



**PERBANDINGAN ALGORITMA *K-MEANS* DAN
ALGORITMA *K-MEDOIDS* DALAM PENGELOMPOKAN
KABUPATEN/KOTA BERDASARKAN JUMLAH PENDUDUK
MISKIN DI JAWA TENGAH DAN DIY**

Tugas Akhir

Disusun sebagai salah satu syarat untuk memperoleh gelah Ahli Madya
Program Studi Statistika Terapan dan Komputasi

Oleh
Wawan Firmansyah
4112320015

**PROGRAM STUDI STATISTIKA TERAPAN DAN KOMPUTASI
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI SEMARANG
SEMARANG, 2023**



**PERBANDINGAN ALGORITMA *K-MEANS* DAN
ALGORITMA *K-MEDOIDS* DALAM PENGELOMPOKAN
KABUPATEN/KOTA BERDASARKAN JUMLAH PENDUDUK
MISKIN DI JAWA TENGAH DAN DIY**

Tugas Akhir

Disusun sebagai salah satu syarat untuk memperoleh gelah Ahli Madya
Program Studi Statistika Terapan dan Komputasi

Oleh
Wawan Firmansyah
4112320015

**PROGRAM STUDI STATISTIKA TERAPAN DAN KOMPUTASI
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI SEMARANG
SEMARANG, 2023**

PERSETUJUAN PEMBIMBING

Tugas Akhir berjudul "Perbandingan Algoritma K-Means dan Algoritma K-Medoids dalam Pengelompokan Kabupaten/Kota Berdasarkan Jumlah Penduduk Miskin di Jawa Tengah dan DIY" yang disusun oleh:

Nama : Wawan Firmansyah

NIM : 4112320015

Prodi : D-III Statistika Terapan dan Komputasi

Telah disetujui untuk diajukan ke sidang ujian tugas akhir

Semarang, 4 September 2023

Pembimbing



Dr. Arief Agoestanto, M.Si.

196807221993031005

PENGESAHAN PENGUJI

Tugas akhir berjudul "Perbandingan Algoritma *K-Means* dan Algoritma *K-Medoids* dalam Pengelompokan Kabupaten/Kota Berdasarkan Jumlah Penduduk Miskin di Jawa Tengah dan DIY" yang disusun oleh

Nama : Wawan Firmansyah

NIM : 4112320015

Prodi : Statistik Terapan dan Komputasi

Telah dipertahankan dalam ujian tugas akhir pada hari Senin, 18 September 2023.

Tim Penguji

Ketua Panitia Prof. Dr. Edy Cahyono, M.Si. NIP 196412051990021001	
Sekretaris Panitia Dr. Putriaji Hendikawati, S.Si., M.Pd., M.Sc. NIP 198208182006042001	
Penguji 1 Dr. Drs. Sugiman, M.Si. NIP 196401111989011001	
Penguji 2 Amidi, S.Si., M.Pd. NIP 198703012014041001	
Penguji 3/Pembimbing Dr. Arief Agoestanto, M.Si. NIP 196807221993031005	

PERNYATAAN

Tugas Akhir yang ditulis berjudul “Perbandingan Algoritma *K-Means* dan Algoritma *K-Medoids* dalam Pengelompokan Kabupaten/Kota Berdasarkan Jumlah Penduduk Miskin di Jawa Tengah dan DIY” Merupakan karya ilmiah asli dan bukan hasil plagiasi dari karya ilmiah orang lain. Pendapat atau temuan orang lain yang dikutip di dalam Tugas Akhir/Skripsi/Publikasi Ilmiah ini telah ditulis berdasarkan kode etik ilmiah.

Semarang, 4 September 2023

Yang menyatakan



Wawan Firmansyah

NIM 4112320015

MOTTO

"Seungguhnya Allah tidak akan mengubah keadaan suatu kaum, sebelum mereka mengubah keadaan diri mereka sendiri."

(QS Ar Rad 11)

“Takdir setiap manusia memang telah ditentukan sejak mereka lahir, tetapi dengan kerja keras kita dapat mengalahkan takdir”

(Uzumaki Naruto)

"Jangan pernah lihat ke belakang, mengenang hari kemarin, tersenyumlah untuk masa depan, esok adalah hari baru."

