

**PERBANDINGAN FUZZY C-MEANS DAN K-MEANS UNTUK
MENGELOMPOKKAN TINGKAT BUTA HURUF BERDASARKAN
PROVINSI DI INDONESIA**

SKRIPSI



Disusun oleh:

MURTADLO ANUGERAH PAMUNGKAS

1610651010

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH JEMBER
2021**

**PERBANDINGAN FUZZY C-MEANS DAN K-MEANS UNTUK
MENGELOMPOKKAN TINGKAT BUTA HURUF BERDASARKAN
PROVINSI DI INDONESIA**

Disusun Untuk Melengkapi dan Memenuhi Syarat Kelulusan

Guna Meraih Gelar Sarjana Komputer

Teknik Informatika Universitas Muhammadiyah Jember

SKRIPSI



Disusun oleh:

MURTADLO ANUGERAH PAMUNGKAS

1610651010

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS MUHAMMADIYAH JEMBER

2021

HALAMAN PERSETUJUAN SIDANG TUGAS AKHIR

PERBANDINGAN *FUZZY C-MEANS* DAN *K-MEANS* UNTUK MENGELOMPOKKAN TINGKAT BUTA HURUF BERDASARKAN PROVINSI DI INDONESIA

Oleh:

Murtadlo Anugerah Pamungkas

1610651010

Telah disetujui bahwa Laporan Tugas Akhir ini untuk diajukan pada sidang Tugas Akhir sebagai salah satu syarat kelulusan dan mendapatkan gelar Sarjana Komputer (S.Kom)
Di Universitas Muhammadiyah Jember

Disetujui oleh,

Pembimbing I

Pembimbing II

Hardian Oktavianto, S.Si.,M.Kom

Reni Umilasari, S.Pd, M.Si

NIDN. 0722108105

NIDN. 0728079101

HALAMAN PENGESAHAN

PERBANDINGAN FUZZY C-MEANS DAN K-MEANS UNTUK MENGELOMPOKKAN TINGKAT BUTA HURUF BERDASARKAN PROVINSI DI INDONESIA

Oleh:

Murtadlo Anugerah Pamungkas

1610651010

Disetujui Oleh,

Pengaji I


Dr. Bagus Setya Rintvarna S.T, M.Kom
NIDN. 0729017904

Pembimbing I


Hardian Oktavianto S.Si., M.Kom
NIDN. 0722108105

Pengaji II


Henny Wahyu Sulistyo S.Kom, M.Kom
NIDN. 0718088309

Pembimbing II


Reni Umilasari, S.Pd, M.Si
NIDN. 0728079101

Mengesahkan,
Dekan Fakultas Teknik



Dr. Nanang Saiful Rizal, S.T, M.T
NPK. 03 08 366

Mengetahui,
Ketua Program Studi Teknik Informatika



Triawan Adi Cahyanto, M.Kom
NPK. 12 03 719

PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Murtadlo Anugerah Pamungkas
NIM : 1610651010
Program Studi : Teknik Informatika
Perguruan Tinggi : Universitas Muhammadiyah Jember

Menyatakan bahwa Tugas Akhir yang berjudul "**“PERBANDINGAN FUZZY C-MEANS DAN K-MEANS UNTUK MENGELOMPOKKAN TINGKAT BUTA HURUF BERDASARKAN PROVINSI DI INDONESIA”**" bukan merupakan Tugas Akhir orang lain baik sebagian maupun keseluruhan kecuali yang diacu dalam Daftar Pustaka dan kutipan yang telah disebutkan sumbernya pada Tugas Akhir ini.

Demikian Surat Pernyataan ini dibuat dengan sebenar-benarnya dan apabila pernyataan ini tidak sesuai, penulis bersedia mendapatkan sanksi dari akademik.

Jember, 11 Maret 2021



Murtadlo Anugerah Pamungkas

NIM. 1610651010

KATA PENGANTAR

Alhamdulillah segala puji bagi Allah yang senantiasa memberikan rahmat dan karunianya terhadap penulis, sehingga penulis dapat menyelesaikan skripsi ini dan diberikan kemudahan dalam menyelesaikan studi di kampus Universitas Muhammadiyah Jember.

Skripsi ini berjudul “Perbandingan *Fuzzy C-Means* dan *K-Means* Untuk Mengelompokkan Tingkat Buta Huruf Berdasarkan Provinsi Di Indonesia” dengan memuat bab I sampai bab V. Bab I berisi pendahuluan, bab II berisi tinjauan pustaka, bab III berisi metode penelitian, bab IV berisi hasil penelitian dan bab V berisi kesimpulan dan saran

Dengan penuh kesadaran penulis menyampaikan permohonan maaf atas kekurangan yang ada pada penelitian skripsi ini. Semoga bisa menjadi koreksi bersama untuk perbaikan selanjutnya.

Jember, 11 Maret 2021

Penulis

UNGKAPAN TERIMAKASIH

Skripsi ini merupakan bukti dari terselesaiannya materi-materi mata kuliah yang telah ditempuh pada jenjang S1 Teknik Informatika (TI) Universitas Muhammadiyah Jember. Atas segala upaya, bimbingan dan arahan dari semua pihak, penulis mengucapkan terimakasih yang sebesar-besarnya kepada:

