

**PENERAPAN *K-MEANS CLUSTERING* GUNA IDENTIFIKASI
PENDUDUK MISKIN EKSTREM YANG BELUM TERAKSES
PROGRAM PERLINDUNGAN SOSIAL**

SKRIPSI

**Untuk Memenuhi Persyaratan
Memperoleh Gelar Sarjana**

Oleh:

**Muhammad Ikhsan
NIM 24.51.0010**



PROGRAM STUDI S-1 SISTEM INFORMASI

**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
PPKIA PRADNYA PARAMITA
MALANG
2025**

LEMBAR PERSETUJUAN

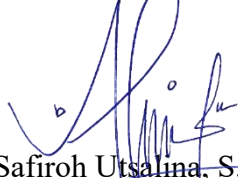
PENERAPAN K-MEANS CLUSTERING GUNA IDENTIFIKASI PENDUDUK MISKIN EKSTREM YANG BELUM TERAKSES PROGRAM PERLINDUNGAN SOSIAL

Oleh:
Muhammad Ikhsan
NIM 24.51.0010

**Telah disetujui oleh pembimbing untuk diujikan
Pada tanggal 21 Juli 2025**

Menyetujui,
Komisi Pembimbing

Pembimbing I



Dr. Dwi Safiroh Utsalina, S.Kom, MMSI
NIK 00.12.01.106

Pembimbing II



Dr. Tb. Mohammad Akhriza, S.Si., MMSI., Ph.D
NIK 00.12.01.002

Mengetahui,
Ketua Program Studi S-1 Sistem Informasi



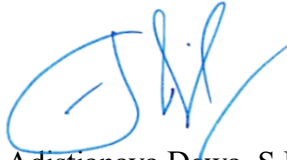
Dr. Dwi Safiroh Utsalina, S.Kom, MMSI
NIK 00.12.01.106

LEMBAR PENGESAHAN

PENERAPAN K-MEANS CLUSTERING GUNA IDENTIFIKASI PENDUDUK MISKIN EKSTREM YANG BELUM TERAKSES PROGRAM PERLINDUNGAN SOSIAL

**Skripsi oleh Nama Mahasiswa ini
Telah dipertahankan di depan dewan penguji
Pada tanggal 04 Agustus 2025**

Dewan Penguji,
Ketua



Dr. Weda Adistianaya Dewa, S.Kom., MMSI
NIK 14.02.25.003

Anggota I



Dr. Mochamad Husni, S.Pd.I, M.M., MMSI
NIK 01.09.10.036

Anggota II



Dr. Tb. Mohammad Akhriza, S.Si., MMSI., Ph.D
NIK 00.12.01.002

Mengetahui,
Ketua Program Studi
S-1 Sistem Informasi



Dr. Dwi Santia Utsalina, S.Kom, MMSI
NIK 00.12.01.106

Mengesahkan,
Ketua STMIK PPKIA
Pradnya Paramita



Dr. Tb. Mohammad Akhriza, S.Si., MMSI., Ph.D
NIK 00.12.01.002

PERNYATAAN ORISINALITAS SKRIPSI

Saya menyatakan dengan sebenar-benarnya bahwa sepanjang pengetahuan saya, di dalam Naskah SKRIPSI ini tidak terdapat karya ilmiah yang pernah diajukan oleh orang lain untuk memperoleh gelar akademik di suatu Perguruan Tinggi, dan tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip dalam naskah ini disebutkan dalam sumber kutipan dan daftar pustaka.

Apabila ternyata di dalam naskah SKRIPSI ini dapat dibuktikan terdapat unsur-unsur PLAGIASI, saya bersedia SKRIPSI ini digugurkan dan gelar akademik yang saya peroleh (SARJANA) dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku (UU No 20 Tahun 2003, pasal 25 ayat 2 dan pasal 70).

Malang, 22 Juli 2025

Mahasiswa,

A handwritten signature in black ink is written over a yellow and red revenue stamp. The stamp features the Garuda Pancasila emblem and the text 'REPUBLIK INDONESIA', '10000', 'METAL TEMPEL', and the serial number 'D526EAM292395167'.

Nama : Muhammad Ikhsan

NIM : 24.51.0010

‘MOTTO’

“Semua Berasal dari Allah dan Akan Kembali Kepada-Nya”

“Ingatlah Allah di Waktu Lapangmu

Niscaya Allah Mengingatmu di Waktu Sempitmu”

ABSTRAK

Ikhsan, Muhammad. 2025. *Penerapan K-Means Clustering Guna Identifikasi Penduduk Miskin Ekstrem Yang Belum Terakses Program Perlindungan Sosial*. Program Studi Sistem Informasi STMIK PPKIA Pradnya Paramita. Pembimbing: (I) Dr. Dwi Safiroh Utsalina, S.Kom, MMSI, (II) Dr. Tb. Mohammad Akhriza, S.Si., MMSI., Ph.D