(Zayn Malik)

PERSEMBAHAN

1. Allah SWT yang telah memberikan Rahmat, Nikmat, dan Karunia-Nya, serta Kemudahan sehingga yang bisa menyelesaikan Tugas Akhir ini dengan baik.
2. Kedua orang tua saya selalu memberikan doa dan dukungannya.
3. Keluarga besar kakek dan nenek saya yang selalu memberikan nasihat dan motivasi dalam menyelesaikan Tugas Akhir.
4. Kakak tercinta yan selalu menghibur.
5. Bapak pembimbing Tugas Akhir yang senantiasa tidak pernah lelah membimbing saya.
6. Teman-teman Program Studi Statistik Terapan dan Komputasi Angkatan 2020
7. Semua pihak yang tidak bisa saya sebutkan satu persatu.

ABSTRAK

Firmansyah, Wawan (2023), *Perbandingan Algoritma K-Means dan Algoritma K-Medoids dalam Pengelompokan Kabupaten/Kota Berdasarkan Jumlah Penduduk Miskin di Jawa Tengah dan DIY.* Tugas Akhir, Statistika Terapan dan Komputasi Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Semarang. Pembimbing Dr. Arief Agoestanto, M.Si.

Kata kunci: *K-Means, K-Medoids, cluster, perbandingan*

Analisis *cluster* merupakan metode data mining yang banyak digunakan untuk menentukan struktur dalam kumpulan data berdasar kesamaan karakteristik terbesar dalam *cluster*. Penelitian ini bertujuan untuk membandingkan hasil analisis *cluster K-Means* dan *K-Medoids* dari data jumlah penduduk miskin Provinsi Jawa Tengah dan DIY tahun 2015-2022. Analisis *cluster* terbagi menjadi herarki dan non-herarki, dalam penelitian ini digunakan *non-herarki* yaitu metode *K-Means* dan *K-Medoids*. *K-Means* adalah metode *clustering* yang menggunakan mean sebagai pusat *cluster*, sedangkan *K-Medoids* menggunakan *medoid*, yang merupakan data poin yang paling mirip dengan pusat *cluster*. Jumlah penduduk miskin merupakan indikator penting dalam perencanaan pembangunan dan kebijakan sosial. Di Pulau Jawa sendiri Provinsi Jawa Tengah menempati peringkat kedua dengan presentase kemiskinan 10,98%, yang mana pada peringkat pertama diperoleh oleh Provinsi Daerah Istimewa Yogyakarta dengan presentase kemiskinan 11,49%. Oleh sebab itu, fokus penulis pada penelitian ini adalah Provinsi Jawa Tengah dan karena presentase kemiskinannya yang cukup tinggi. Untuk membandingkan hasil dari analisis maka dilakukan uji validitas yaitu dengan metode *Silhouette Index* dan *Davies Buildin Index* (DBI).

Hasil *cluster* menunjukkan bahwa metode *K-Means* menghasilkan 5 *cluster* yaitu *cluster* 1 terdiri dari 6 Kabupaten/kota, kelompok 2 terdiri dari 5 kabupaten/kota, *cluster* 3 terdiri dari 13 kabupaten/kota, *cluster* 4 terdiri dari 8 kabupaten/kota, dan *cluster* 5 terdiri dari 8 kabupaten/kota. Sedangkan hasil *cluster* menggunakan algoritma *K-Medoids* menghasilkan 4 *cluster* yaitu *cluster* 1 terdiri dari 5 kabupaten/kota, *cluster* 2 terdiri dari 12 kabupaten/kota, *cluster* 3 terdiri dari 17 kabupaten/kota, *cluster* 4 terdiri dari 6 kabupaten/kota. Untuk hasil perbandingan yaitu nilai *Davies Bouldin Indeks* (DBI) *K-Means* adalah 0,4036484 lebih kecil dibandingkan *K-Medoids* yaitu 0,5274465. Untuk nilai *Silhouette index K-Means* adalah 0,6162525 lebih besar dibandingkan *K-Medoids* yaitu 0,6085647. Nilai DBI algoritma *K-Means* lebih kecil dari algoritma *K-Medoids* sedangkan nilai *Silhouette Index* algoritma *K-Means* lebih besar dari algoritma *K-Medoids*, hal ini menunjukkan bahwa algoritma *K-Means* lebih baik daripada algoritma *K-Medoids*.

PRAKATA

Segala Puji dan syukur kehadiran Allah SWT yang telah melimpahkan Rahmat dan Hidayah-Nya sehingga peneliti dimudahkan dan dapat menyelesaikan Tugas Akhir dengan judul “*Perbandingan Algoritma K-Means dan Algoritma K-Medoids dalam Pengelompokan Kabupaten/Kota Berdasarkan Jumlah Penduduk Miskin di Jawa Tengah dan DIY*”.

