

**ANALISIS KLASTERISASI PENERIMA PUPUK SUBSIDI TAHUN 2025
MENGUNAKAN ALGORITMA *K-MEANS++* DAN *FUZZY C-MEANS***

SKRIPSI

Diajukan Untuk Memenuhi Sebagian Syarat Guna
Memperoleh Gelar Sarjana Komputer (S.Kom.)
Pada Program Studi Sistem Informasi



OLEH :

UMAR AL FARUQ

NPM: 2113030003

FAKULTAS TEKNIK DAN ILMU KOMPUTER (FTIK)
UNIVERSITAS NUSANTARA PERSATUAN GURU REPUBLIK INDONESIA
UN PGRI KEDIRI

2025

Skripsi oleh :

UMAR AL FARUQ

NPM : 2113030003

Judul

**ANALISIS KLASTERISASI PENERIMA PUPUK SUBSIDI TAHUN 2025
MENGUNAKAN ALGORITMA *K-MEANS++* DAN *FUZZY C-MEANS***

Telah disetujui untuk diajukan Kepada
Panitia Ujian/Sidang Skripsi Program Studi Sistem Informasi
FTIK UN PGRI Kediri

Tanggal: 29 desember 2025

Pembimbing I



Rini Indriati, M.Kom

NIDN. 0725057003

Pembimbing II



Aidina Ristyan, M.Kom

NIDN. 0721018801

Skripsi oleh :

UMAR AL FARUQ

NPM : 2113030003

Judul

**ANALISIS KLASTERISASI PENERIMA PUPUK SUBSIDI TAHUN 2025
MENGUNAKAN ALGORITMA *K-MEANS++* DAN *FUZZY C-MEANS***

Telah dipertahankan di depan Panitia Ujian/Sidang Skripsi

Program Studi Sistem Informasi FTIK UN PGRI Kediri

Pada Tanggal: 29 Desember 2025

Dan Dinyatakan telah Memenuhi Persyaratan

Panitia Penguji :

1. Ketua : Rini Indriati, S.Kom, M.Kom

[.....]

2. Penguji 1 : Rina Firliana, S.Kom, M.Kom

[.....]

3. Penguji 2 : Aidina Ristyawan, S.Kom, M.Kom

[.....]



Mengetahui,

Dekan Fakultas Teknik dan Ilmu Komputer

Dr. Sutisno, M.Si

NIDN. 0007076801

PERNYATAAN

Yang bertanda tangan di bawah ini saya,

Nama : Umar Al Faruq
Jenis Kelamin : Laki-laki
Tempat/tgl. lahir : Nganjuk / 12 Desember 2001
NPM : 2113030003
Fak/Prodi. : Teknik Dan Ilmu Komputer/ S1-Sistem Informasi

menyatakan dengan sebenarnya, bahwa dalam Skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu perguruan tinggi, dan sepanjang pengetahuan saya tidak terdapat karya tulis atau pendapat yang pernah diterbitkan oleh orang lain, kecuali yang secara sengaja dan tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Kediri, 29 Desember 2025

Yang Menyatakan



UMAR AL FARUQ

NPM: 2113030003

MOTTO

“Seburuk Apapun Hidup Tetaplah Jadi Diri Sendiri”.

ABSTRAK

Umar Al Faruq: *ANALISIS KLASTERISASI PENERIMA PUPUK SUBSIDI TAHUN 2025 MENGGUNAKAN ALGORITMA K-MEANS++ DAN FUZZY C-MEANS*, Skripsi, Program Studi Sistem Informasi, Fakultas Teknik dan Ilmu Komputer, Universitas Nusantara PGRI Kediri, 2025.

Kata Kunci: Pupuk Subsidi, Data Mining, *Clustering*, *K-Means++*, *Fuzzy C-Means*, CRISP-DM.

