 51.936 ton dengan luas panen rata-rata 10.913 hektare, menunjukkan kondisi pertanian yang cukup stabil namun masih dipengaruhi oleh fluktuasi faktor iklim.

Tingkat Tingkat Kerentanan Tinggi ditandai oleh wilayah yang memiliki produksi dan luas panen relatif rendah. Kota Pagar Alam merupakan satu-satunya wilayah yang secara konsisten berada pada kelompok kerentanan tinggi selama periode 2018–2025. Selain itu, Kabupaten Ogan Komering Ulu, Kabupaten Lahat, Kabupaten Musi Rawas, Kabupaten Empat Lawang, dan Kota Lubuklinggau juga termasuk dalam kelompok ini pada sebagian besar periode pengamatan. Kelompok ini memiliki rata-rata produksi terendah sebesar 46.094 ton dengan luas panen rata-rata 8.872 hektare. Karakteristik iklim pada kelompok ini ditunjukkan oleh curah hujan rata-rata sebesar 3.373 mm, kelembapan udara sebesar 87,27%, dan suhu rata-rata sebesar 24,68°C. Kombinasi luas panen yang relatif rendah serta karakteristik iklim tersebut menunjukkan bahwa wilayah dalam kelompok ini memiliki tingkat kerentanan produksi padi yang lebih tinggi dibandingkan kelompok lainnya sehingga memerlukan perhatian dan strategi pengelolaan pertanian yang lebih intensif.

Secara keseluruhan, penerapan algoritma *K-Means* terbukti mampu mengidentifikasi pola kerentanan produksi padi berdasarkan karakteristik luas panen dan faktor iklim secara terstruktur pada tingkat kabupaten/kota. Hasil pengelompokan yang diperoleh dapat dimanfaatkan sebagai dasar pendukung pengambilan keputusan dalam penyusunan kebijakan pertanian yang lebih tepat sasaran, terutama pada wilayah dengan tingkat kerentanan tinggi. Dengan demikian, hasil penelitian ini diharapkan dapat mendukung upaya peningkatan produktivitas padi serta memperkuat ketahanan pangan di Provinsi Sumatera Selatan.

DAFTAR PUSTAKA

[1] R. Hesananda, “Penerapan Klasterisasi K-Means terhadap Produktivitas Padi di Pulau Sumatera sebagai Strategi Pendukung Ketahanan Pangan,” vol. 6, no. 1, 2025.

[2] N. H. Aryadi, Bagas Putra, “Penerapan algoritma k-means untuk melakukan klasterisasi pada varietas padi,” vol. 7, no. 1, pp. 124–131, 2024.

[3] B. O. Ndasak *et al.*, “KLASTERISASI DATA HASIL PRODUKSI PERTANIAN DAN PETERNAKAN PROVINSI NUSA TENGGARA TIMUR MENGGUNAKAN METODE K-MEANS,” pp. 415–426, 2021.

[4] I. Sanela, A. Nazir, F. Syafria, E. Haerani, and L. Oktavia, “Penerapan Metode Clustering Dengan K-Means Untuk Memetakan Potensi Tanaman Padi di Sumatera,” vol. 5, no. 1, pp. 82–92, 2023, doi: 10.47065/josyc.v5i1.4523.

[5] H. Mukhtar, T. M. Syafutri, R. A. Rahman, and A. Putra, “Jurnal Software Engineering and Information System (SEIS) ANALISIS KESUBURAN PERTANIAN MELALUI IRIGASI DENGAN MENGGUNAKAN METODE K-MEANS CLUSTERING,” vol. 4, no. 2, pp. 102–107, 2024.

[6] My. F. Sena Wijayanto, “Pengelompokkan Produktivitas Tanaman Padi di Jawa Tengah Menggunakan Metode Clustering K-Means,” vol. 13, no. 2, pp. 212–219, 2021.

[7] N. Nurfaidah, Y. C. Hafizha, and H. Yeni, “Transformasi Efisiensi Layanan Kesehatan dengan Sistem Informasi Manajemen Rumah Sakit (SIMRS) di RSUD Provinsi Sulawesi Barat,” vol. 6, no. 3, pp. 1929–1941, 2025.

