E-ISSN : 3025-1311
https://journal.cib.institute/index.php/kohesi

Kohesi: Jurnal Multidisiplin Saintek
Volume 11 No 3 Tahun 2026

PENGELOMPOKAN DATA PROGRAMA PENYULUHAN PERTANIAN BERDASARKAN KARAKTERISTIK PERTANIAN MENGGUNAKAN METODE CLUSTERING (STUDI KASUS: BPP SELESAI KABUPATEN LANGKAT)

Dwi Anjani [1*], Yani Maulita [2], I Gusti Prahmana [3]
[1,2,3] Program Studi Sistem Informasi, STMIK Kaputama Binjai
E-mail: dwiianjanii18@gmail.com

ABSTRACT

The Agricultural Extension Office (BPP) of Selesai District, Langkat Regency holds a large volume of Agricultural Extension Program data with varying characteristics of Commodity, Land Area, and Production, which makes manual grouping difficult and hinders the preparation of well-targeted extension programs. This study aims to apply a Data Mining approach using the Clustering technique with the K-Means Algorithm to group Agricultural Extension Program data at BPP Selesai, Langkat Regency, based on the variables of Commodity, Land Area (Ha), and Production (Ton). The dataset consists of 555 records from 2023-2025, processed through the Knowledge Discovery in Database (KDD) stages of data selection, preprocessing, transformation, data mining, and evaluation, and computed using the MATLAB application. The developed program allows the user to select the number of clusters (K), run the clustering process repeatedly, and keep the result with the smallest SSE (Sum of Squared Error) as the best result. With K=3 clusters, the results yield a best SSE of 294.3742, where Cluster 1 consists of 309 records characterized by Corn with low land area and production, Cluster 2 consists of 201 records characterized by Rice with moderate land area and production, and Cluster 3 consists of 45 records characterized by Rice with high land area and production. These clustering results can serve as supporting information for agricultural extension workers in formulating more targeted extension programs.

Keywords: Data Mining, Clustering, K-Means, Agricultural Extension Program, MATLAB

ABSTRAK

Balai Penyuluhan Pertanian (BPP) Kecamatan Selesai, Kabupaten Langkat memiliki data Programa Penyuluhan Pertanian dalam jumlah besar dengan karakteristik Komoditas, Luas Lahan, dan Jumlah Produksi yang beragam, sehingga sulit dikelompokkan

Article History
Received: Agustus 2026
Reviewed:Agustus 2026
Published: Agustus 2026

Plagiarism Checker No 237.u873247.fh8374

Copyright : Author
Publish by : Kohesi

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

secara manual untuk mendukung penyusunan program penyuluhan yang tepat sasaran. Penelitian ini bertujuan menerapkan metode Data Mining dengan teknik Clustering menggunakan Algoritma K-Means untuk mengelompokkan Data Programa Penyuluhan Pertanian pada BPP Selesai Kabupaten Langkat berdasarkan variabel Komoditas, Luas Lahan (Ha), dan Jumlah Produksi (Ton). Data yang digunakan berjumlah 555 data Tahun 2023-2025 yang diolah melalui tahapan Knowledge Discovery in Database (KDD), meliputi data selection, preprocessing, transformation, data mining, dan evaluation, kemudian diproses menggunakan aplikasi MATLAB. Program yang dibangun memungkinkan pengguna memilih jumlah cluster (K), menjalankan proses clustering secara berulang, serta menyimpan hasil dengan nilai SSE (Sum of Squared Error) terkecil sebagai hasil terbaik. Hasil penelitian dengan jumlah cluster K=3 menghasilkan nilai SSE terbaik sebesar 294,3742, dengan Cluster 1 sebanyak 309 data berkarakteristik Jagung serta luas lahan dan produksi rendah, Cluster 2 sebanyak 201 data berkarakteristik Padi serta luas lahan dan produksi menengah, dan Cluster 3 sebanyak 45 data berkarakteristik Padi serta luas lahan dan produksi tinggi. Hasil pengelompokan tersebut dapat digunakan sebagai informasi pendukung bagi penyuluh pertanian dalam menyusun program penyuluhan pertanian yang lebih tepat sasaran.

Kata Kunci: Data Mining, Clustering, K-Means, Programa Penyuluhan Pertanian, MATLAB

1. PENDAHULUAN

Sektor pertanian merupakan salah satu sektor yang memiliki peran penting dalam mendukung ketahanan pangan serta perekonomian masyarakat di Indonesia. Balai Penyuluhan Pertanian (BPP) Kecamatan Selesai, Kabupaten Langkat mengelola data Programa Penyuluhan Pertanian dalam jumlah besar dengan karakteristik Komoditas, Luas Lahan, dan Jumlah Produksi yang beragam. Namun, data tersebut selama ini masih sebatas digunakan sebagai laporan administratif dan belum dimanfaatkan secara optimal dalam proses analisis data, sehingga pihak BPP belum memiliki gambaran yang jelas mengenai pola dan karakteristik pertanian di setiap wilayah binaan.