Kata Kunci: *K-Means, Silhouette Coefficient, WCSS, Clustering, Segmentasi*

Kemiskinan merupakan masalah utama dalam menghambat pencapaian kesejahteraan sosial masyarakat. Mengacu pada data Badan Pusat Statistik Kabupaten Malang, tahun 2023 jumlah penduduk miskin mencapai 251,36 ribu. Penghapusan kemiskinan ekstrem merupakan salah satu tujuan strategis di Daerah di mana intervensinya memanfaatkan data Pensasaran Percepatan Penghapusan Kemiskinan Ekstrem (P3KE) yang dirilis oleh MenkoPMK dan dilakukan pemutakhiran setiap tahun. Belum adanya model segmentasi data pada sasaran penerima program perlindungan sosial rentan menimbulkan duplikasi penerima bantuan sosial, penduduk yang tidak layak menerima bantuan (*inclusion error*), penduduk yang belum teridentifikasi sebagai sasaran (*exclusion error*), serta data yang tidak valid dalam penyaluran bantuan sosial. Segmentasi penduduk miskin menggunakan metode K-Means Clustering dengan kerangka kerja CRISP-DM. Adapun atribut data ditetapkan yaitu jenis pekerjaan, kepemilikan rumah, kepemilikan Program Keluarga Harapan (PKH), dan jumlah anggota keluarga beresiko stunting. Hasil penerapan metode K-Means pada 7857 record dengan penentuan jumlah k optimal sebesar k=5 setelah dilakukan uji metode *Silhouette Coefficient* dan WCSS dengan k=2 sampai dengan k=15 menghasilkan 5 kluster. Clustering penduduk miskin ekstrem menghasilkan 5 segmen prioritas penduduk miskin yaitu Sangat Prioritas sebanyak 2332 KK, Prioritas Tinggi sebanyak 1353 KK, Prioritas Sedang sebanyak 1545 KK, Kurang Prioritas sebanyak 284 KK, dan Bukan Prioritas sebanyak 2343 KK. Selanjutnya segmentasi penduduk miskin ekstrem dapat digunakan sebagai bahan penyusunan kebijakan dalam menyusun prioritas penerima program perlindungan sosial.

KATA PENGANTAR

Puji syukur kami panjatkan kehadirat Tuhan Yang Maha Esa atas segala rahmat dan hidayah-Nya penulis dapat menyelesaikan skripsi dengan baik sebagai syarat untuk memperoleh gelar Sarjana Komputer..

Kami menyadari tanpa adanya dukungan dan kerja sama dari berbagai pihak, kegiatan skripsi ini tidak akan dapat berjalan baik. Untuk itu, kami ingin menyampaikan rasa terima kasih kepada:

1. Tuhan Yang Maha Esa yang telah memberikan petunjuk dan hidayah dalam pembuatan skripsi dan penyusunan laporan sehingga dapat berjalan dengan baik dari awal hingga akhir.
2. Kedua orang tua kami yang telah memberikan doa dan dukungannya.
3. Bapak Dr. Tb. Mohammad Akhriza, S.Si, M.M.S.I, selaku Ketua STMIK PPKIA Pradnya Paramita.
4. Ibu Dr. Dwi Safiroh Utsalina, S.Kom, MMSI., selaku Ketua Program Studi S-1 Sistem Informasi.
5. Ibu Dr. Dwi Safiroh Utsalina, S.Kom, MMSI dan Bapak Dr. Tb. Mohammad Akhriza, S.Si., MMSI., Ph.D, selaku dosen pembimbing skripsi.
6. Bapak Dr. Weda Adistianaya Dewa, S.Kom., MMSI, dan Bapak Dr. Mochamad Husni, S.Pd.I, M.M., MMSI, selaku dosen penguji skripsi
7. Dan seluruh pihak yang telah membantu dan mendukung lancarnya pembuatan Skripsi dari awal hingga akhir yang tidak dapat kami sebutkan satu persatu.

Penulis mengharapkan saran dan kritik dari berbagai pihak yang bersifat membangun demi penyempurnaan skripsi ini. Semoga skripsi ini berguna bagi pembaca secara umum dan penulis secara khusus. Akhir kata, penulis ucapkan banyak terima kasih.

Malang, 21 Juli 2025



Nama : Muhammad Ikhsan

NIM : 24.51.0010

DAFTAR ISI

LEMBAR PERSETUJUAN	i
LEMBAR PENGESAHAN	ii
PERNYATAAN ORISINALITAS SKRIPSI.....	iii
ABSTRAK	v
KATA PENGANTAR.....	vi
DAFTAR ISI.....	vii
DAFTAR TABEL	ix
DAFTAR GAMBAR.....	x
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	5
1.3. Tujuan Penelitian	5
1.4. Batasan Masalah	5
1.5. Kontribusi Penelitian	6
BAB II TINJAUAN PUSTAKA.....	7
2.1. Kajian Terkait	7
2.2. Kemiskinan Ekstrem.....	8
2.3. Data Mining	9
2.4. <i>Clustering</i>	10
2.5. CRISP-DM.....	16
2.6. Metode K-Means	19
2.7. KNIME Analytics Platform	21
BAB III METODOLOGI PENELITIAN	23
3.1. Analisis Permasalahan	23
3.2. Solusi yang Diusulkan	24
3.3. Pengumpulan Dataset.....	24
3.4. Business Understanding.....	25
3.5. Data Understanding	25
3.6. Data Preparation.....	25

