

# Penerapan Klasterisasi K-Means terhadap Produktivitas Padi di Pulau Sumatera sebagai Strategi Pendukung Ketahanan Pangan

Rizki Hesananda<sup>1</sup>

<sup>1</sup>Teknologi Informasi / Universitas Siber Indonesia

[hessananda@cyber-univ.ac.id](mailto:hessananda@cyber-univ.ac.id)<sup>1</sup>

**Abstrak** — Produksi padi merupakan aspek krusial dalam ketahanan pangan nasional, khususnya di wilayah Pulau Sumatera yang berperan strategis sebagai lumbung pangan Indonesia. Penelitian ini dilakukan untuk mengidentifikasi pola produksi padi berdasarkan karakteristik luas panen dan hasil produksi tahunan guna menghasilkan segmentasi wilayah produktivitas yang lebih informatif. Tujuan utama penelitian ini adalah melakukan segmentasi data produksi padi dengan metode clustering, serta mengevaluasi hasil klaster menggunakan metrik statistik dan visualisasi data mining. Metodologi yang digunakan mengikuti pendekatan *CRISP-DM*, dengan tahapan eksplorasi data, visualisasi, modeling, dan evaluasi. Algoritma yang digunakan adalah *K-Means Clustering* dengan jumlah klaster sebanyak tiga. Data diperoleh dari Kaggle dan BMKG, mencakup 8 provinsi di Pulau Sumatera dalam rentang waktu 1995 hingga 2020, dengan total 208 entri data observasi tahunan. Hasil penelitian menunjukkan bahwa model clustering yang dibangun memiliki kualitas sangat baik dengan nilai *Silhouette Score* sebesar 0,7147 dan *Davies-Bouldin Index (DBI)* sebesar 0,3752. Visualisasi scatter plot menunjukkan segmentasi yang jelas antara klaster berdasarkan luas panen dan produksi. Kesimpulannya, pendekatan clustering efektif dalam mengelompokkan wilayah berdasarkan karakteristik produksi padi. Implikasi penelitian ini mendukung optimalisasi kebijakan pengelolaan lahan pertanian berbasis data. Bagi pemangku kepentingan industri dan pengambil keputusan, hasil penelitian ini dapat menjadi dasar dalam menentukan prioritas intervensi, distribusi bantuan pertanian, serta perencanaan ketahanan pangan yang berkelanjutan di tingkat regional.

**Kata Kunci:** produksi padi, clustering, K-Means, data mining

## I. PENDAHULUAN

Padi merupakan komoditas pangan utama yang memiliki peran vital dalam menjaga ketahanan pangan nasional [1]. Di Indonesia, sebagian besar kebutuhan karbohidrat masyarakat masih bergantung pada beras sebagai makanan pokok [2]. Oleh karena itu, produksi padi menjadi perhatian utama dalam perencanaan strategis pembangunan pertanian [3]. Pulau Sumatera, sebagai salah satu wilayah agraris dengan keragaman agroekologi yang tinggi, menyumbang produksi padi dalam skala signifikan untuk mendukung suplai pangan nasional [4].

Namun, produksi padi tidak merata antar provinsi dan cenderung dipengaruhi oleh berbagai faktor seperti kondisi iklim, luas lahan, serta praktik pertanian yang diterapkan [5], [6].

Dalam konteks dinamika iklim dan perubahan penggunaan lahan, analisis pola produksi padi secara berkala menjadi hal yang esensial untuk mendukung pengambilan kebijakan yang berbasis data [7], [8]. Upaya segmentasi atau pengelompokan wilayah berdasarkan karakteristik produksi dapat membantu dalam merancang program intervensi yang lebih tepat sasaran [9], [10]. Dengan mengetahui klaster provinsi yang memiliki produktivitas tinggi, sedang, atau rendah, maka pemerintah maupun pemangku kepentingan dapat menetapkan prioritas dalam hal bantuan pertanian, peningkatan kapasitas, maupun pembangunan infrastruktur pendukung [11], [12], [13].

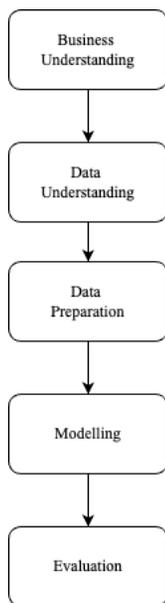
Penelitian sebelumnya telah menampilkan tren visualisasi produksi padi di Pulau Sumatera menggunakan Google Looker Studio, dengan data yang bersumber dari Kaggle dan Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) untuk periode tahun 1995 hingga 2020. Visualisasi tersebut menunjukkan adanya perbedaan signifikan dalam rata-rata produksi per provinsi serta indikasi penurunan produksi sejak tahun 2017 [14]. Namun, pendekatan eksplorasi visual semata belum cukup untuk menyimpulkan adanya segmentasi struktural dalam data. Oleh karena itu, diperlukan pendekatan analitis yang lebih sistematis seperti *data clustering*.