[8] N. R. Risnawati1, "IMPLEMENTASI METODE K-MEANS CLUSTERING TUNGGAKAN," vol. 4307, no. 1, pp. 118–124, 2022.

[9] W. Priatna, I. Sembiring, A. Setiawan, I. Setyawan, J. Warta, and T. S. Lestari, "NETWORK INTRUSION DETECTION USING TRANSFORMER MODELS AND NATURAL LANGUAGE PROCESSING FOR ENHANCED WEB APPLICATION ATTACK DETECTION Jurnal Nasional Pendidikan Teknik Informatika : JANAPATI | 483," vol. 13, no. 3, pp. 482–493, 2024.

[10] F. A. Muharam, D. Novita, S. Informasi, U. Multi, and D. Palembang, "Penerapan Metode K-Means Clustering Untuk Menentukan Prioritas Persediaan Barang," pp. 72–85.

[11] J. Pratama *et al.*, "Implementasi Algoritma K-Means Clustering untuk Mengelompokkan Wilayah Produksi Kopi di Jawa Barat," pp. 733–745.

[12] V. R. Sari and E. Buulolo, "Implementasi Algoritma K-Means dengan Normalisasi Sigmoidal Untuk Klastering Data Ternak Sapi," vol. 02, no. 01, pp. 1–10, 2023.

[13] J. Novaldi and A. W. Wijayanto, "' Analisis Cluster Kualitas Pemuda di Indonesia pada Tahun 2022 dengan Agglomerative Hierarchical dan K-Means ' ' Clustering Analysis Using Agglomerative Hierarchical and K-Means on Youth Quality Data in Indonesia in 2022 ,'" vol. 12, no. 148, 2023, doi: 10.34010/komputika.v12i2.10348.

[14] M. N. Yohanes Eka Wibawa, "PEMBANGUNAN PERANGKAT LUNAK KOMUNITAS MEDIA MUSIK DENGAN FRAMEWORK REACT NATIVE," vol. VI, pp. 160–170, 2023.

[15] I. Mujahidin, S. H. Hasanah, U. Terbuka, F. C-means, and D. Index, "Comparative analysis of k-means, k-medoids, and fuzzy c-means for clustering provinces in indonesia based on rice production in 2024 1,2," vol. 14, pp. 356–365, 2025, doi: 10.14710/j.gauss.14.2.356-365.

[16] A. S. Putriyekti, W. Setyonugroho, and H. Suroto, "Implementation of hospital information system from 2012-2022 : a bibliometric analysis," vol. 16, no. 2, pp. 158–163, 2022, doi: 10.15562/ijbs.v16i2.471.

[17] B. Hartono, V. Lusiana, I. Husni, and A. Amin, "Perbandingan Proses Klasterisasi Data Menggunakan K-Means Clustering dan Agglomerative Hierarchical Clustering," vol. 12, no. 4, pp. 628–635, 2025, doi: 10.30865/jurikom.v12i4.8766.

[18] Samsuddin and S. dan T. H. , Zulfan, Ade Irfan Saputra, Munawir, "Sistem Informasi Penggunaan Anggaran Takterduga Pada Kantor Badan Pengelolaan Keuangan Aceh Berbasis Web Menggunakan Codeigniter," vol. 5, no. 3, pp. 592–600, 2022.

[19] N. A. Maori, "METODE ELBOW DALAM OPTIMASI JUMLAH CLUSTER PADA K-MEANS CLUSTERING," vol. 14, no. 2, pp. 277–287, 2023.

[20] D. Hartama, S. Oktaviani, and I. Engineering, "OPTIMIZATION OF K-MEANS AND K-MEDOIDS CLUSTERING USING DBI In the increasingly advanced digital era , the industry ' s need for computer science graduates with high technical skill competence is increasing .," vol. XI, no. 2, 2025.

[21] S. A. Muhammad Hafiz*, "KLASTERISASI PENYAKIT MENULAR DI INDONESIA MENGGUNAKAN METODE K-MEANS CLUSTERING," vol. 4, no. 1, pp. 50–57, 2024.

[22] A. Susilo, "EVALUASI EFEKTIFITAS ALGORTIMA K-MEANS DAN DBSCAN DALAM CLUSTERING DATA RUMAH SAKIT UNTUK," vol. 19, no. 1, pp. 43–51, 2024.