

SKRIPSI

**PENERAPAN DATA MINING UNTUK CLUSTERING WILAYAH
PRODUKSI KAKAO MENGGUNAKAN METODE K-MEANS
(STUDI KASUS: KABUPATEN SIKKA, NUSA TENGGARA TIMUR)**



Disusun Oleh:

MADELBERTHA FRIDOLIN PUKA

21.18.132

**PROGRAM STUDI TEKNIK INFORMATIKA S-1
FAKULTAS TEKNOLOGI INDUSTRI
INSTITUT TEKNOLOGI NASIONAL MALANG**

2025

LEMBAR PERSETUJUAN

**PENERAPAN DATA MINING UNTUK CLUSTERING WILAYAH
PRODUKSI KAKAO MENGGUNAKAN METODE K-MEANS
(STUDI KASUS: KABUPATEN SIKKA, NUSA TENGGARA TIMUR)**

SKRIPSI

*Disusun dan Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar
Sarjana Komputer Strata Satu (S-1)*

Disusun Oleh :

Madelbertha Fridolin Puka

21.18.132

Diperiksa dan Disetujui,

Dosen Pembimbing I

Dosen Pembimbing II

(Joseph Dedy Irawan, ST., MT)

NIP. 197404162005 011002

(Nurlaiti Vendvansyah, ST., MT)

NIP. 1031900557

Mengetahui,

Program Studi Teknik Informatika S-1

Ketua

Yosep Agus Pranoto ST., MT

NIP.P 1031000432

**PROGRAM STUDI TEKNIK INFORMATIKA S-1
FAKULTAS TEKNOLOGI INDUSTRI
INSTITUT TEKNOLOGI NASIONAL MALANG**

2025

LEMBAR PERNYATAAN KEASLIAN SKRIPSI

Sebagai mahasiswa Program Studi Teknik Informatika S-1 Fakultas Teknologi Industri Institut Teknologi Nasional Malang, yang bertanda tangan di bawah ini, saya:

Nama : Madelbertha Fridolin Puka

NIM : 2118132

Program Studi : Teknik Informatika S-1

Fakultas : Teknologi Industri

Menyatakan dengan sesungguhnya bahwa skripsi saya dengan judul “Penerapan Data Mining Untuk Clustering Wilayah Produksi Kakao Menggunakan Metode K-Means (Studi Kasus: Kabupaten Sikka, Nusa Tenggara Timur)” merupakan karya asli dan bukan merupakan duplikat dan mengutip seluruhnya karya orang lain. Apabila di kemudian hari, karya asli saya disinyalir bukan merupakan karya asli saya, maka saya akan bersedia menerima segala konsekuensi apapun yang diberikan Program Studi Teknik Informatika S-1 Fakultas Teknologi Industri Institut Teknologi Nasional Malang.

Demikian surat pernyataan ini saya buat dengan sebenar-benarnya.

Malang, Januari 2025

Yang membuat pernyataan,



(Madelbertha Fridolin Puka)

NIM 2118132

**PENERAPAN DATA MINING UNTUK CLUSTERING WILAYAH
PRODUKSI KAKAO MENGGUNAKAN METODE K-MEANS
(STUDI KASUS: KABUPATEN SIKKA, NUSA TENGGARA TIMUR)**

Madelbertha Fridolin Puka, Joseph Dedy Irawan, Nurlaily Vendyansyah

Teknik Informatika, Institut Teknologi Nasional Malang

Jalan Raya Karanglo km 2 Malang, Indonesia

madelberthafridolinpuka@gmail.com

ABSTRAK

Dengan total produksi 7.993,17 ton pada tahun 2022, Kabupaten Sikka menduduki peringkat ketiga untuk produksi perkebunan dan merupakan salah satu wilayah penghasil kakao di Provinsi Nusa Tenggara Timur. Namun, menurut sebuah studi pada tahun 2023 oleh Pusat Kajian Dinamika Sistem Pembangunan (PKDSP), potensi kakao di Kabupaten Sikka belum berkembang secara maksimal karena tidak dapat menentukan potensi masing-masing daerah. Tujuan dari penelitian yang dilakukan ini adalah untuk mengidentifikasi dan mengklasifikasikan daerah penghasil kakao di Kabupaten Sikka dengan menerapkan teknik pada data mining dan algoritma K-Means clustering. Badan Pusat Statistik (BPS) Kabupaten Sikka menyediakan data yang meliputi informasi luas lahan perkebunan kakao di 21 kecamatan dan statistik produksi. Data yang digunakan adalah data luas lahan perkebunan kakao di 21 kecamatan di Kabupaten Sikka.