Metode *clustering* dalam data mining dapat mengelompokkan data yang memiliki karakteristik serupa ke dalam satu klaster tanpa perlu adanya label awal (unsupervised learning) [15], [16]. Salah satu algoritma yang umum digunakan adalah *K-Means Clustering*, yang mampu membentuk klaster berdasarkan jarak antar titik dalam ruang fitur. Dalam konteks produksi padi, penggunaan *K-Means* memungkinkan pembentukan kelompok wilayah berdasarkan kombinasi variabel seperti luas panen dan total produksi, sehingga dapat ditemukan pola-pola tersembunyi yang tidak tampak secara kasat mata [17], [18], [19].

Penelitian ini bertujuan untuk melakukan analisis segmentasi wilayah produksi padi di Pulau Sumatera menggunakan algoritma K-Means, dengan pendekatan CRISP-DM sebagai kerangka kerja metodologis. Evaluasi hasil kluster dilakukan menggunakan dua metrik validasi internal, yaitu *Silhouette Score* dan *Davies-Bouldin Index (DBI)*, yang menunjukkan kualitas pemisahan dan kekompakan antar kluster [20]. Penelitian ini diharapkan dapat memberikan kontribusi dalam pemetaan produktivitas wilayah secara komputasional dan mendukung perencanaan strategis pertanian yang berbasis data. Selain itu, temuan dari penelitian ini dapat menjadi acuan awal bagi kajian lanjutan yang mengintegrasikan klasifikasi prediktif dan peramalan produksi padi di masa depan.

## II. METODE

Penelitian ini menggunakan pendekatan Cross Industry Standard Process for Data Mining (CRISP-DM) sebagai kerangka kerja utama dalam pelaksanaan analisis data. CRISP-DM merupakan metodologi yang umum digunakan dalam proses data mining karena mencakup seluruh tahapan secara sistematis, mulai dari pemahaman konteks bisnis hingga evaluasi dan penerapan model. Penelitian ini dilakukan menggunakan platform Google Colab dengan bahasa pemrograman Python, serta pustaka pendukung seperti Pandas, Matplotlib, Seaborn, dan Scikit-learn. Sumber data utama berasal dari Kaggle dan BMKG yang mencakup informasi produksi padi, luas panen, serta variabel iklim selama periode tahun 1995 hingga 2020 dari delapan provinsi di Pulau Sumatera.



Gambar 1. Metode Penelitian CRISP-DM

Adapun penjelasan tahapan CRISP-DM pada Gambar 1 yang dirancang dan digunakan dalam penelitian ini adalah sebagai berikut:

### A. Business Understanding

Tahapan ini bertujuan untuk memahami konteks permasalahan yang dihadapi, yaitu ketidakraturan dalam pola produksi padi antar wilayah di Pulau Sumatera. Penelitian diarahkan untuk mengeksplorasi pola tersebut dan menghasilkan segmentasi wilayah yang dapat digunakan dalam pengambilan keputusan terkait kebijakan pangan dan pertanian.

### B. Data Understanding

Data yang digunakan meliputi produksi padi tahunan, luas panen, curah hujan, suhu rata-rata, dan kelembapan udara. Seluruh data dikompilasi dari sumber terbuka dengan cakupan waktu 26 tahun. Analisis awal dan visualisasi dilakukan untuk memahami struktur data dan distribusi antar variabel, serta mengidentifikasi potensi anomali atau ketidakwajaran dalam data.

### C. Data Preparation

Data dibersihkan dari kolom non-numerik seperti nama provinsi dan tahun untuk menghasilkan matriks fitur yang sesuai dengan kebutuhan algoritma clustering. Pemilihan variabel dilakukan berdasarkan relevansi dengan tujuan segmentasi. Jika diperlukan, proses transformasi atau normalisasi data akan diterapkan untuk menjaga konsistensi skala antar fitur.

### D. Modeling

Tahapan ini melibatkan penerapan algoritma K-Means Clustering untuk mengelompokkan data berdasarkan kemiripan karakteristik antar observasi. Jumlah kluster ditentukan berdasarkan eksplorasi awal dan pertimbangan evaluasi model. Visualisasi hasil model akan digunakan untuk mengamati distribusi kluster pada ruang dua dimensi.

### E. Evaluation

Evaluasi model dilakukan dengan menggunakan metrik validasi internal, yaitu *Silhouette Score* dan *Davies-Bouldin Index (DBI)*. Kedua metrik ini akan digunakan untuk menilai kualitas pemisahan dan kekompakan antar kluster yang terbentuk, guna memastikan bahwa model yang dihasilkan memiliki performa yang layak untuk dianalisis lebih lanjut.

