

TUGAS AKHIR

**ANALISIS PADA TINGKAT PENYELESAIAN PENDIDIKAN
DI 34 PROVINSI INDONESIA MENGGUNAKAN METODE
ALGORITMA *K-MEANS CLUSTERING***



NOERALITA VEFILIANA AUDYA WULANDARI

2010651122

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH JEMBER**

2025

TUGAS AKHIR

**ANALISIS PADA TINGKAT PENYELESAIAN PENDIDIKAN
DI 34 PROVINSI INDONESIA MENGGUNAKAN METODE
ALGORITMA *K-MEANS CLUSTERING***

Disusun untuk Memenuhi Persyaratan Guna Meraih Gelar Sarjana Komputer
Teknik Informatika Universitas Muhammadiyah Jember



NOERALITA VEFILIANA AUDYA WULANDARI

2010651122

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH JEMBER**

2025

SURAT PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Noeralita Vefiliana Audya Wulandari
NIM : 2010651122
Program Studi : S1 Teknik Informatika
Perguruan Tinggi : Universitas Muhammadiyah Jember

Menyatakan dengan sesungguhnya bahwa Tugas Akhir berjudul “ANALISIS PADA TINGKAT PENYELESAIAN PENDIDIKAN DI 34 PROVINSI INDONESIA MENGGUNAKAN METODE ALGORITMA K-MEANS CLUSTERING” adalah benar merupakan karya sendiri. Hal-hal yang bukan karya saya, dalam skripsi tersebut diberi tanda citasi dan ditunjukkan dalam Daftar Pustaka. Apabila di kemudian hari terbukti pernyataan saya tidak benar dan ditemukan pelanggaran atas karya Tugas Akhir ini, saya bersedia menerima sanksi akademik jika pernyataan ini tidak benar.

Jember, 08 Februari 2025

Yang menyatakan,



Noeralita Vefiliana Audya Wulandari
2010651122

HALAMAN PERSETUJUAN TUGAS AKHIR

**ANALISIS PADA TINGKAT PENYELESAIAN PENDIDIKAN
DI 34 PROVINSI INDONESIA MENGGUNAKAN METODE
ALGORITMA K-MEANS CLUSTERING**

Oleh:

Noeralita Vefiliana Audya Wulandari

2010651122

Telah disetujui bahwa Laporan Tugas Akhir ini diajukan pada sidang Tugas Akhir sebagai salah satu syarat kelulusan dan mendapatkan gelar Sarjana Komputer

(S.Kom)

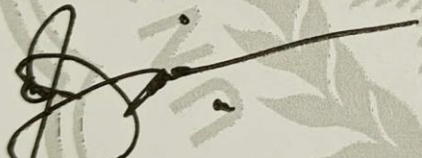
di

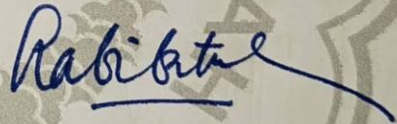
Universitas Muhammadiyah Jember

Disetujui oleh,

Pembimbing 1

Pembimbing 2


Garuh Wijaya, S.T, M.Kom.
NIDN.0729017601


Habibatul Azizah Al Faruq, M.Pd.
NIDN. 0718128901

LEMBAR PENGESAHAN

ANALISIS PADA TINGKAT PENYELESAIAN PENDIDIKAN
DI 34 PROVINSI INDONESIA MENGGUNAKAN METODE
ALGORITMA K-MEANS CLUSTERING


Oleh,

Noeralita Vefiliana Audya Wulandari
2010651122

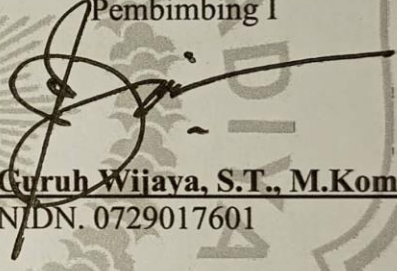
Telah mempertanggung jawabkan Laporan Tugas Akhir pada Sidang Tugas Akhir
tanggal 08 Februari 2025 sebagai salah satu syarat kelulusan dan mendapatkan
gelar Sarjana Komputer (S.Kom) di Universitas Muhammadiyah Jember

Disetujui oleh,

Dosen Penguji:
Penguji I


Ari Eko Wardoyo, S.T., M.Kom
NIDN. 0014027501

Dosen Pembimbing:
Pembimbing I


Guruh Wijaya, S.T., M.Kom
NIDN. 0729017601

Penguji II


Dewi Lusiana Pater, Ir. MT
NIDN. 0712086702

Pembimbing II


Habibatul Azizah Al-Faruq, M.Pd
NIDN. 0718128901

Mengesahkan,
Dekan Fakultas Teknik



Dr. Ir. Muhtar, S.T., M.T., IPM
NIDN. 19736102005011001

Mengesahkan,
Ketua Program Studi
Teknik Informatika



Rosita Yanuarti, S.Kom., M.Cs
NIDN. 0629018601

KATA PENGANTAR

Puji Syukur penulis panjatkan kehadirat Allah S.W.T., atas segala rahmat dan hidayah-Nya yang telah dilimpahkan, sehingga penulis dapat menyelesaikan tugas akhir dengan judul: “Analisis Pada Tingkat Penyelesaian Pendidikan Di 34 Provinsi Indonesia Menggunakan Metode Algoritma *K-Means Clustering*”