Beberapa penelitian terdahulu menunjukkan bahwa Algoritma K-Means efektif digunakan untuk mengelompokkan data pertanian berdasarkan karakteristik produksi, luas lahan, maupun ketahanan pangan antarwilayah. Meskipun demikian, penelitian-penelitian tersebut umumnya berfokus pada data produksi pertanian secara umum di tingkat kabupaten atau provinsi, sehingga masih terdapat *research gap* berupa belum adanya penerapan metode clustering pada data Programa Penyuluhan Pertanian di tingkat BPP kecamatan, khususnya BPP Selesai Kabupaten

Langkat. *Novelty* dari penelitian ini terletak pada objek penelitian, yaitu Data Programa Penyuluhan Pertanian pada BPP Selesai dengan variabel Komoditas, Luas Lahan (Ha), dan Jumlah Produksi (Ton) yang diolah menggunakan aplikasi MATLAB.

Penelitian ini bertujuan menerapkan metode Data Mining dengan teknik Clustering menggunakan Algoritma K-Means untuk mengelompokkan Data Programa Penyuluhan Pertanian pada BPP Selesai Kabupaten Langkat berdasarkan variabel Komoditas, Luas Lahan (Ha), dan Jumlah Produksi (Ton), menggunakan data sebanyak 555 data pada tahun **2023-2025**. Hasil pengelompokan tersebut diharapkan dapat digunakan sebagai informasi pendukung bagi penyuluh pertanian dalam menyusun program penyuluhan pertanian yang lebih tepat sasaran.

2. METODE PENELITIAN

Metode penelitian yang digunakan dalam penelitian ini adalah pendekatan Knowledge Discovery in Database (KDD), yaitu serangkaian proses sistematis untuk memperoleh informasi atau pengetahuan dari sekumpulan data. Tahapan KDD yang diterapkan meliputi data selection, data preprocessing, transformation, data mining, dan evaluation. Data yang digunakan adalah Data Programa Penyuluhan Pertanian yang diperoleh dari Balai Penyuluhan Pertanian (BPP) Kecamatan Selesai, Kabupaten Langkat, sebanyak 555 data pada tahun 2023–2025, dengan variabel penelitian berupa Komoditas, Luas Lahan (Ha), dan Jumlah Produksi (Ton).Pada tahap data selection dipilih atribut Nama Kelompok, Komoditas, Luas Lahan (Ha), dan Jumlah Produksi (Ton) yang relevan dengan tingkat produktivitas pertanian. Tahap data preprocessing dilakukan untuk membersihkan data dari ketidakkonsistenan penulisan, misalnya “cabe” diseragamkan menjadi “cabai” dan “padii” menjadi “Padi”, tanpa mengurangi jumlah data yang tetap 555 data. Pada tahap transformation, atribut Komoditas, Luas Lahan, dan Jumlah Produksi diubah menjadi nilai numerik (kode) agar dapat diproses menggunakan Algoritma K-Means, yang mensyaratkan data dalam bentuk angka.

Proses data mining dilakukan menggunakan Algoritma K-Means Clustering dengan bantuan aplikasi MATLAB. Program yang dibangun memungkinkan pengguna memilih jumlah cluster (K) melalui dropdown (2, 3, atau 4), kemudian menjalankan proses clustering menggunakan fungsi kmeans dengan parameter Start='sample', MaxIter=100, EmptyAction='singleton', Replicates=20, dan rng(0) agar hasil konsisten setiap kali dijalankan. Setiap kali tombol “Proses Cluster” ditekan, program menjalankan satu kali percobaan K-Means dan menghitung nilai *Sum of Squared Error (SSE)* sebagai ukuran evaluasi. Hasil dengan SSE terkecil disimpan sebagai hasil terbaik (bestSSE, bestIdx, bestC). Apabila SSE tidak mengalami perbaikan selama 5 kali percobaan berturut-turut, program akan menandai bahwa proses clustering telah mencapai kondisi optimal.

Setelah cluster dinyatakan optimal, label cluster diurutkan ulang menggunakan fungsi urutkanCluster() berdasarkan nilai centroid Komoditas secara menurun, kemudian Luas Lahan dan Produksi secara menaik, sehingga “Cluster 1” selalu merujuk pada kelompok data yang sama pada setiap kali program dijalankan. Hasil akhir clustering ditampilkan dalam bentuk tabel ringkasan serta grafik tiga dimensi (scatter3) menggunakan aplikasi MATLAB.