## III. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil dari implementasi tahapan *data mining* berdasarkan kerangka CRISP-DM yang telah dirancang pada bab sebelumnya. Seluruh proses analisis dilakukan menggunakan Google Colab dengan dataset produksi padi Pulau Sumatera tahun 1995 hingga 2020. Hasil-hasil pada tiap tahapan ditampilkan dalam bentuk visualisasi, tabel, dan interpretasi mendalam untuk menggambarkan proses serta temuan dari penelitian ini.

### A. Business Understanding

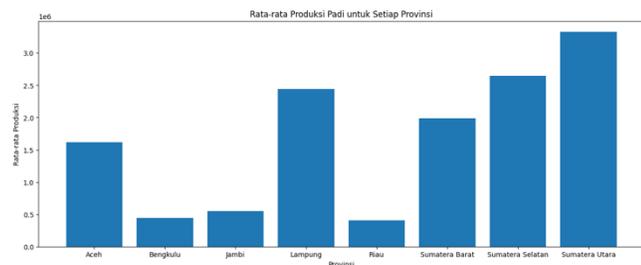
Produksi padi merupakan indikator utama ketahanan pangan di Indonesia. Pulau Sumatera sebagai salah satu sentra produksi padi nasional memiliki karakteristik geografis dan iklim yang beragam, yang berpengaruh langsung terhadap variasi produksi antar provinsi. Meskipun data historis produksi padi telah tersedia secara terbuka, belum banyak dilakukan analisis segmentasi berbasis data mining yang mampu mengelompokkan provinsi berdasarkan karakteristik produksi dan luas panen. Hal ini menyebabkan strategi pengembangan pertanian seringkali bersifat umum dan tidak mempertimbangkan pola spesifik tiap wilayah.

Kebutuhan utama dalam konteks ini adalah menghasilkan kluster provinsi yang memiliki pola produktivitas padi yang serupa, guna mendukung perumusan kebijakan pertanian yang lebih terarah. Dengan mengetahui kelompok-kelompok wilayah yang memiliki karakteristik mirip, misalnya dari sisi luas panen atau volume produksi, maka dapat dilakukan alokasi sumber daya, program peningkatan produksi, serta intervensi teknologi secara lebih efisien. Selain itu, pendekatan ini dapat membuka peluang untuk membandingkan antar kluster guna mengidentifikasi praktik terbaik dari kelompok dengan performa tinggi.

Penelitian ini menggunakan pendekatan data mining dengan metode unsupervised learning untuk menemukan pola tersembunyi dalam data produksi padi. Proses analisis mengikuti kerangka CRISP-DM, dimulai dari eksplorasi data hingga evaluasi model. Dengan menerapkan algoritma K-Means Clustering, diharapkan dapat diperoleh segmentasi yang bermakna secara statistik dan dapat divisualisasikan secara jelas. Pemahaman yang dihasilkan dari proses ini menjadi pondasi penting untuk menyusun kebijakan pertanian berbasis data (data-driven agriculture policy), yang lebih adaptif terhadap kondisi wilayah masing-masing.

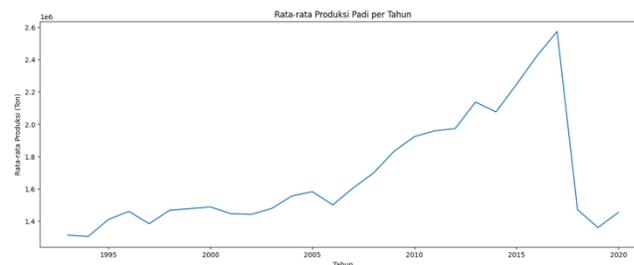
### B. Data Understanding

Tahap ini bertujuan untuk memahami struktur, sebaran, dan pola awal dari data produksi padi yang dianalisis. Data diperoleh dari sumber terbuka yang terpercaya, mencakup delapan provinsi di Pulau Sumatera selama periode 1995 hingga 2020. Variabel utama yang digunakan adalah Produksi Padi (dalam ton) dan Luas Panen (dalam hektar), sedangkan variabel pendukung mencakup curah hujan, kelembapan, dan suhu rata-rata. Analisis awal dilakukan untuk mengamati konsistensi data, identifikasi pola waktu, serta potensi hubungan antar variabel yang dapat memengaruhi proses clustering.