Penelitian tugas akhir ini mendasarkan pada isu Tingkat Penyelesaian Pendidikan menurut Jenjang, Tugas akhir ini merupakan karya ilmiah yang disusun dalam upaya untuk menyelesaikan pendidikan sarjana (S1) pada Fakultas Teknik Prodi Teknik Informatika Universitas Muhammadiyah Jember.

Penulis sangat berterimakasih kepada Bapak Guruh Wijaya, S.T.,M.Kom selaku pembimbing utama, Ibu Habibatul Azizah Al-Faruq, M.Pd, selaku pembimbing kedua, atas segala perhatian dan bimbingannya serta arahan-arahan yang diberikan kepada penulis dalam upaya menyelesaikan tugas akhir ini.

Terimakasih penulis disampaikan pula kepada Bapak Ari Eko Wardoyo, S.T.,M.Kom dan Ibu Dewi Lusiana Pater, Ir.,M.t, atas bantuan dan kesediaan serta saran-saran yang diberikan kepada penulis dalam ujian tugas akhir.

Ucapan terimakasih yang sebesar-besarnya penulis sampaikan kepada Bapak Dr .Ir. Muhtar, S.T., M.T., IPM selaku Dekan Fakultas Teknik, atas kesediaannya penulis belajar di Fakultas Teknik Prodi Teknik Informatika Universitas Muhammadiyah Jember.

Tidak lupa mengucapkan banyak terimakasih kepada para teman-teman prodi Teknik Informatika yang telah memberikan bantuan data dan informasi selama pelaksanaan penelitian. Harapan penulis semoga laporan hasil penelitian tugas akhir ini bisa bermanfaat bagi pembaca dan berguna bagi pengembangan Ilmu Informatika.

DAFTAR ISI

HALAMAN SAMPUL.....	ii
SURAT PERNYATAAN	iii
HALAMAN PERSETUJUAN TUGAS AKHIR.....	iv
LEMBAR PENGESAHAN	v
MOTTO	vi
PERSEMBAHAN.....	vii
ABSTRAK	viii
ABSTRACT	ix
KATA PENGANTAR.....	x
DAFTAR ISI.....	xi
DAFTAR TABEL	xiii
DAFTAR GAMBAR.....	xv
BAB 1 PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Tujuan Penelitian.....	2
1.4 Manfaat Penelitian.....	3
1.5 Batasan Penelitian.....	3
BAB 2 KAJIAN PUSTAKA	4
2.1 <i>Data Mining</i>	4
2.2 Algoritma <i>K-Means Clustering</i>	4
2.3 <i>RapidMiner</i>	5
2.4 Tingkat Penyelesaian Pendidikan.....	7
2.5 Penelitian Terdahulu	8
BAB 3 METODOLOGI PENELITIAN	9
3.1 Tahapan Penelitian	12
3.2 Teknik Pengumpulan Data	12
3.3 Jenis Data.....	13
3.4 Sumber Data.....	13
3.5 Implementasi Algoritma <i>K-Means Clustering</i>	14
BAB 4 HASIL DAN PEMBAHASAN	26

4.1 Perhitungan dan Implementasi Menggunakan <i>Tools RapidMiner</i>	26
BAB 5 KESIMPULAN DAN SARAN	33
5.1 Kesimpulan	33
5.2 Saran	34
DAFTAR PUSTAKA	35
LAMPIRAN-LAMPIRAN	38