Penelitian Tugas Akhir ini diajukan sebagai salah satu syarat untuk memperoleh gelar Diploma III Program Studi Statistik Terapan dan Komputasi. Keberhasilan dalam penyelesaian Tugas Akhir ini tidak terlepas dari bantuan dan dukungan berbagai pihak. Oleh karena itu, peneliti mengucapkan terima kasih kepada:

1. Prof. Dr. S Martono, M.Si., selaku Rektor Universitas Negeri Semarang.
2. Prof. Dr. Edy Cahyono, M.Si., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Semarang.
3. Dr. Putriaji Hendikawati, S.Si., M.Pd., M.Sc. selaku Koordinator Program Studi Statistika Terapan dan Komputasi Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Semarang.
4. Dr. Arief Agoestanto, M.Si. selaku dosen pembimbing utama yang telah memberikan bimbingan dan saran serta motivasi dalam penyusunan Tugas Akhir ini.
5. Bapak dan Ibu dosen Program Studi Statistika Terapan dan Komputasi Universitas Negeri Semarang yang telah memberikan bekal ilmu yang bermanfaat kepada peneliti.
6. Kedua orang tua yang telah memberikan doa dan dukungan.
7. Keluarga besar kakek dan nenek yang telah memberikan motivasi dan nasihat.
8. Kakak saya yang selalu menghibur dalam penyelesaian Tugas akhir ini.
9. Teman-teman Program Studi Statistika Terapan dan Komputasi Angkatan 2020.
10. Semua pihak yang telah membantu dalam penyusunan tugas akhir ini yang tidak bisa saya sebutkan satu persatu.

Penulis menyadari bahwa dalam penyusunan tugas akhir ini memiliki banyak kekurangan dan masih jauh dari kata sempurna, sehingga penulis berharap kritik dan saran yang membangun guna untuk memperbaiki penelitian ini.

Semarang, 4 September 2023

Wawan Firmansyah

DAFTAR ISI

COVER.....	i
PERSETUJUAN PEMBIMBING.....	ii
PENGESAHAN PENGUJI.....	iii
PERNYATAAN.....	iv
MOTTO.....	v
PERSEMBAHAN.....	vi
ABSTRAK.....	vii
PRAKATA.....	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xii
DAFTAR GAMBAR.....	xiii
DAFTAR LAMPIRAN.....	xiv
BAB 1 PENDAHULUAN.....	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	4
1.3. Tujuan Penelitian.....	4
1.4. Batasan Masalah.....	4
1.5. Manfaat Penelitian.....	4
1.6. Keaslian Penelitian.....	5
BAB II KAJIAN PUSTAKA.....	7
2.1. Tinjauan Pustaka.....	7
2.2. Landasan Teori.....	8
2.2.1. <i>Data Mining</i>	8
2.2.2. Statistik Deskriptif.....	9
2.2.3. <i>Analisis Cluster</i>	9
2.2.4. Analisis Komponen Utama.....	11
2.2.5. Teknik-teknik Analisis Dalam <i>Cluster</i>	12
2.2.6. <i>Jarak Euclidean</i>	15
2.2.7. Metode <i>Elbow</i>	15
2.2.8. <i>Silhouette Coefficient</i>	15
2.2.9. <i>Davies-Bouldin Index (DBI)</i>	16

2.2.10. Software RStudio	16
2.2.11. Kemiskinan	17
2.2.12. Penduduk Miskin	20
2.3. Kerangka Berpikir	22
BAB III METODE PENELITIAN.....	24
3.1. Sumber Data	24
3.1.1. Jenis Data	24
3.1.2. Sumber Data	24
3.2. Jenis Penelitian	24
3.3. Teknik Pengumpulan Data	24
3.4. Langkah-langkah Analisi Data	24
3.5. Tahapan Penelitian.....	25
BAB IV HASIL DAN PEMBAHASAN.....	30
4.1. Asumsi Analisis <i>Cluster</i>	30
4.2. Analisis Komponen Utama	31
4.3. Penentuan Banyak <i>Cluster</i>	35
4.4. Visualisasai cluster <i>K-Means</i> dan <i>K-Medoids</i>	39
4.5. Interpretasi <i>Cluster</i>	42
4.6. Perbandingan <i>Cluster</i>	48
BAB V PENUTUP.....	51
5.1. Simpulan.....	51
5.2. Saran.....	51
DAFTAR PUSTAKA	52
Lampiran	55