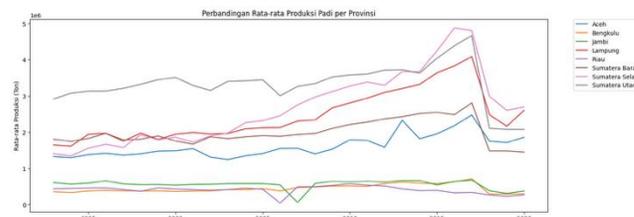
Gambar 2. Rata-rata Produksi Padi per Provinsi

Visualisasi pada Gambar 2 menunjukkan rata-rata produksi padi dari masing-masing provinsi selama periode penelitian. Tampak bahwa Provinsi Sumatera Utara memiliki rata-rata produksi paling tinggi dibandingkan dengan provinsi lainnya, diikuti oleh Sumatera Selatan dan Lampung. Perbedaan nilai produksi yang mencolok antar provinsi menunjukkan adanya potensi segmentasi atau klusterisasi berdasarkan kapasitas produksi masing-masing wilayah. Selain itu, hasil ini mengindikasikan bahwa penggunaan teknik clustering dapat memberikan insight tentang kelompok provinsi dengan produktivitas serupa.



Gambar 3. Rata-rata Produksi Padi per Tahun (1995–2020)

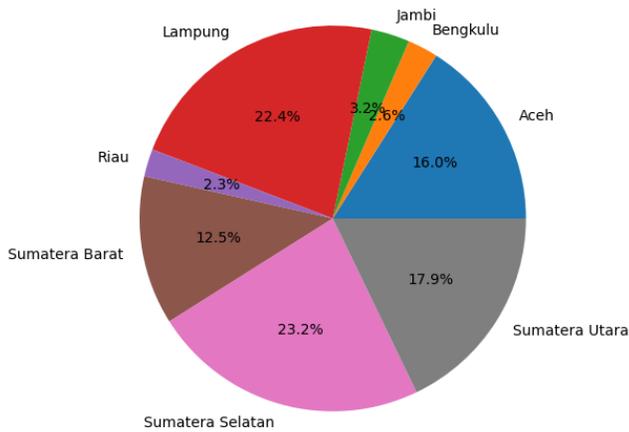
Grafik pada Gambar 3 memperlihatkan fluktuasi produksi padi rata-rata dari seluruh provinsi selama kurun waktu 26 tahun. Terlihat adanya tren menurun pada rentang tahun 2017 hingga 2020, yang kemungkinan berkaitan dengan faktor eksternal seperti perubahan iklim, kebakaran hutan, atau degradasi lahan. Informasi ini memberikan konteks penting dalam interpretasi tren dan dapat menjadi dasar untuk integrasi data waktu dalam studi lanjutan. Meski fluktuatif, tren umum menunjukkan stabilitas pada sebagian besar periode awal.



Gambar 4. Perbandingan Tren Produksi Antar Provinsi

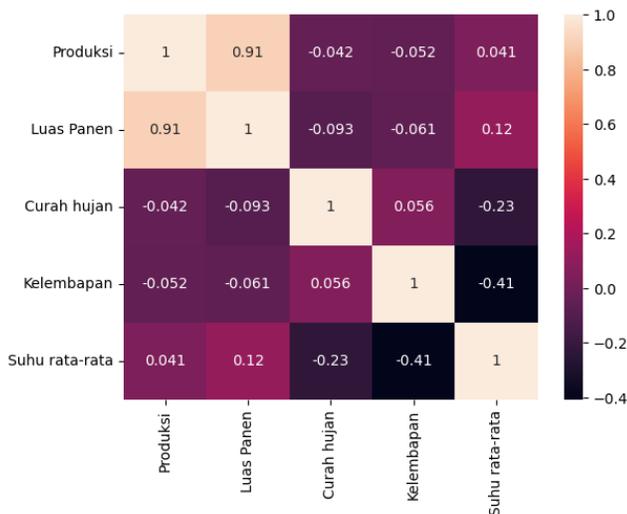
Visualisasi pada Gambar 4 menyajikan perbandingan tren produksi antar provinsi dari tahun ke tahun. Sumatera Utara kembali menunjukkan performa yang paling stabil dan tinggi, sementara provinsi lain seperti Bengkulu dan Riau memiliki pola produksi yang lebih fluktuatif. Pola-pola ini memperkuat asumsi awal bahwa provinsi-provinsi memiliki karakteristik produksi yang berbeda, dan metode clustering dapat digunakan untuk mengelompokkan provinsi-provinsi dengan pola yang serupa secara kuantitatif. Grafik ini juga mempermudah identifikasi perubahan performa wilayah dari waktu ke waktu.