Kata Kunci : *Data mining, Clustering, Produksi Kakao, K-Means, Kabupaten Sikka*

KATA PENGANTAR

Dengan memanjatkan puji syukur kehadirat Tuhan Yang Maha Esa atas limpahan rahmat dan Karunia-Nya, sehingga penulis dapat menyelesaikan laporan skripsi dengan judul **“Penerapan Data Mining Untuk Clustering Wilayah Produksi Kakao Menggunakan Metode K-Means (Studi Kasus: Kabupaten Sikka, Nusa Tenggara Timur)”** dan dapat diselesaikan dengan baik. Laporan skripsi ini disusun sebagai salah satu syarat untuk mengerjakan skripsi pada program Strata-1 di Jurusan Teknik Informatika Institut Teknologi Nasional Malang.

Penulis menyadari banyak pihak yang memberikann dukungan dan bantuan selama menyelesaikan studi dan tugas akhir ini. Oleh karena itu, sudah sepantasnya penulis dengan penuh hormat mengucapkan terimakasih yang sangat besar kepada:

1. Tuhan Yesus Kristus yang telah memberikan kasih setia, rahmat dan karunia-Nya bagi penyusun sehingga dapat mengerjakan skripsi dengan lancar.
2. Kedua Orang Tua dan Keluarga Besar yang telah memberikan doa dan berbagai macam dorongan dalam penyelesaian skripsi ini.
3. Bapak Yosep Agus Pranoto, S.T., MT, selaku Ketua Program Studi Teknik Informatika S-1 Institut Teknologi Nasional Malang.
4. Bapak Joseph Dedy Irawan, ST.,MT, selaku Dosen Pembimbing I yang telah membimbing dalam penyusunan skripsi.
5. Ibu Nurlaily Vendyansyah, ST.,MT, selaku Dosen Pembimbing II yang telah membimbing dalam penyusunan skripsi.
6. Segenap dosen Program Studi Teknik Informatika Institut Teknologi Nasional Malang yang selalu memberikan dukungan kepada penyusun.
7. Semua teman-teman Keluarga Mahasiswa Katolik institut Teknologi Nasional Malang yang telah memberikan semangat dan dorongan dalam pelaksanaan dan penyusunan skripsi ini.
8. Semua rekan-rekan kerja skripsi yang selalu menjadi penyemangat serta motivasi dalam mengerjakan skripsi.
9. Semua teman-teman Kos Sadewa yang telah membantu dalam pelaksanaan dan penyusunan skripsi ini.

10. Para mahasiswa Informatika yang telah memberikan dukungan, evaluasi dan saran yang sangat berharga bagi penyusun dalam menyelesaikan skripsi ini.
11. Diri sendiri karena sudah bertahan dan mampu melewati segala kesulitan dalam penyusunan skripsi ini sehingga dapat menyelesaikannya dengan baik.

Dengan segala kerendahan hati, penulis menyadari masih banyak kekurangan dalam penyusunan skripsi ini. Oleh karena itu, penulis meminta maaf yang sedalam-dalamnya atas kesalahan yang dilakukan penulis. Penulis berharap semoga skripsi ini dapat bermanfaat bagi pembaca dan dapat dijadikan referensi demi pengembangan ke arah yang lebih baik.