3. HASIL DAN PEMBAHASAN

Penerapan Algoritma K-Means dengan bantuan software MATLAB terhadap 555 data Programa Penyuluhan Pertanian BPP Selesai Kabupaten Langkat Tahun 2023–2025 menghasilkan pengelompokan optimal pada jumlah cluster K=3, dengan nilai SSE (Sum of Squared Error) terbaik sebesar 294,3742 yang diperoleh setelah 6 kali percobaan clustering. Cluster 1 beranggotakan 309 data dengan titik pusat centroid (2,27; 1,00; 1,05), didominasi Komoditas Jagung dengan luas lahan 0,42–28,25 Ha dan produksi 1,96–184,20 Ton. Cluster 2 beranggotakan 201 data dengan centroid (1,27; 2,28; 2,36), didominasi Komoditas Padi dengan luas lahan 28,26–56,09 Ha dan produksi 184,21–366,44 Ton. Cluster 3 beranggotakan 45 data dengan centroid (1,07; 4,73; 4,73), didominasi Komoditas Padi dengan luas lahan 83,94–111,77 Ha dan produksi 548,69–730,92 Ton.

Hasil ini sejalan dengan penelitian sebelumnya yang menunjukkan bahwa Algoritma K-Means efektif digunakan untuk mengelompokkan data pertanian berdasarkan karakteristik komoditas dan tingkat produksi, sehingga mampu mengubah data yang sebelumnya hanya berfungsi sebagai arsip administratif menjadi informasi yang lebih bermakna. Kebaruan (*novelty*) penelitian ini terletak pada penerapan clustering K-Means secara khusus pada data Programa Penyuluhan Pertanian tingkat BPP kecamatan, dengan mekanisme pemilihan hasil terbaik berdasarkan SSE terkecil dari beberapa kali percobaan serta pengurutan ulang label cluster agar konsisten pada setiap proses.

Cluster 1 dengan karakteristik Jagung berlahan sempit dapat menjadi prioritas program penyuluhan intensifikasi lahan dan peningkatan produktivitas, Cluster 2 dengan karakteristik Padi berlahan menengah dapat diarahkan pada penguatan pola tanam dan akses sarana produksi, sedangkan Cluster 3 dengan karakteristik Padi berlahan luas dan produksi tinggi dapat dijadikan percontohan (best practice) bagi kelompok tani lain. Keterbatasan penelitian ini adalah variabel yang digunakan hanya mencakup Komoditas, Luas Lahan, dan Jumlah Produksi, tanpa memperhitungkan faktor lain seperti kondisi tanah, iklim, maupun aspek sosial ekonomi petani, sehingga hasil pengelompokan perlu dilengkapi dengan variabel lain pada penelitian selanjutnya.

KESIMPULAN

Penerapan metode Data Mining dengan teknik Clustering menggunakan Algoritma K-Means berhasil mengelompokkan 555 Data Programa Penyuluhan Pertanian pada BPP Selesai Kabupaten Langkat berdasarkan variabel Komoditas, Luas Lahan (Ha), dan Jumlah Produksi (Ton). Dengan jumlah cluster K=3, diperoleh nilai SSE terbaik sebesar 294,3742, dengan Cluster 1 sebanyak 309 data berkarakteristik Jagung serta luas lahan dan produksi rendah, Cluster 2 sebanyak 201 data berkarakteristik Padi serta luas lahan dan produksi menengah, dan Cluster 3 sebanyak 45 data berkarakteristik Padi serta luas lahan dan produksi tinggi.

Hasil pengelompokan tersebut memiliki implikasi praktis bagi BPP Selesai Kabupaten Langkat dalam menyusun program penyuluhan pertanian yang lebih tepat sasaran, karena data yang sebelumnya hanya digunakan sebagai arsip administratif kini dapat dimanfaatkan sebagai dasar pengambilan keputusan. Implikasi teoretisnya adalah memperkaya bukti empiris mengenai efektivitas Algoritma K-Means dalam pengelompokan data pertanian pada level unit kerja

penyuluhan yang lebih spesifik. Keterbatasan penelitian ini adalah cakupan variabel yang hanya meliputi komoditas, luas lahan, dan produksi, tanpa memperhitungkan faktor kondisi tanah, iklim, maupun sosial ekonomi petani.

Bagi penelitian selanjutnya, disarankan untuk menggunakan data pertanian yang lebih lengkap dan terbaru serta menambahkan variabel lain yang berkaitan dengan karakteristik pertanian agar hasil clustering memberikan informasi yang lebih luas. Bagi BPP Selesai Kabupaten Langkat, hasil pengelompokan ini dapat dijadikan salah satu dasar pertimbangan dalam menentukan prioritas program penyuluhan pertanian pada setiap kelompok tani. Program MATLAB yang telah dibangun juga disarankan untuk dikembangkan lebih lanjut dengan menambahkan fitur penyimpanan hasil clustering dan pelaporan agar dapat digunakan secara lebih praktis oleh penyuluh pertanian.

