

# **PENDEKATAN CLUSTERING UNTUK ANALISIS KARAKTERISTIK PERTANIAN DESA DI KABUPATEN LEMBATA**

**HALIDA FIADNIN**



**PROGRAM SARJANA ILMU KOMPUTER  
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA  
INSTITUT PERTANIAN BOGOR  
BOGOR  
2026**



@Hak cipta milik IPB University

IPB University



IPB University  
— Bogor Indonesia —

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Perpustakaan IPB University

## PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Pendekatan *Clustering* untuk Analisis Karakteristik Pertanian Desa di Kabupaten Lembata” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Agustus 2026

Halida Fiadnin  
G6401211142



@Hak cipta milik IPB University

IPB University



IPB University  
— Bogor Indonesia —

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Perpustakaan IPB University

## ABSTRAK

HALIDA FIADNIN. Pendekatan *Clustering* untuk Analisis Karakteristik Pertanian Desa di Kabupaten Lembata. Dibimbing oleh SRI WAHJUNI dan HARI AGUNG ADRIANTO.

Pembangunan pedesaan memerlukan pemahaman karakteristik pertanian sebagai penghidupan utama masyarakat. Penelitian ini bertujuan mengidentifikasi pola kesamaan karakteristik pertanian desa di Kabupaten Lembata melalui *clustering* dan pemetaan spasial. Data survei dari 881 rumah tangga di 29 desa diolah menjadi delapan indikator tingkat desa. Melalui *feature selection*, terpilih tiga variabel paling berkontribusi, yaitu persentase pengairan irigasi, pembelian pupuk, dan pola tanam lebih dari sekali dalam setahun. Proses *clustering* menggunakan algoritma K-Means, *Agglomerative Hierarchical Clustering* (AHC), dan DBSCAN menghasilkan pengelompokan yang konsisten dengan hasil *Silhouette score* sebesar 0,8113. Hasilnya, 19 desa pada *cluster* besar (karakteristik serba rendah), 5 desa *cluster* sedang (pembelian pupuk tinggi, pola tanam rendah), 3 desa *cluster* kecil (irigasi dan pola tanam lebih baik), serta 2 desa sebagai *outlier* (Desa Bour dan Merdeka). Visualisasi QGIS menunjukkan kemiripan karakteristik pertanian desa tidak dipengaruhi kedekatan geografis. Ketersediaan irigasi menjadi kunci utama yang mendukung optimalisasi pola tanam di Kabupaten Lembata.

Kata kunci: *clustering*, Lembata, pemetaan spasial, pertanian desa.

## ABSTRACT

HALIDA FIADNIN. Clustering Approach for Analysis of Village Agricultural Characteristics in Lembata Regency. Supervised by SRI WAHJUNI and HARI AGUNG ADRIANTO.

Rural development requires an understanding of the characteristics of agriculture as the primary livelihood of the community. This study aims to identify patterns of similarity in village agricultural characteristics in Lembata Regency through clustering and spatial mapping. Survey data from 881 households in 29 villages were processed into eight village-level indicators. Through feature selection, three variables were selected with the highest contribution: the percentage of irrigation, fertilizer purchases, and cropping patterns more than once a year. The clustering process using the K-Means algorithm, Agglomerative Hierarchical Clustering (AHC), and DBSCAN produced consistent groupings with a Silhouette score of 0.8113. As a result, 19 villages were in the large cluster (low characteristics), 5 villages in the medium cluster (high fertilizer purchases, low cropping patterns), 3 villages in the small cluster (improved irrigation and cropping patterns), and 2 villages as outliers (Bour and Merdeka Villages). QGIS visualization shows that the similarities in village agricultural characteristics are not influenced by geographic proximity. The availability of irrigation is the main key that supports the optimization of cropping patterns in Lembata Regency.

*Keywords:* clustering, Lembata, spatial mapping, village agricultural.



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2026  
Hak Cipta dilindungi Undang-Undang

*Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.*

*Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.*

# **PENDEKATAN CLUSTERING UNTUK ANALISIS KARAKTERISTIK PERTANIAN DESA DI KABUPATEN LEMBATA**

**HALIDA FIADNIN**

Skripsi  
sebagai salah satu syarat untuk memperoleh gelar  
Sarjana pada  
Program Studi Ilmu Komputer

**PROGRAM SARJANA ILMU KOMPUTER  
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA  
INSTITUT PERTANIAN BOGOR  
BOGOR  
2026**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.