DAFTAR TABEL

Tabel 3. 1 Sumber Data BPS	13
Tabel 3. 2 Penentuan <i>Cluster</i> Dari Data.....	14
Tabel 3. 3 <i>Centroid</i> Awal.....	15
Tabel 3. 4 Menghitung Jarak Setiap Data SD ke <i>Centroid</i>	15
Tabel 3. 5 Menghitung Jarak Setiap Data SMP ke <i>Centroid</i>	17
Tabel 3. 6 Menghitung Jarak Setiap Data SMA ke <i>Centroid</i>	18
Tabel 3. 7 Iterasi ke 1	19
Tabel 3. 8 <i>Centroid</i> Baru Iterasi 2.....	20
Tabel 3. 9 Menghitung Jarak Setiap Data SD ke <i>Centroid</i> (Iterasi 2).....	20
Tabel 3. 10 Menghitung Jarak Setiap Data SMP ke <i>Centroid</i> (Iterasi 2).....	21
Tabel 3. 11 Menghitung Setiap Jarak Data SMA ke <i>Centroid</i> (Iterasi 2).....	22
Tabel 3. 12 Iterasi ke 2.....	24
Tabel 3. 13 Hasil Pengelompokan Data SD.....	25
Tabel 3. 14 Hasil Pengelompokan Data SMP.....	25
Tabel 3. 15 Hasil Pengelompokan Data SMA.....	25
Tabel 4. 1 Prosentase Kelulusan.....	30
Tabel 3. 16 <i>Centroid</i> Baru Iterasi 3.....	38
Tabel 3. 17 Menghitung Jarak Setiap Data SD ke <i>Centroid</i> (Iterasi ke 3).....	38
Tabel 3. 18 Menghitung Jarak Setiap Data SMP ke <i>Centroid</i> (Iterasi 3).....	39
Tabel 3. 19 Menghitung Jarak Setiap Data SMA ke <i>Centroid</i> (Iterasi 3).....	40
Tabel 3. 20 Iterasi ke 3.....	41
Tabel 3. 21 <i>Centroid</i> Baru Iterasi 4.....	42
Tabel 3. 22 Menghitung Jarak Setiap Data SD ke <i>Centroid</i> (Iterasi 4).....	43
Tabel 3. 23 Menghitung Jarak Setiap Data SMP ke <i>Centroid</i> (Iterasi 4).....	44
Tabel 3. 24 Menghitung Jarak Setiap Data SMA ke <i>Centroid</i> (Iterasi 4).....	45
Tabel 3. 25 Iterasi ke 4.....	46
Tabel 3. 26 <i>Centroid</i> Baru Iterasi 5.....	47
Tabel 3. 27 Menghitung Jarak Setiap Data SD ke <i>Centroid</i> (Iterasi 5).....	47
Tabel 3. 28 Menghitung Setiap Jarak Data SMP ke <i>Centroid</i> (Iterasi 5).....	49
Tabel 3. 29 Menghitung Jarak Setiap Data SMA ke <i>Centroid</i> (Iterasi 5).....	50
Tabel 3. 30 Iterasi ke 5.....	51
Tabel 3. 31 <i>Centroid</i> Baru Iterasi 6.....	52
Tabel 3. 32 Menghitung Jarak Setiap Data SD ke <i>Centroid</i> (Iterasi 6).....	52
Tabel 3. 33 Menghitung Jarak Setiap Data SMP ke <i>Centroid</i> (Iterasi 6).....	53
Tabel 3. 34 Menghitung Jarak Setiap Data SMA ke <i>Centroid</i> (Iterasi 6).....	54
Tabel 3. 35 Iterasi ke 6.....	56
Tabel 3. 36 <i>Centroid</i> Baru Iterasi 7.....	57
Tabel 3. 37 Menghitung Jarak Setiap Data SD ke <i>Centroid</i> (Iterasi 7).....	57
Tabel 3. 38 Menghitung Jarak Setiap Data SMP ke <i>Centroid</i> (Iterasi 7).....	58
Tabel 3. 39 Menghitung Jarak Setiap Data SMA ke <i>Centroid</i> (Iterasi 7).....	59
Tabel 3. 40 Iterasi ke 7.....	60
Tabel 3. 41 <i>Centroid</i> Baru Iterasi 8.....	61
Tabel 3. 42 Menghitung Jarak Setiap Data SD ke <i>Centroid</i> (Iterasi 8).....	62
Tabel 3. 43 Menghitung Jarak Setiap Data SMP ke <i>Centroid</i> (Iterasi 8).....	63
Tabel 3. 44 Menghitung Jarak Setiap Data SMA ke <i>Centroid</i> (Iterasi 8).....	64

Tabel 3. 45 Iterasi ke 8.....	65
Tabel 3. 46 <i>Centroid</i> Baru Iterasi 9.....	66
Tabel 3. 47 Menghitung Jarak Setiap Data SD ke <i>Centroid</i> (Iterasi 9)	66
Tabel 3. 48 Menghitung Jarak Setiap Data SMP ke <i>Centroid</i> (Iterasi 9)	68
Tabel 3. 49 Menghitung Jarak Setiap Data SMA ke <i>Centroid</i> (Iterasi 9).....	69
Tabel 3. 50 Iterasi ke 9.....	70



DAFTAR GAMBAR

Gambar 3. 1 Diagram Alir Penelitian	12
Gambar 4. 1 Tampilan awal <i>RapidMiner Studio</i>	26
Gambar 4. 2 Operator Data dan <i>Clustering</i>	26
Gambar 4. 3 Parameter <i>Clustering</i>	27
Gambar 4. 4 Pengujian <i>Cluster</i> Optimum $K=4$	27
Gambar 4. 5 Hasil <i>Clustering</i> Data	28
Gambar 4. 6 Data Dalam Setiap <i>Cluster</i>	28
Gambar 4. 7 Hasil Evaluasi <i>Cluster</i>	29
Gambar 4. 8 Hasil Visualisasi <i>Clustering</i>	30