Proporsi Produksi Padi untuk Setiap Provinsi pada Tahun 2020



Gambar 5. Proporsi Produksi Padi per Provinsi pada Tahun 2020

Pada grafik pada Gambar 5 ditampilkan proporsi kontribusi produksi padi dari masing-masing provinsi untuk tahun terakhir dalam dataset, yaitu 2020. Hasil menunjukkan bahwa kontribusi terbesar masih berasal dari Sumatera Utara, Sumatera Selatan, dan Lampung. Visualisasi ini berguna untuk menyoroti dominasi produksi pada tahun tertentu, serta sebagai pembandingan dalam evaluasi tren waktu. Informasi ini juga membantu dalam menganalisis keberlanjutan produksi dan ketahanan pangan di masing-masing wilayah.



Gambar 6. Heatmap Korelasi antar Variabel Numerik

Heatmap korelasi pada Gambar 6 memberikan gambaran hubungan linear antar variabel dalam dataset. Ditemukan bahwa Produksi memiliki korelasi sangat kuat dengan Luas Panen, sementara hubungan dengan variabel iklim seperti suhu, curah hujan, dan kelembapan cenderung lemah hingga sedang. Hasil ini menjadi landasan untuk pemilihan fitur dalam proses clustering, di mana variabel Produksi dan Luas Panen diprioritaskan karena memberikan kontribusi paling signifikan dalam membedakan karakteristik observasi.

### C. Data Preparation

Tahap *data preparation* merupakan proses penting dalam memastikan bahwa data yang akan digunakan dalam pemodelan memiliki kualitas dan format yang sesuai dengan kebutuhan algoritma. Dalam penelitian ini, data disiapkan agar kompatibel dengan metode *K-Means Clustering*, yang memerlukan input dalam bentuk numerik dan memiliki distribusi nilai yang seragam.

Langkah pertama dalam proses ini adalah seleksi fitur. Dari keseluruhan atribut yang tersedia, dua variabel utama dipilih sebagai input untuk proses clustering, yaitu Produksi (dalam ton) dan Luas Panen (dalam hektar). Kedua variabel ini dipilih berdasarkan korelasi yang signifikan dan peran langsungnya dalam menentukan karakteristik produktivitas wilayah. Sementara itu, atribut non-numerik seperti Provinsi dan Tahun dihapus karena tidak relevan dalam proses pemodelan *unsupervised learning*.

Selanjutnya dilakukan pembersihan data, dengan pengecekan nilai hilang (*missing values*) dan pencilan (*outliers*). Berdasarkan eksplorasi awal pada tahap *data understanding*, dataset tidak mengandung nilai kosong maupun pencilan ekstrem yang mengganggu, sehingga tidak diperlukan proses imputasi atau transformasi tambahan. Hal ini memungkinkan analisis dilanjutkan secara langsung pada data asli.

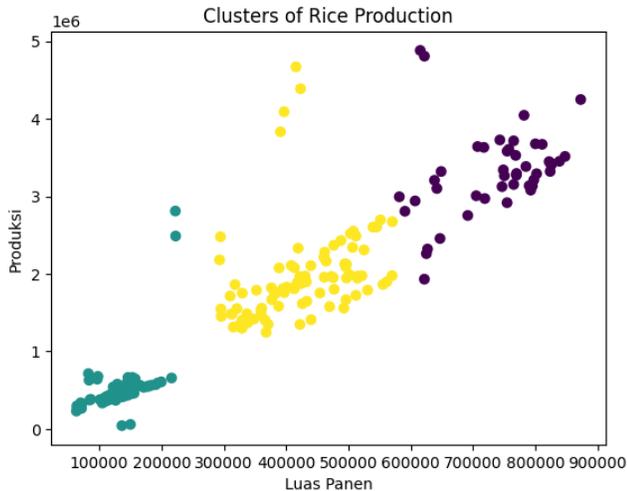
Terakhir, dilakukan penskalaan fitur apabila diperlukan. Meskipun variabel Produksi dan Luas Panen memiliki satuan yang berbeda, distribusinya cukup sebanding dan tidak menunjukkan dominasi yang mengganggu keseimbangan model. Oleh karena itu, pada tahap awal ini, proses *normalisasi* atau *standarisasi* tidak diterapkan. Namun, evaluasi model tetap mempertimbangkan performa visual dan statistik untuk memastikan bahwa hasil clustering tidak bias terhadap skala variabel.

### D. Modeling

Tahap pemodelan dalam pendekatan *CRISP-DM* bertujuan untuk membentuk model *clustering* yang mampu mengelompokkan data ke dalam beberapa segmen berdasarkan kesamaan karakteristik. Dalam penelitian ini, algoritma yang digunakan adalah *K-Means Clustering*, yang merupakan salah satu metode *unsupervised learning* paling populer untuk segmentasi data numerik.