Judul Skripsi : Pendekatan *Clustering* untuk Analisis Karakteristik Pertanian  
Desa di Kabupaten Lembata

Nama : Halida Fiadnin

NIM : G6401211142

Disetujui oleh

Pembimbing 1:

Dr. Ir. Sri Wahjuni, M.T.

---

Pembimbing 2:

Hari Agung Adrianto, S.Kom., M.Si., Ph.D.

---

Diketahui oleh

Ketua Program Sarjana Ilmu Komputer:

Dr. Sony Hartono Wijaya, S.Kom., M.Kom

19810809 200812 1 002

---

Tanggal Ujian:

7 Juli 2026

Tanggal Lulus:

## PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Judul yang dipilih dalam penelitian yang dilaksanakan sejak bulan Januari 2025 sampai bulan Agustus 2026 ini ialah “Pendekatan *Clustering* untuk Analisis Karakteristik Pertanian Desa di Kabupaten Lembata”.

Terima kasih penulis ucapkan sebesar-besarnya kepada

1. Ibu Dr. Ir. Sri Wahjuni, M.T. dan Bapak Hari Agung Adrianto, S.Kom., M.Si., Ph.D. selaku dosen pembimbing yang telah meluangkan waktu, tenaga, masukan, serta memberikan dukungan yang luar biasa kepada penulis.
2. Bapak Endang Purnama Giri, S.Kom., M.Kom. selaku moderator seminar hasil, dan Bapak Muhammad Asyhar Agmalaro, S.Si., M.Kom. selaku dosen penguji sidang yang memberikan kontribusi penting dalam penyelesaian tugas akhir ini.
3. Seluruh dosen dan sivitas akademika Ilmu Komputer IPB yang telah memberikan ilmu dan pengalaman selama masa studi.
4. Ayah Zubaidy, Ibu Mukti, Mbah Wardiwati, Kak Hurin, Hibban, dan Dedek Zia serta seluruh keluarga yang senantiasa memberikan dukungan, doa, dan kasih sayang yang tiada henti.
5. Bapak Ferry Lasar selaku mentor kegiatan magang MSIB yang telah membimbing penulis dan rekan-rekan selama kegiatan magang, membantu proses pengumpulan data, serta mengenalkan budaya dan nilai-nilai masyarakat setempat.
6. Mama Mia, Bapak Edi, Daling, Iska, Ian, serta seluruh keluarga di Waienga, Lembata yang telah menerima kehadiran penulis selama tiga bulan, memberikan pengalaman menjelajahi Pulau Lembata yang begitu indah, serta mengajarkan penulis banyak makna kehidupan.
7. Sahabat-sahabat penulis di “Puncak Komedi” yaitu Ismy Fana Fillah, Viragita Athaya Haura, Tita Madriyanti, dan Jihan Febriharvianti Wirawan yang menjadi teman penulis sejak hari pertama menginjakkan kaki di IPB sebagai teman belajar, bermain, bersantai, berkuliner, juga sebagai teman yang selalu hadir baik di masa-masa sulit maupun senang, terima kasih atas waktu dan kebersamaannya.
8. Teman-teman dekat penulis, Sarah, Rahmah, Faiza, Hani, Nurul, Rina, Jundi, Alfian, Taqi, yang telah menemani penulis sejak masa sekolah menengah hingga kini. Tak lupa Illinia sebagai teman seperjuangan selama studi independen dan penulisan tugas akhir. Terima kasih telah memberikan banyak *core memory* bagi penulis.
9. Rekan-rekan magang di Lembata, Firdaus, Ais, Zabrina, Sabil, Puput, David, dan Indra atas kerja sama, dukungan, dan kebersamaannya selama berada di tempat perantauan.
10. Seluruh teman-teman Ilkomerz 58, rekan divisi Edukasi Himalkom 2023, rekan BPH Himalkom 2024, rekan kepanitiaan IT Today, CPSC, serta teman-teman lainnya yang tidak dapat disebutkan satu per satu atas dukungan, kerja sama, dan kebersamaannya selama masa perkuliahan.