Distribusi pupuk subsidi merupakan kebijakan strategis pemerintah dalam mendukung peningkatan produktivitas pertanian dan menjaga ketahanan pangan nasional. Namun, dalam pelaksanaannya masih ditemukan berbagai permasalahan, seperti ketidaktepatan sasaran penerima, ketimpangan distribusi antarpetani, serta pengelolaan data penerima yang belum optimal. Kompleksitas data penerima pupuk subsidi yang semakin meningkat menuntut adanya pendekatan berbasis data untuk mengelompokkan penerima secara objektif berdasarkan karakteristik yang dimiliki. Penelitian ini bertujuan untuk melakukan klasterisasi penerima pupuk subsidi tahun 2025 menggunakan algoritma K-Means++ dan Fuzzy C-Means. Metode penelitian yang digunakan adalah data mining dengan pendekatan CRISP-DM (Cross Industry Standard Process for Data Mining) yang meliputi tahapan pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan deployment. Dataset yang digunakan berjumlah 729 data penerima pupuk subsidi dengan variabel luas lahan, pupuk urea, pupuk NPK, dan pupuk organik. Seluruh data dinormalisasi menggunakan metode Min–Max sebelum dilakukan proses klasterisasi.

Hasil penelitian menunjukkan bahwa penerapan algoritma K-Means++ menghasilkan delapan cluster ($K = 8$) dengan nilai Silhouette Score sebesar 0,756, Davies–Bouldin Index sebesar 0,241, dan Calinski–Harabasz Index sebesar 7.901,15. Sementara itu, penerapan algoritma Fuzzy C-Means menghasilkan tujuh cluster ($K = 7$) dengan nilai Silhouette Score sebesar 0,749, Davies–Bouldin Index sebesar 0,320, dan Calinski–Harabasz Index sebesar 5.338,80. Cluster yang terbentuk menggambarkan pengelompokan penerima pupuk subsidi berdasarkan perbedaan karakteristik luas lahan dan jumlah pupuk yang diterima.

Kesimpulan dari penelitian ini menunjukkan bahwa kedua proses klasterisasi mampu mengelompokkan penerima pupuk subsidi tahun 2025 ke dalam beberapa kelompok dengan karakteristik kebutuhan pupuk yang berbeda. Hasil klasterisasi dapat dimanfaatkan sebagai informasi pendukung dalam perencanaan dan evaluasi distribusi pupuk subsidi agar lebih tepat sasaran.

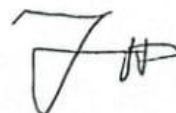
KATA PENGANTAR

Puji syukur kehadiran Allah Tuhan Yang Maha Kuasa atas rahmat-Nya sehingga skripsi berjudul "*Analisis Klasterisasi Penerima Pupuk Subsidi Tahun 2025 Menggunakan Algoritma K-Means++ dan Fuzzy C-Means*" dapat diselesaikan sebagai salah satu syarat memperoleh gelar Sarjana Komputer pada Program Studi Sistem Informasi Fakultas Teknik dan Ilmu Komputer Universitas Nusantara PGRI Kediri. Pada kesempatan ini, penulis menyampaikan terima kasih dan penghargaan yang setulus-tulusnya kepada:

1. Bapak Dr. Zainal Afandi, M.Pd selaku Rektor UN PGRI Kediri.
2. Bapak Dr. Sulistiono, M.Si. selaku Dekan Fakultas Teknik dan Ilmu Komputer UN PGRI Kediri.
3. Bapak Dr. Sucipto, M.Kom. selaku Ka Prodi Sistem Informasi Fakultas Teknik dan Ilmu Komputer UN PGRI Kediri.
4. Ibu Rini Indriati, M. Kom. selaku Pembimbing pertama skripsi ini.
5. Bapak Aidina Ristyan, M.Kom. selaku Pembimbing kedua skripsi ini.
6. Kedua orang tua dan adik-adik saya yang telah mendukung dalam do'a, dan motivasi dalam penyelesaian skripsi ini.
7. Teman-teman yang selalu membantu memberi semangat dalam penyelesaian skripsi ini.
8. Pihak-pihak lain yang tidak dapat disebutkan satu persatu, yang telah banyak membantu menyelesaikan Skripsi ini.

Penulis menyadari bahwa skripsi ini masih memiliki keterbatasan, sehingga kritik dan saran yang membangun sangat diharapkan. Semoga skripsi ini dapat memberikan manfaat, khususnya bagi dunia pendidikan.