Algoritma *K-Means* bekerja dengan cara membagi data ke dalam sejumlah  $k$  kluster berdasarkan jarak Euclidean antar titik terhadap pusat kluster (*centroid*). Nilai  $k$  ditentukan secara eksploratif berdasarkan jumlah kelompok yang diharapkan dari pola distribusi data. Dalam konteks penelitian ini, nilai  $k$  ditetapkan sebanyak 3 kluster, yang diasumsikan dapat merepresentasikan tiga tingkat kategori wilayah produksi: tinggi, sedang, dan rendah.

Model dilatih menggunakan dua fitur utama, yaitu Produksi dan Luas Panen. Data dua dimensi ini memungkinkan representasi visual yang intuitif, sehingga memudahkan interpretasi hasil. Setelah proses pelatihan model, setiap observasi diberi label kluster tertentu sesuai kedekatannya dengan pusat kluster masing-masing.



Gambar 7. Scatter Plot Hasil Clustering Berdasarkan Produksi dan Luas Panen

Visualisasi pada Gambar 7 menunjukkan penyebaran data dalam ruang dua dimensi dengan pewarnaan berbeda untuk setiap kluster. Titik-titik pada grafik merepresentasikan data observasi tahunan per provinsi, dengan koordinat berdasarkan Luas Panen dan Produksi. Perbedaan warna menunjukkan bahwa model berhasil memisahkan data ke dalam tiga kelompok yang cukup terpisah. Kluster dengan nilai produksi dan luas panen tinggi terkonsentrasi pada satu sisi, sementara kelompok dengan karakteristik produksi lebih rendah terletak pada sisi lainnya.

#### E. Evaluation

Tahap evaluasi dilakukan untuk menilai kualitas model K-Means yang telah dibangun pada tahap sebelumnya. Dalam penelitian ini, digunakan dua metrik validasi internal, yaitu Silhouette Score dan Davies-Bouldin Index (DBI). Kedua metrik ini memberikan gambaran tentang seberapa baik model mengelompokkan data dalam ruang fitur yang telah disiapkan.

Silhouette Score mengukur seberapa mirip suatu data dengan klasternya sendiri dibandingkan dengan kluster lain. Nilai berkisar antara -1 hingga 1, dengan nilai mendekati 1 menunjukkan kluster yang sangat baik (semakin besar semakin baik). Sementara itu, Davies-Bouldin Index mengevaluasi kekompakan dan pemisahan antar kluster. Semakin kecil nilai DBI, semakin baik kualitas kluster yang terbentuk.

Tabel 1. Evaluasi Efektivitas Sistem Informasi Perhitungan Lembur

Metrik Evaluasi	Nilai	Interpretasi
Silhouette Score	0.7147	Sangat Baik
Davies-Bouldin Index	0.3752	Sangat Baik

Hasil evaluasi pada Tabel 1 menunjukkan bahwa model *K-Means* dengan tiga kluster memberikan pemisahan yang sangat baik dan struktur kluster yang jelas. Nilai Silhouette Score yang tinggi menunjukkan bahwa observasi lebih dekat dengan kluster mereka sendiri dibandingkan dengan kluster lain. Sementara itu, nilai DBI yang rendah menunjukkan bahwa jarak antar kluster cukup besar dan distribusi dalam kluster relatif padat. Hasil ini mengonfirmasi bahwa struktur kluster yang dihasilkan layak untuk dianalisis lebih lanjut dan digunakan dalam pengambilan keputusan.

#### IV. KESIMPULAN

Penelitian ini bertujuan untuk melakukan segmentasi wilayah produksi padi di Pulau Sumatera menggunakan algoritma *K-Means Clustering* dengan pendekatan kerangka kerja CRISP-DM. Dataset yang digunakan mencakup periode 1995–2020, dengan dua fitur utama yaitu Produksi dan Luas Panen. Visualisasi dan eksplorasi awal menunjukkan adanya variasi signifikan antar provinsi, baik dari segi tren tahunan maupun proporsi kontribusi terhadap produksi nasional. Hasil dari proses pemodelan menunjukkan bahwa pembentukan tiga kluster dapat mengelompokkan observasi data ke dalam kategori produksi tinggi, sedang, dan rendah secara jelas. Evaluasi model menggunakan Silhouette Score sebesar 0,7147 dan Davies-Bouldin Index sebesar 0,3752 menunjukkan kualitas clustering yang sangat baik, dengan pemisahan kluster yang optimal dan kekompakan internal yang tinggi.