Penulis juga menyadari bahwa dalam proses penyusunan tugas akhir ini, masih banyak kekurangan dan keterbatasan baik dalam hal penyusunan materi maupun penyampaian informasi. Oleh karena itu, penulis memohon maaf atas segala kekurangan atau kekeliruan yang mungkin ada. Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Agustus 2026

*Halida Fiadnin*

@Hak cipta milik IPB University

IPB University



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



## DAFTAR ISI

|                                                                                 |    |
|---------------------------------------------------------------------------------|----|
| DAFTAR TABEL                                                                    | ix |
| DAFTAR GAMBAR                                                                   | ix |
| PENDAHULUAN                                                                     | 1  |
| 1.1 Latar Belakang                                                              | 1  |
| 1.2 Rumusan Masalah                                                             | 2  |
| 1.3 Tujuan                                                                      | 2  |
| 1.4 Manfaat                                                                     | 2  |
| 1.5 Ruang Lingkup                                                               | 2  |
| TINJAUAN PUSTAKA                                                                | 4  |
| 2.1 <i>Clustering</i>                                                           | 4  |
| 2.2 K-Means                                                                     | 6  |
| 2.3 <i>Agglomerative Hierarchical Clustering</i>                                | 7  |
| 2.4 DBSCAN                                                                      | 7  |
| III METODE                                                                      | 11 |
| 3.1 Tahapan Penelitian                                                          | 11 |
| 3.2 Pengumpulan Data                                                            | 11 |
| 3.3 Persiapan Data                                                              | 12 |
| 3.4 <i>Data Aggregation</i>                                                     | 13 |
| 3.5 <i>Feature Selection</i>                                                    | 14 |
| 3.6 <i>Modelling</i>                                                            | 16 |
| 3.7 Evaluasi Hasil <i>Clustering</i>                                            | 22 |
| 3.8 Visualisasi                                                                 | 23 |
| IV HASIL DAN PEMBAHASAN                                                         | 24 |
| 4.1 <i>Feature Selection</i>                                                    | 24 |
| 4.2 <i>Modelling</i>                                                            | 25 |
| 4.3 Interpretasi Kesamaan Hasil <i>Clustering</i> pada K-Means, AHC, dan DBSCAN | 33 |
| 4.4 Visualisasi dan Analisis Profil Desa                                        | 34 |
| V SIMPULAN DAN SARAN                                                            | 42 |
| 5.1 Simpulan                                                                    | 42 |
| 5.2 Saran                                                                       | 42 |
| DAFTAR PUSTAKA                                                                  | 43 |
| RIWAYAT HIDUP                                                                   | 45 |



## DAFTAR TABEL

|    |                                                                                                                                               |    |
|----|-----------------------------------------------------------------------------------------------------------------------------------------------|----|
| 1  | Tahapan yang dilakukan pada <i>data transformation</i>                                                                                        | 12 |
| 2  | Parameter yang digunakan pada metode <i>grid-search</i> DBSCAN                                                                                | 21 |
| 3  | Nilai rata-rata variabel <i>k</i> pada <i>cluster t</i> dalam skala yang sudah terstandarisasi (ctk) dan jumlah anggota <i>cluster t</i> (St) | 24 |
| 4  | Nilai akhir <i>contribution weight</i> masing-masing variabel dan persenan kontribusinya                                                      | 25 |
| 5  | Statistik setiap <i>cluster</i> untuk algoritma K-Means dalam skala asli                                                                      | 26 |
| 6  | Statistik setiap <i>cluster</i> untuk algoritma AHC dalam skala asli                                                                          | 29 |
| 7  | Hasil metode <i>grid-search</i> pada algoritma DBSCAN                                                                                         | 30 |
| 8  | Statistik setiap <i>cluster</i> untuk algoritma DBSCAN dalam skala asli                                                                       | 32 |
| 9  | Rangkuman hasil <i>clustering</i> dan nilai evaluasi untuk setiap algoritma                                                                   | 32 |
| 10 | Perbandingan hasil <i>clustering</i> jika <i>feature selection</i> dilakukan pada tiga dan empat variabel                                     | 33 |