Kediri, 29 Desember 2025



UMAR AL FARUQ

NPM: 2113030003

DAFTAR ISI

Halaman Sampul	i
Halaman Persetujuan.....	ii
Halaman Pengesahan	iii
PERNYATAAN.....	iv
Halaman Motto.....	v
ABSTRAK	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR.....	x
DAFTAR TABEL	xi
DAFTAR LAMPIRAN	xii
BAB I PENDAHULUAN	1
A. Latar Belakang	1
B. Identifikasi Masalah	3
C. Batasan Masalah.....	3
D. Rumusan Masalah	3
E. Tujuan Penelitian	4
F. Manfaat Penelitian	4
BAB II KAJIAN PUSTAKA	5
A. Kajian Teori	5
B. Kajian Hasil Penelitian Terdahulu	14
C. Kerangka Berpikir.....	17
BAB III METODE PENELITIAN	19
BAB IV HASIL DAN PEMBAHASAN	27
A. Data Hasil Penelitian.....	27
B. Pembahasan.....	45
BAB V PENUTUP.....	48
A. Kesimpulan	48

B. Saran.....	49
DAFTAR PUSTAKA	50
LAMPIRAN.....	54

DAFTAR GAMBAR

Gambar 2. 1 Tahapan Data Mining.....	6
Gambar 2. 2 Kerangka Berpikir	17
Gambar 3. 1 Tahapan CRISP-DM	19
Gambar 4. 1 Boxplot Variabel Penelitian	33
Gambar 4. 2 Distribusi Jumlah Data K-Means++.....	41
Gambar 4. 3 Distribusi Jumlah Data Fcm.....	43

DAFTAR TABEL

Tabel 4. 1 variabel penelitian	27
Tabel 4. 2 Data Penelitian	28
Tabel 4. 3 Statistik Deskriptif Data Penerima Pupuk Subsidi Tahun 2025	31
Tabel 4. 4 Hasil penerapan normalisasi Min-Max	36
Tabel 4. 5 Statistik Data Setelah Normalisasi	36
Tabel 4. 6 Hasil Analisis Cluster K-Means++	37
Tabel 4. 7 Distribusi Cluster K-Means++	38
Tabel 4. 8 Hasil Evaluasi Cluster Fuzzy C-Means	39
Tabel 4. 9 Distribusi Cluster Fuzzy C-Means	40
Tabel 4. 10 Tabel Hasil Indikator Kualitas Cluster.....	40
Tabel 4. 11 Hasil Clustering K-Means++ Pada Data Penerima Pupuk Subsidi ...	42
Tabel 4. 12 Hasil Clustering FCM Pada Data Penerima Pupuk Subsidi	44

DAFTAR LAMPIRAN

Lampiran 1 Surat Izin Penelitian.....	54
Lampiran 2 Surat Balasan Laporan Penelitian.....	55
Lampiran 3 Surat Keterangan Bebas Similarity.....	56
Lampiran 4 Bukti Halaman Awal Similarity	57
Lampiran 5 Bukti Screenshot Submit Artikel/Loa.....	58

BAB I

PENDAHULUAN

A. Latar Belakang

Subsidi pupuk merupakan kebijakan penting pemerintah untuk mendukung peningkatan produktivitas pertanian dan menjaga ketahanan pangan nasional. Pemerintah mengalokasikan anggaran setiap tahun untuk memastikan kebutuhan pupuk subsidi dapat terpenuhi oleh petani. Selain pupuk anorganik, pemanfaatan pupuk organik juga menjadi bagian dari upaya peningkatan kualitas tanah dan efisiensi penggunaan input pertanian (Sucipto et al. 2019). Meskipun demikian, implementasi kebijakan tersebut masih menghadapi berbagai kendala, antara lain ketidaktepatan sasaran penerima, ketimpangan distribusi antar petani, dan keterbatasan pembaruan data penerima yang menyebabkan distribusi tidak merata dan tidak sesuai kebutuhan di lapangan (Ika Ari Sasmita, Rini Indriati, and M. Najibulloh Muzaki 2021).