Secara umum, pendekatan *unsupervised learning* dalam penelitian ini berhasil mengungkap pola tersembunyi dalam data produksi padi, yang sebelumnya sulit diidentifikasi hanya melalui visualisasi konvensional. Segmentasi ini dapat menjadi dasar dalam merancang strategi pembangunan pertanian berbasis data, baik untuk pendistribusian bantuan, peningkatan kapasitas petani, maupun investasi infrastruktur pertanian. Bagi pemangku kepentingan industri pangan, hasil penelitian ini memberikan wawasan yang relevan dalam penyusunan kebijakan yang responsif terhadap kondisi wilayah. Sementara itu, bagi pembuat kebijakan, informasi kluster dapat digunakan untuk mengidentifikasi wilayah prioritas dalam penguatan ketahanan pangan regional. Dengan demikian, penelitian ini berkontribusi pada pendekatan *data-driven policy making* dalam sektor pertanian Indonesia.

DAFTAR PUSTAKA

- [1] A. Juansa, A. Maulana, M. Lubis, and A. Wijaya, "Ketahanan Pangan: Swasembada Pangan dan Implikasinya terhadap Pertumbuhan Ekonomi di Indonesia," 2025, Accessed: Jul. 28, 2025. [Online]. Available: <https://books.google.com/books?hl=id&lr=&id=jo5dEQAAQBAJ&oi=fnd&pg=PA16&dq=ketahanan+pangan+indonesia&ots=igl2UikQVj&sig=mkv5eX9tP3yEQ8YiAvuZOsmSUIU>
- [2] M. H.-J. 'ULŪM A.-Q. I. Pengetahuan and undefined 2024, "Ketahanan Pangan Menuju Kedaulatan Pangan Bagi Indonesia Adalah Keniscayaan," *journal.cidescimi.org*, vol. 13, pp. 347–358, doi: 10.14710/jnc.v13i4.42107.
- [3] S. Fatimah, A. Rasyid, ... A. A.-... of G. and, and undefined 2024, "Kebijakan Makan Bergizi Gratis di Indonesia Timur: Tantangan, Implementasi, dan Solusi untuk Ketahanan Pangan," *journal.intelekmadani.org*, vol. 4, no. 1, pp. 14–21, 2024, Accessed: Jul. 28, 2025. [Online]. Available: <https://journal.intelekmadani.org/index.php/jgpi/article/view/641>
- [4] L. Budiman, D. S.-J. P. Masyarakat, and undefined 2024, "Resiliensi penguatan ketahanan pangan daerah di Indonesia," *ejournal.ipdn.ac.id*, vol. 01, no. 2, pp. 63–71, 2024, Accessed: Jul. 28, 2025. [Online]. Available: <https://ejournal.ipdn.ac.id/jpa/article/view/5113>
- [5] D. Andaresta, D. Retnowati, ... A. F.-P. S., and undefined 2024, "Analisis Faktor Yang Mempengaruhi Ketahanan Pangan Di Indonesia," *unars.ac.id*, Accessed: Jul. 28, 2025. [Online]. Available: <https://unars.ac.id/ojs/index.php/prosidingSDGs/article/view/4983>
- [6] R. Rahim, N. Utami, R. Nurfalah, ... Y. A.-I. J. O., and undefined 2024, "Dinamika ketahanan pangan: analisis pengaruh luas panen padi, konsumsi beras, harga beras, dan jumlah penduduk terhadap produksi padi di wilayah sentra padi," *j-innovative.org*, vol. 01, no. 2, pp. 63–71, 2024, Accessed: Jul. 28, 2025. [Online]. Available: <https://j-innovative.org/index.php/Innovative/article/view/12524>
- [7] D. Sukmawati, R. Rivaldi, D. Mahmiludin, P. Agribisnis, P. Universitas, and W. Mukti, "ANALISIS KETAHANAN PANGAN INDONESIA: TANTANGAN DAN STRATEGI BERKELANJUTAN DALAM ERA TRANSFORMASI SOSIAL-EKONOMI," *ejournal.papanda.org*, vol. 04, no. 1, 2025, Accessed: Jul. 28, 2025. [Online]. Available: <https://ejournal.papanda.org/index.php/jira/article/view/1811>
- [8] M. H.-J. E. P. D. Agribisnis and undefined 2024, "Food Estate: Ancaman Ataukah Peluang Bagi Ketahanan Pangan Indonesia?," *jepa.ub.ac.id*, vol. 8, pp. 1313–1326, 2024, doi: 10.21776/ub.jepa.2024.008.04.7.
- [9] D. Agustinngtyas *et al.*, "Analisis Spasial Ketahanan Pangan di Indonesia Tahun 2022 dengan Model Geographically Weighted Logistic Regression," *journal.sragenkab.go.id*, vol. 8, no. 1, pp. 45–62, 2024, doi: 10.32630/sukowati.v8i1.438.
- [10] A. Zahra, S. Ada, A. A.-J. of C. and, and undefined 2024, "Implementasi Clustering Algoritma K-Means Pada Produksi Beras di Provinsi Jawa Timur Tahun 2022," *journal-computing.org*, vol. 5, no. 3, pp. 2775–2496, 2024, Accessed: Jul. 28, 2025. [Online]. Available: <https://www.journal-computing.org/index.php/journal-cisa/article/view/485>
- [11] K. I. Hafsa, M. Toriq, A. Hijrah, T. K. Wijayanti, and R. Kurniawan, "Implementasi Algoritma K-Means pada Pengelompokan Ketahanan Pangan di Indonesia Menurut Kabupaten/Kota," *prosiding-senada.upnjatim.ac.id*, vol. 2024, Accessed: Jul. 28, 2025. [Online]. Available: <https://prosiding-senada.upnjatim.ac.id/index.php/senada/article/view/161>
- [12] N. Hanan Cahyaningrum, N. Mas, and R. Yotenka, "Analisis Aspek Ketahanan Pangan Indonesia dengan Hard dan Soft Clustering, 2022," *journal.trunojoyo.ac.id*, vol. 3, no. 2, 2025, Accessed: Jul. 28, 2025. [Online]. Available: <https://journal.trunojoyo.ac.id/rekayasa/article/view/21774>
- [13] D. Saputra, ... K. A.-J. (Jurnal M., and undefined 2024, "PENERAPAN METODE K-MEANS CLUSTERING UNTUK PEMETAAN PENGELOMPOKAN LAHAN PRODUKSI JAGUNG," *mail.ejournal.itn.ac.id*, vol. 3, no. 2, 2025, Accessed: Jul. 28, 2025. [Online]. Available: <https://mail.ejournal.itn.ac.id/index.php/jati/article/view/10541>
- [14] R. Hesana, D. R.-I. J. I. Komputer, and undefined 2024, "IMPLEMENTASI GOOGLE LOOKER STUDIO UNTUK ANALISIS TREN DAN VISUALISASI DATA (STUDI KASUS: PRODUKSI PADI PULAU SUMATERA)," *ejournal.cyber-univ.ac.id* R Hesana, DF RacmaInnotech: Jurnal Ilmu Komputer, Sistem Informasi dan, 2024•*ejournal.cyber-univ.ac.id*, vol. 1, no. 2, 2024, Accessed: Jul. 28, 2025. [Online]. Available: <https://ejournal.cyber-univ.ac.id/index.php/innotech/article/view/45>
- [15] A. Putri, ... N. S.-J. J. of, and undefined 2024, "Pengelompokan Provinsi di Indonesia Menggunakan Time Series Clustering pada Sektor Ekspor Nonmigas," *ejournal.ung.ac.id*, vol. 6, no. 1, pp. 16–22, 2024, doi: 10.37905/jjom.v6i1.21921.
- [16] E. Susilawati, A. H.-J. I. U. B. Jambi, and undefined 2024, "Pengaruh Luas Lahan Panen Padi dan Tenaga Kerja Sektor Pertanian terhadap Produksi Padi di Provinsi Jambi," *ji.unbari.ac.id*, vol. 24, no. 2, pp. 1829–1832, 2024, doi: 10.33087/jiubj.v24i2.5334.
- [17] S. Abdullah, Z. F.-J. I. SAINS, and undefined 2025, "Analisis Produksi Cabai Rawit Indonesia