DAFTAR PUSTAKA

[1] Ryana, D., Sundari, S., Kustiari, T., & Mawardi, A. K. (2025). Digitalisasi Informasi Penyuluhan: Upaya Peningkatan Peran Balai Penyuluhan Pertanian (BPP) Sumberwringin Melalui Website dan Instagram.
https://doi.org/10.25047/agropross.2025.811

[2] Septianto, M. A., Faqih, A., & Rinaldi, A. R. (2025). Klasterisasi Data Produksi Pertanian di Kabupaten Cirebon dengan Algoritma K-Means. Jurnal Informatika dan Teknik Elektro Terapan, 13(2).
https://doi.org/10.23960/jitet.v13i2.6174

[3] Nuraeni, F., & Firdaus, M. H. M. (2025). Model Klasterisasi Data Pertanian Menggunakan Algoritma K-Means. Jurnal Algoritma, 22(2), 890–898.
https://doi.org/10.33364/algoritma/v.22-2.2054

[4] Mujahidin, I., & Hasanah, S. H. (2025). Comparative Analysis of K-Means, K-Medoids, and Fuzzy C-Means for Clustering Provinces in Indonesia Based on Rice Production in 2024. Jurnal Gaussian, 14(2), 356–365.
https://doi.org/10.14710/j.gauss.14.2.356-365

[5] Prestisya, Y. E., Rahma, N. N., & Aningthia, D. (2025). Analisis Pengelompokan Tingkat Ketahanan Pangan Provinsi di Indonesia Tahun 2019–2023 Menggunakan Algoritma K-Means.

[6] Widyatami, A. I., & Reistiani, V. M. (2023). Clustering Wilayah Potensi dan Strategi Pengembangan Komoditas Unggulan Tanaman Hortikultura dan Palawija Level Kecamatan di Sumatera Barat Tahun 2021. Seminar Nasional Official Statistics, 2023(1), 41–52.
https://doi.org/10.34123/semnasoffstat.v2023i1.1737

[7] Aldi, M. A., & Fatah, Z. (2025). Implementasi K-Means Clustering dalam Pengelompokan Data Kunjungan. Jurnal Ilmiah, 2(1), 13–19.

[8] Annas, M., & Wahab, S. N. (2023). Data Mining Methods: K-Means Clustering Algorithms. International Journal of Cyber and IT Service Management (IJCITSM), 3(1), 40–47.
https://iiast.iaicpublisher.org/ijcitsm/index.php/IJCITSM/article/view/122

[9] Dicha Mutia Dhani, Relita Buaton, & I Gusti Prahmana. (2024). Penerapan Metode K-Means Clustering untuk Menentukan Kepuasan Mahasiswa Terhadap Fasilitas Sarana dan Prasarana Kampus di STMIK Kaputama Binjai. Bridge: Jurnal Publikasi Sistem Informasi dan Telekomunikasi, 2(3), 229–243.

https://doi.org/10.62951/bridge.v2i3.170

[10] Sulistiyawati, A., & Supriyanto, E. (2021). Implementasi Algoritma K-Means Clustering dalam Penentuan Siswa Kelas Unggulan. Jurnal Tekno Kompak, 15(2), 25. https://doi.org/10.33365/jtk.v15i2.1162

[11] Nasution, A. N. (2024). Systematic Literature Review: Software Matematika MATLAB Sebagai Media Belajar untuk Meningkatkan Kemampuan Komputasi Siswa. 2(4), 20–27.

[12] Setyaningtyas, S., Nugroho, B. I., & Arif, Z. (2022). Penerapan Data Mining Teknik Clustering Algoritma K-Means. Jurnal Teknoif Teknik Informatika Institut Teknologi Padang, 10(2), 52–61.

[13] Tapsi, S. A., & Baga, L. M. (2025). Analisis Klaster Pengembangan Sektor Pertanian di Kabupaten Bogor Berbasis Karakteristik Kewirausahaan Pertanian. https://doi.org/10.20961/region.v20i1.87909

[14] Khairunnisa, N. F., Saidah, Z., Hapsari, H., & Wulandari, E. (2021). The Agricultural Extension Agent's Role on the Level of Corn Farming Production. Jurnal Penyuluhan, 17(2), 113–125. https://jurnal.ipb.ac.id/index.php/jupe/article/view/33656

[15] Pengelompokan Populasi Hewan Ternak Menggunakan Metode Clustering (Studi Kasus: Dinas Pertanian dan Ketahanan Pangan Kabupaten Langkat). (2021). STMIK Kaputama Binjai.