Kendala distribusi pupuk subsidi menunjukkan bahwa pendataan dan mekanisme verifikasi penerima belum berjalan optimal. Sistem pengelolaan data penerima pupuk subsidi masih bersifat manual dan kurang terintegrasi sehingga rawan terjadi bias dan kesalahan pencatatan. Kompleksitas data penerima semakin meningkat seiring bertambahnya jumlah petani penerima pupuk subsidi serta variasi kebutuhan pupuk pada tiap wilayah. Kondisi ini menuntut adanya metode pengelolaan dan analisis data yang lebih objektif dan berbasis komputer.

Teknik *Clustering* merupakan salah satu pendekatan analisis data yang dapat digunakan untuk mengelompokkan penerima pupuk subsidi berdasarkan karakteristik tertentu sehingga mendukung penyusunan kebijakan distribusi yang lebih tepat sasaran. *Clustering* mampu mengidentifikasi kelompok petani dengan karakteristik serupa melalui pengolahan data numerical dan digunakan secara luas dalam bidang pertanian untuk pemetaan kebutuhan input dan segmentasi petani (Aldino et al. 2021). Penerapan *Clustering* pada data pupuk subsidi memungkinkan pemetaan kebutuhan pupuk yang lebih terukur dan sistematis.

Metode *Clustering* memiliki beragam algoritma dengan karakteristik yang berbeda. *K-Means++* merupakan metode hard *Clustering* yang mampu

menghasilkan *Cluster* yang tegas karena proses inisialisasi pusat *Cluster* dilakukan secara lebih terarah sehingga mengurangi bias awal (Daoudi et al. 2021). Sebaliknya, *Fuzzy C-Means* memberikan derajat keanggotaan pada tiap *Cluster* sehingga fleksibel untuk menangani data dengan karakteristik tumpang tindih dan heterogen (Krasnov et al. 2023). Penggunaan kedua metode ini memberikan pendekatan analisis yang saling melengkapi untuk memahami pola data penerima pupuk subsidi.

Analisis hasil *Clustering* diperlukan untuk menentukan jumlah *Cluster* optimal dan memastikan kualitas *Cluster* yang terbentuk. Validasi *Cluster* dapat dilakukan menggunakan beberapa indeks evaluasi seperti *Silhouette Score*, *Davies-Bouldin Index*, dan *Calinski-Harabasz Index* (Zhang, LiuY and Wang 2023). Penggunaan multi-indeks validasi memberikan gambaran yang lebih komprehensif terhadap kualitas *Cluster* dibandingkan hanya menggunakan satu indeks (Meng et al. 2023). Pendekatan ini diperlukan untuk menghasilkan *Cluster* yang akurat, stabil, dan representatif.

Penelitian terdahulu terkait penerapan clustering pada distribusi pupuk subsidi menunjukkan bahwa metode ini mampu mengelompokkan penerima pupuk berdasarkan karakteristik tertentu, seperti luas lahan dan jumlah pupuk yang diterima, sehingga dapat dijadikan dasar dalam pemerataan distribusi pupuk subsidi (Zalfa, Nugroho, and Harini 2023). Namun, sebagian besar penelitian masih berfokus pada penggunaan satu algoritma clustering serta evaluasi menggunakan indikator terbatas, sehingga analisis terhadap kualitas dan karakteristik cluster yang dihasilkan belum dilakukan secara menyeluruh. Selain itu, pemanfaatan data terbaru dengan pendekatan evaluasi multi-indeks validitas internal masih relatif terbatas, sehingga membuka peluang untuk dilakukan analisis klasterisasi yang lebih komprehensif.

Berdasarkan kesenjangan tersebut, penelitian ini dilakukan untuk menganalisis klasterisasi penerima pupuk subsidi tahun 2025 menggunakan algoritma *K-Means++* dan *Fuzzy C-Means*, serta mengevaluasi kinerja hasil klasterisasi berdasarkan *Silhouette Score*, *Davies-Bouldin Index*, dan *Calinski-Harabasz Index*. Analisis ini diharapkan dapat menghasilkan cluster yang akurat, stabil, dan

mudah diinterpretasikan sebagai dasar pendukung pengambilan keputusan dalam distribusi pupuk subsidi yang lebih tepat sasaran.