## DAFTAR GAMBAR

|    |                                                                                                                                                                                                       |    |
|----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 1  | Kategorisasi dan contoh algoritma <i>clustering</i> (Mostafa 2020)                                                                                                                                    | 4  |
| 2  | Metode pemotongan pada algoritma <i>hierarchical-based</i> dapat dilakukan pada tingkat jumlah <i>cluster</i> yang diinginkan (Sisodia <i>et al.</i> 2012)                                            | 5  |
| 3  | Tahapan algoritma K-Means (Berkhin 2006)                                                                                                                                                              | 6  |
| 4  | Perbandingan nilai <i>epsilon</i> yang semakin besar (dan menjadi terlalu besar) menyebabkan <i>noise</i> masuk ke dalam <i>cluster</i> (Fahim 2023)                                                  | 8  |
| 5  | Konsep dasar DBSCAN: (a) dan (b) pendefinisian <i>directly density-reachable</i> , (c) pendefinisian <i>density-reachable</i> , (d) pendefinisian <i>density-connected</i> (Ester <i>et al.</i> 1996) | 9  |
| 6  | Contoh penentuan <i>core point</i> , <i>border point</i> , dan <i>noise point</i> pada algoritma DBSCAN (Birant dan Kut 2006)                                                                         | 9  |
| 7  | Tahapan penelitian                                                                                                                                                                                    | 11 |
| 8  | Tahapan <i>feature selection</i>                                                                                                                                                                      | 15 |
| 9  | Tahapan algoritma K-Means                                                                                                                                                                             | 17 |
| 10 | Contoh grafik untuk memilih nilai <i>k</i> terbaik menggunakan metode <i>Silhouette</i> (Saputra <i>et al.</i> 2020)                                                                                  | 18 |
| 11 | Tahapan algoritma AHC                                                                                                                                                                                 | 19 |
| 12 | Tahapan algoritma DBSCAN                                                                                                                                                                              | 20 |
| 13 | Contoh hasil pencarian nilai MinPts menggunakan <i>grid-search</i> pada nilai <i>eps</i> = 0,357                                                                                                      | 22 |
| 14 | Hasil analisis <i>Silhouette score</i> untuk tahapan <i>feature selection</i>                                                                                                                         | 24 |
| 15 | Hasil metode <i>Silhouette</i> pada algoritma K-Means                                                                                                                                                 | 26 |
| 16 | Dendrogram hasil algoritma AHC                                                                                                                                                                        | 27 |
| 17 | Hasil <i>Silhouette score</i> algoritma AHC                                                                                                                                                           | 28 |
| 18 | Perbedaan <i>t</i> = 3 dan <i>t</i> = 5 pada <i>clustering</i> algoritma AHC dalam bentuk visualisasi tiga dimensi                                                                                    | 28 |
| 19 | Hasil metode K-Neighbors pada algoritma DBSCAN                                                                                                                                                        | 30 |

|    |                                                                                                                              |    |
|----|------------------------------------------------------------------------------------------------------------------------------|----|
| 20 | Perbedaan MinPts = 4 dan MinPts = 3 pada nilai $eps = 0,229$ algoritma DBSCAN dalam bentuk visualisasi tiga dimensi          | 31 |
| 21 | Visualisasi tiga dimensi dan pemetaan QGIS untuk seluruh <i>cluster</i>                                                      | 35 |
| 22 | Analisis curah hujan dasarian III per Februari 2025 (Dinas Pertanian dan Ketahanan Pangan Provinsi Nusa Tenggara Timur 2025) | 35 |
| 23 | Pemetaan untuk <i>cluster</i> besar                                                                                          | 36 |
| 24 | Pemetaan untuk <i>cluster</i> sedang                                                                                         | 37 |
| 25 | Analisis curah hujan Lembata dan hasil <i>clustering</i>                                                                     | 38 |
| 26 | Pemetaan untuk <i>cluster</i> kecil                                                                                          | 39 |
| 27 | Pemetaan untuk <i>outlier</i>                                                                                                | 41 |