- Menggunakan Algoritma K-Means Clustering,” *jurnal.alimspublishing.co.id*, vol. 3, no. 1, pp. 66–74, 2025, doi: 10.59024/jiti.v3i1.1024.
- [18] H. Mukhtar *et al.*, “Perbandingan Algoritma K-Means Dan K-Medoids Untuk Pengelompokan Hasil Pertanian Di Kabupaten Cirebon,” *ejournal.itn.ac.id*, vol. 4, no. 2, pp. 2089–3272, 2024, Accessed: Jul. 28, 2025. [Online]. Available: <https://www.ejournal.itn.ac.id/index.php/jati/article/view/8326>
- [19] H. Mukhtar *et al.*, “Implementasi Clustering Algoritma K-Means Pada Produksi Beras di Provinsi Jawa Timur Tahun 2022,” *journal-computing.org*, vol. 4, no. 2, pp. 2089–3272, 2024, Accessed: Jul. 28, 2025. [Online]. Available: <https://www.journal-computing.org/index.php/journal-cisa/article/view/485>
- [20] R. F.-J. of I. System, undefined Applied, and undefined 2024, “Penerapan K-Means Clustering Untuk Pemetaan Produktivitas Padi Dan Prediksi Panen Di Kabupaten Indramayu,” *journal.stmikjayakarta.ac.id*, vol. 8, no. 3, pp. 589–605, 2024, doi: 10.52362/jisamar.v8i3.1572.