B. Identifikasi Masalah

Permasalahan dalam distribusi pupuk subsidi muncul karena pendataan penerima belum dilakukan secara optimal. Distribusi pupuk subsidi masih menghadapi ketidaktepatan sasaran penerima, ketidakseimbangan distribusi antartetani, dan pembaruan data penerima yang belum dilakukan secara berkala. Proses verifikasi penerima pupuk sebagian besar masih dilakukan secara manual sehingga rawan kesalahan pencatatan dan bias. Selain itu, pengelompokan penerima pupuk belum memanfaatkan analisis berbasis data yang objektif, sehingga kebutuhan pupuk belum dapat dipetakan sesuai karakteristik penerimanya. Metode *Clustering* belum digunakan secara komparatif dalam menganalisis data penerima pupuk subsidi, dan penentuan jumlah *Cluster* optimal belum divalidasi menggunakan multi-indeks sehingga kualitas *Cluster* yang terbentuk belum dinilai secara menyeluruh.

C. Batasan Masalah

Untuk menjaga fokus penelitian dan menghindari pembahasan yang terlalu luas, penelitian ini dibatasi pada proses pengelompokan penerima pupuk subsidi tahun 2025 berdasarkan variabel numerik yang terdapat dalam dataset. Variabel yang dianalisis meliputi luas lahan dan jumlah pupuk subsidi yang diterima, yaitu pupuk urea, pupuk NPK, dan pupuk organik. Analisis *Clustering* dibatasi pada dua metode, yaitu *K-Means++* dan *Fuzzy C-Means*, serta evaluasi kualitas *Cluster* dilakukan menggunakan tiga indeks validitas internal, yaitu *Silhouette Score*, *Davies-Bouldin Index*, dan *Calinski-Harabasz Index*. Penelitian ini hanya berfokus pada proses pengelompokan data dan analisis kualitas *Cluster* tanpa membahas aspek implementasi kebijakan distribusi di lapangan ataupun faktor eksternal di luar variabel penelitian.

D. Rumusan Masalah

Berdasarkan latar belakang masalah maka permasalahan sebagai berikut : Bagaimana analisis klasterisasi penerima pupuk subsidi tahun 2025 menggunakan algoritma *K-Means++* dan *Fuzzy C-Means* dalam menentukan jumlah *cluster*

optimal berdasarkan indeks validitas internal sehingga dihasilkan *cluster* yang akurat dan dapat mendukung distribusi pupuk subsidi yang tepat sasaran?

E. Tujuan Penelitian

Secara khusus penelitian ini bertujuan untuk:

1. Menganalisis klasterisasi data penerima pupuk subsidi tahun 2025 menggunakan algoritma *K-Means++* dan *Fuzzy C-Means*.
2. Menentukan jumlah cluster optimal berdasarkan *Silhouette Score*, *Davies–Bouldin Index*, dan *Calinski–Harabasz Index*.
3. Menganalisis kinerja dan karakteristik cluster yang dihasilkan oleh masing-masing algoritma dalam pengelompokan penerima pupuk subsidi.

F. Manfaat Penelitian

Penelitian ini memberikan manfaat sebagai informasi pengelompokan penerima pupuk subsidi tahun 2025 yang dapat digunakan sebagai bahan pertimbangan dalam perencanaan distribusi pupuk subsidi yang lebih tepat sasaran. Serta, Menjadi referensi bagi penelitian selanjutnya yang berkaitan.