Malang, Januari 2025

Penulis

DAFTAR ISI

DAFTAR ISI	vii
DAFTAR GAMBAR	viii
DAFTAR TABEL	x
BAB I PENDAHULUAN	1
1.2 Latar Belakang	1
1.3 Rumusan Masalah	3
1.4 Tujuan	3
1.5 Batasan Masalah	3
1.6 Manfaat	4
1.7 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	6
2.2 Penelitian Terkait	6
2.3 Dasar Teori	7
BAB III ANALISIS DAN PERANCANGAN	20
3.1 Analisis Kebutuhan	20
3.2 Diagram Blok Sistem	21
3.3 Use Case Diagram	21
3.4 Struktur Menu	22
3.5 Flowchart Sistem	23
3.6 Flowchart K-Means Clustering	24
3.7 Desain Prototype	25
BAB IV IMPLEMENTASI DAN PENGUJIAN	30
4.1 Implementasi	30
4.2 Implementasi Metode K-Means Clustering	34
4.3 Pengujian Sistem	39
4.4 Evaluasi Metode K-Means Clustering Menggunakan Silhoutte Score	42
4.5 Evaluasi Korespondensi Data Manual dengan Hasil Sistematis	44
4.6 Evaluasi Keselarasan Keluaran Sistem dan Penilaian Pakar	44
BAB V PENUTUP	45
5.2 Kesimpulan	45
5.3 Saran	45
DAFTAR PUSTAKA	46

DAFTAR GAMBAR

Gambar 1.1 Produksi Komoditi Perkebunan Kab. Sikka.....	1
Gambar 2.1 Flowchart Algoritma K-Means	11
Gambar 2.2 Kakao	14
Gambar 2.3 Peta Administrasi Kabupaten Sikka	15
Gambar 3.1 Diagram Blok	21
Gambar 3.2 <i>Use Case</i> Diagram	22
Gambar 3.3 Struktur Menu.....	22
Gambar 3.4 Flowchart Sistem	23
Gambar 3.5 Flowchart K-Means Clustering	24
Gambar 3.6 Menu Home.....	25
Gambar 3.7 Menu Login.....	25
Gambar 3.8 Menu Register	25
Gambar 3.9 Menu Dashboard.....	26
Gambar 3.10 Menu Data Kecamatan.....	26
Gambar 3.11 Menu Tambah Data Kecamatan	26
Gambar 3.12 Edit Data Kecamatan	27
Gambar 3.13 Menu Data Luas Lahan	27
Gambar 3.14 Tambah Data Luas Lahan	27
Gambar 3.15 Edit Data Luas Lahan.....	28
Gambar 3.16 Menu Data Produksi	28
Gambar 3.17 Menu Tambah Data Produksi.....	28
Gambar 3.18 Menu Hasil Data Produksi	29
Gambar 3.19 Menu Clustering Data	29
Gambar 4.1 Halaman Home.....	30
Gambar 4.2 Halaman Login	30
Gambar 4.3 Halaman Register	31
Gambar 4.4 Halaman Dashboard.....	31
Gambar 4.5 Halaman Data Kecamatan.....	31
Gambar 4.6 Halaman Luas Lahan	32
Gambar 4.7 Halaman Data Produksi	32
Gambar 4.8 Halaman Clustering	33
Gambar 4.9 Grafik Elbow	33

Gambar 4.10 Grafik SSE.....	36
-----------------------------	----

DAFTAR TABEL

Tabel 4.1 Data Kecamatan	34
Tabel 4.2 Data Sampel Perhitungan	34
Tabel 4.3 Pencarian Jarak $K=2$	34
Tabel 4.4 Tabel SSE $K=2$	35
Tabel 4.5 Selisih Nilai K	35
Tabel 4.6 Cluster.....	36
Tabel 4.7 Centroid Awal	36
Tabel 4.8 Iterasi Ke-1	37
Tabel 4.9 Centroid Baru	37
Tabel 4.10 Hasil Clustering.....	38
Tabel 4.11 Pengujian Fungsional	39
Tabel 4.12 Pengujian Blackbox.....	40
Tabel 4.13 Cluster 1	42
Tabel 4.14 Cluster 2.....	42
Tabel 4.15 Cluster 3.....	42
Tabel 4.16 Menghitung Jarak Antar Cluster ($b(i)$)	43
Tabel 4.17 Menghitung Silhoutte	43
Tabel 4.18 Perbandingan hasil data manual dengan sistem	44
Tabel 4.19 Perbandingan hasil sistem dengan hasil pakar	44