3.7.	Modelling	28
3.8.	Evaluation	29
3.9.	Deployment.....	30
3.10.	Rancangan Eksperimen.....	30
3.10.1.	Kebutuhan Perangkat.....	30
3.10.2.	Kebutuhan Data	31
BAB IV PENGUJIAN DAN HASIL.....		32
4.1.	Business Understanding.....	32
4.2.	Data Understanding	32
4.3.	Data Preparation.....	36
4.4.	Modelling	43
4.5.	Evaluation	48
4.6.	Deployment.....	68
BAB V KESIMPULAN DAN SARAN		80
5.1.	Kesimpulan	80
5.2.	Saran	80
DAFTAR PUSTAKA		

DAFTAR TABEL

No Nama Tabel	Halaman
4.1 Sampel Data Mentah	32
4.2 Penjelasan Atribut Data.....	34
4.3 Sampel Data Mentah	36
4.4 Hasil Seleksi Data Mentah	38
4.5 Kriteria Atribut Jenis Pekerjaan	39
4.6 Kriteria Atribut Kepemilikan Rumah.....	39
4.7 Kriteria Atribut Kepemilikan PKH	40
4.8 Dataset setelah dilakukan transformasi	40
4.9 Hasil Rekapitulasi Nilai Silhouette Coefficient	57
4.10 Hasil perhitungan WCSS	65
4.11 Hasil <i>Clustering</i> menggunakan K-Means	74

DAFTAR GAMBAR

No Nama Gambar	Halaman
2.1 Siklus CRISP-DM.....	16
3.1 Tahapan Penelitian.....	24
3.3 Sampel Data Mentah.....	31
4.6 Tangkapan layar pengaturan node Excel Reader.....	41
4.7 Node Excel Reader yang sudah dikonfigurasi.....	41
4.8 Tangkapan layar pengaturan node Normalizer.....	42
4.9 Tangkapan Layar node Normalizer.....	43
4.10 Tampilan Data hasil normalisasi.....	43
4.11 Instalasi Node K-Means.....	44
4.12 Pengaturan node.....	44
4.13 Hasil Klasterisasi Dengan Node K-Means.....	45
4.14 Nilai Centroid dari Masing-Masing Klaster.....	45
4.15 Node Denormalizer.....	46
4.16 Hasil <i>Denormalizer</i>	46
4.17 Node Color Manager.....	47
4.18 Konfigurasi node Color Manager.....	47
4.19 Penambahan label warna.....	48
4.20 Workflow Evaluasi Menggunakan <i>Silhouette Coefficient</i>	48
4.21 Node Konversi <i>String to Number</i>	49
4.22 Data Hasil <i>Normalizer</i>	50
4.23 Pengaturan node <i>Interval Loop Start</i>	51
4.24 <i>Clustering</i> K-Means menggunakan <i>Interval Loop</i>	52
4.25 Tangkapan Layar tabel <i>Silhouette Coefficient</i>	52

4.26 Nilai <i>Silhouette Coefficient</i> pada $k=15$	53
4.27 Mengambil nilai <i>Silhouette Coefficient</i> dan Iterasi	54
4.28 <i>Grouping Data Silhouette Coefficient</i>	55
4.29 Pengaturan <i>End Loop</i>	56
4.30 Tampilan tabel <i>Loop End</i>	56
4.31 Grafik <i>Silhouette Coefficient</i>	58
4.32 Workflow evaluasi menggunakan WCSS	59
4.33 Konfigurasi Node <i>Joiner</i>	60
4.34 Hasil join	60
4.35 Pengaturan node <i>Column Filter</i>	61
4.36 Hitung Jarak Antar Klaster dengan WCSS	62
4.37 Hitung jumlah <i>squared errors</i>	62
4.38 Nilai WCSS Pada k Iterasi Pertama	63
4.39 Mengubah Variabel Menjadi Kolom	63
4.40 Pengaturan Loop End	64
4.41 Hasil WCSS Tiap k yang Diulang	64
4.42 Menemukan Titik Siku pada grafik Elbow	65
4.43 Clustering menggunakan jumlah k yang optimal	66
4.44 Hasil <i>clustering</i> dengan nilai k yang optimal	66
4.45 Menambahkan hasil klustering kedalam file excel	67
4.46 Grafik Parallel Coordinates Plot	67
4.47 Visualisasi dengan <i>Parallel Coordinates Plot</i>	68
4.48 Visualisasi klaster 0 dengan <i>Parallel Coordinates Plot</i>	69
4.49 Visualisasi klaster 1 dengan <i>Parallel Coordinates Plot</i>	70
4.50 Visualisasi klaster 2 dengan <i>Parallel Coordinates Plot</i>	71
4.51 Visualisasi klaster 3 dengan <i>Parallel Coordinates Plot</i>	72

4.52 Visualisasi klaster 4 dengan <i>Parallel Coordinates Plot</i>	73
--	----

DAFTAR LAMPIRAN

No Nama Lampiran

1. Berita Acara Bimbingan Skripsi
2. Riwayat Hidup