DAFTAR PUSTAKA

- Aldino, A. A., D. Darwis, A. T. Prastowo, and C. Sujana. 2021. "Implementation of K-Means Algorithm for Clustering Corn Planting Feasibility Area in South Lampung Regency." *Journal of Physics: Conference Series* 1751(1):012038. doi:10.1088/1742-6596/1751/1/012038.
- Arbelaitz, O., Gurrutxaga, I., Muguerza, J., Pérez, J. M., & Perona, I. (2013). *An extensive comparative study of cluster validity indices. Pattern Recognition*, 46(1), 243–256. <https://doi.org/10.1016/j.patcog.2012.07.021>
- Caliński, T., & Harabasz, J. (1974). *A dendrite method for cluster analysis. Communications in Statistics*, 3(1), 1–27. <https://doi.org/10.1080/03610927408827101>
- Daoudi, Sara, Chakib Mustapha Anouar Zouaoui, Miloud Chikr El-Mezouar, and Nasreddine Taleb. 2021. "Parallelization of the *K-Means++* Clustering Algorithm." *Ingenierie Des Systemes d'Information* 26(1):59–66. doi:10.18280/ISI.260106.
- Davies, David L., and Donald W. Bouldin. 1979. "A Cluster Separation Measure." *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-1(2):224–27. doi:10.1109/TPAMI.1979.4766909.
- Hashemi, Seyed Emadedin, Fatemeh Gholian-Jouybari, and Mostafa Hajiaghaei-Keshteli. 2023. "A Fuzzy C-Means Algorithm for Optimizing Data Clustering." *Expert Systems with Applications* 227:120377. doi:10.1016/J.ESWA.2023.120377.
- Hermantria, Irwan. 2021. "Segmentasi Konsumen Pupuk Non-Subsidi Di Provinsi Nusa Tenggara Timur." *BALANCE: Economic, Business, Management and Accounting Journal* 18(2):8–21. doi:10.30651/BLC.V18I2.7301.
- Ika Ari Sasmita, Rini Indriati, and M. Najibullo Muzaki. 2021. "Rekomendasi Penerima Bantuan Program Keluarga." *Jambura Journal of Electrical and Electronics Engineering* 3:84–88.
- Jain, A. K., M. N. Murty, and P. J. Flynn. 1999. "Data Clustering: A Review." *ACM Computing Surveys* 31(3):264–323. doi:10.1145/331499.331504.

- Kodinariya, Trupti m; Prashant Makwana. 2013. "Review on Determining of Cluster in K-Means Clustering." https://www.researchgate.net/publication/313554124_Review_on_Determining_of_Cluster_in_K-means_Clustering.
- Krasnov, D ;, D ; Davis, K ; Malott, Y ; Chen, X ; Shi, A Wong, Daniel Krasnov, Dresya Davis, Keiran Malott, Yiting Chen, Xiaoping Shi, and Augustine Wong. 2023. "Fuzzy C-Means Clustering: A Review of Applications in Breast Cancer Detection." *Entropy* 2023, Vol. 25, Page 1021 25(7):1021. doi:10.3390/E25071021.
- Kurnia, Ovelina Devi, Elisa Triammah, A. '. Fena, Dea Yuliana, Ayu Ningrum, Erna Daniati, and Aidina Ristyawan. 2024. "Analisis Perbandingan Algoritma Naïve Bayes Dengan K-Nearest Neighbor (KNN) Pada Dataset Mobile Price Classification." *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)* 8(2):1174–83. doi:10.29407/INOTEK.V8I2.5053.
- Meng, Jie, Ziyang Yu, Yuxin Cai, and Xiuling Wang. 2023. "K-Means++ Clustering Algorithm in Categorization of Glass Cultural Relics." *Applied Sciences* 2023, Vol. 13, Page 4736 13(8):4736. doi:10.3390/APP13084736.
- Mujahidin, Ilham, and Siti Hadijah Hasanah. 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):356–65. doi:10.14710/J.GAUSS.14.2.356-365.
- Nuraeni, Fitri, Muhammad Hasby, and Maulana Firdaus. 2025. "Model Klasterisasi Data Pertanian Menggunakan Algoritma K-Means." *Jurnal Algoritma* 22(2):890-898–890–898. doi:10.33364/ALGORITMA/V.22-2.2054.
- Oyewole, Gbeminiyi John and Thopil, George Alex. 2022. "Data Clustering: Application and Trends." *Artificial Intelligence Review* 56:6439--6475. <https://doi.org/10.1007/s10462-022-10325-y>.
- Prabowo Angga, Guruh Putro, and Dirgantoro. 2023. "Pengelompokan Pupuk Bersubsidi Di Kabupaten Bojonegoro Menggunakan Metode K -Means Clustering." *Multidisciplinary Applications of Quantum Information Science*

- (*Al-Mantiq*) 3(2). doi:10.32665/ALMANTIQ.V2I2.2006.
- Prakash, L. N. C. K., G. Suryanarayana, Mohd Dilshad Ansari, and Vinit Kumar Gunjan. 2022. "Instantaneous Approach for Evaluating the Initial Centers in the Agricultural Databases Using K-Means Clustering Algorithm." *Journal of Mobile Multimedia* 18(1):43–60. doi:10.13052/JMM1550-4646.1813.
- Ristyawan, Aidina, Arie Nugroho, and Theo Krisna Amarya. 2025. "Optimasi Preprocessing Model Random Forest Untuk Prediksi Stroke." *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)* 12(1). doi:10.35957/JATISI.V12I1.9587.
- Rousseeuw, P. J. (1987). *Silhouettes: A graphical aid to the interpretation and validation of cluster analysis*. *Journal of Computational and Applied Mathematics*, 20, 53–65.
- Saxena, Banshidhar Choudhary, Vipin. 2024. "Clustering of Agriculture Data through Fuzzy C-Means Technique." *African Journal of Biological Sciences (Afr. J. Bio. Sc.)* 6(13):3312–29. doi:10.48047/AFJBS.6.13.2024.
- Schwertman, Neil C., Margaret Ann Owens, and Robiah Adnan. 2004. "A Simple More General Boxplot Method for Identifying Outliers." *Computational Statistics & Data Analysis* 47(1):165–74. doi:10.1016/J.CSDA.2003.10.012.
- Sucipto, Sucipto, Teguh Andriyanto, Kuni Nadliroh, Antowi Bhima P, Dimas Indrajaya, and M. Ary Mustofa. 2019. "PELATIHAN PENGEMASAN DAN PENYULUHAN PEMBUATAN PUPUK KOMPOS: DESA PUS HARANG." *Jurnal Terapan Abdimas* 4(1):13. doi:10.25273/JTA.V4I1.3801.
- Susanto, Eri Sasmita, Novitasari Novitasari, Yudi Mulyanto, and Fahri Hamdani. 2024. "Clustering Petani Penerima Pupuk Berdasarkan Luas Lahan Dengan Menggunakan Algoritma K-Means Dan K-Medoids." *Digital Transformation Technology* 4(2):744–52. doi:10.47709/DIGITECH.V4I2.4264.
- Tamaela, Jemaictry, Eko Sedyono, Adi Setiawan, Magister Sisteminformasiuniversitas, and Kristen Satyawacana. 2017. "Cluster Analysis Menggunakan Algoritma Fuzzy C-Means Dan K-Means Untuk Klasterisasi Dan Pemetaan Lahan Pertanian Di Minahasa Tenggara." *Jurnal*

Buana Informatika 8(3). doi:10.24002/JBI.V8I3.1317.

- Watson, Hugh J., Daryl G. Greich, Colin Shearer, Stacey A. Herdlein, Joseph Fong, H. K. Wong, and Anthony Fong. 2000. "Statement of Purpose 2000 Best Practices and Leadership in Data Warehousing Awards Winners The CRISP-DM Model: The New Blueprint for Data Mining E-Business and the New Demands on Data Warehousing Technology: The New Demands E-Commerce Places on Data Ware...." *JOURNAL* 5. 5(4), 13–22 www.dw-institute.com/journal.htm.
- Wongoutong, Chantha. 2024. "The Impact of Neglecting Feature Scaling in K-Means Clustering." *PLOS ONE* 19(12):e0310839. doi:10.1371/JOURNAL.PONE.0310839.
- Xu, Rui, and Donald Wunsch. 2005. "Survey of Clustering Algorithms." *IEEE Transactions on Neural Networks* 16(3):645–78. doi:10.1109/TNN.2005.845141.
- Yahya, Muh. Zarkawi, Sitti Nurhaliza, Morina A Fathan, Muhammad Edy Rizal, and Andi Harismahyanti A. 2025. "Perbandingan Ukuran Jarak Pada Analisis Kluster Hirarki." *Leibniz: Jurnal Matematika* 5(02):93–111. doi:10.59632/leibniz.v5i02.538.
- Zalfa, Darina, Arie Nugroho, and Dwi Harini. 2023. "Analisa Clustering Pada Penerima Pupuk Subsidi Menggunakan Algoritma K-Means." *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)* 7(1):212–19. doi:10.29407/INOTEK.V7I1.3430.
- Zhang, LiuY and Wang, J. (2023). *A survey of fuzzy clustering validity evaluation methods.* *Information Sciences*, 618, 270–297. <https://doi.org/10.1016/j.ins.2022.11.010>