1. Kedua orang tua ibu Sulastri dan ayah Sajat yang telah menjaga, menyayangi, membimbing, memberi semangat serta doa kepada penulis dengan tulus dan ikhlas. Gelar sarjana ini penulis persembahkan untuk kalian.
2. Dr. Nanang Saiful Rizal, M.T. selaku Dekan Fakultas Teknik Universitas Muhammadiyah Jember.
3. Bapak Triawan Adi Cahyanto, M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas Muhammadiyah Jember.
4. Bapak Hardian Oktavianto, S.Si., M.Kom selaku dosen pembimbing I dan Ibu Reni Umilasari, S.Pd, M.Si selaku dosen pembimbing II yang telah meluangkan waktu untuk memberi arahan dan membimbing penulis dalam menyelesaikan tugas akhir ini.
5. Bapak Dr. Bagus Setya Rintyarna, ST, M.Kom selaku dosen penguji I dan Bapak Henny Wahyu Sulistyo, S.Kom, M.Kom selaku dosen penguji II yang telah memberikan saran dan masukan dalam penyelesaian tugas akhir ini.
6. Kepada Bapak dan Ibu Dosen Program Studi Teknik Informatika Universitas Muhammadiyah Jember yang telah memberikan ilmunya kepada penulis.
7. Teruntuk Mas tercinta Moch. Taufiq Ridlo, Mohammad Nuruddin Effendi, dan Mbak Eny Muhimmatul Fitriyani yang telah memberi semangat kepada penulis.
8. Terimakasih juga kepada Chindy Precillia yang telah memberikan support dan motivasi kepada penulis sehingga skripsi ini terselesaikan.

9. Juga kepada teman-teman seperjuangan yang selalu ada di kala suka maupun duka selama perkuliahan.

Dan juga kepada semua pihak yang tidak bisa disebutkan satu persatu. Disampaikan terimakasih atas semua bantuan dan bimbingannya.

Jember, 11 Maret 2021

Penulis



MOTTO

“Barang siapa bertakwa kepada Allah maka Dia akan menjadikan jalan keluar baginya, dan memberinya rezeki dari jalan yang tidak ia sangka, dan barang siapa bertawakal kepada Allah maka cukuplah Allah baginya, sesungguhnya Allah melaksanakan kehendak-Nya Dia telah menjadikan untuk setiap sesuatu kadarnya”

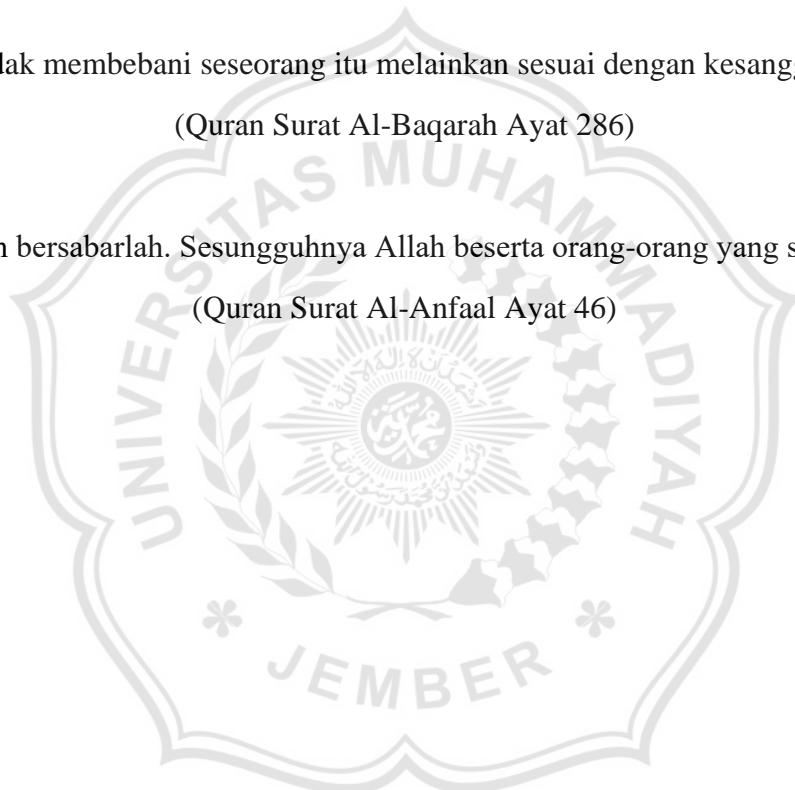
(Quran Surat Ath-Thalaq Ayat 2-3)

“Allah tidak membebani seseorang itu melainkan sesuai dengan kesanggupannya”

(Quran Surat Al-Baqarah Ayat 286)

“Dan bersabarlah. Sesungguhnya Allah beserta orang-orang yang sabar”

(Quran Surat Al-Anfaal Ayat 46)



DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN.....	iii
PERNYATAAN.....	iv
KATA PENGANTAR.....	v
UNGKAPAN TERIMAKASIH	vi
MOTTO	viii
ABSTRAK	ix
ABSTRACT	x
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
BAB II TINJAUAN PUSTAKA	5
2.1 Provinsi di Indonesia	5
2.2 Buta Huruf	5
2.3 Data Mining.....	6
2.4 <i>Clustering</i>	8
2.5 <i>Fuzzy C-Means</i>	9
2.6 <i>K-Means</i>	11
2.7 Uji Validitas	12
2.7.1 Dunn Index.....	12
2.7.2 Partition Coefficient Index.....	13
2.7.3 Silhouette Index	13

BAB III METODOLOGI PENELITIAN	15
3.1 Metode penelitian	15
3.2 Diagram Alur Penelitian.....	15
3.3 Studi Awal	16
3.4 Pengumpulan Data	16
3.5 Implementasi Algoritma <i>Fuzzy C-Means</i>	16
3.6 Implementasi Algoritma <i>K-Means</i>	24
3.7 <i>Cluster Optimum</i>	30
3.8 Uji Validitas	31
3.8.1 Algoritma Fuzzy C-Means dengan Partition Coefficient Index	32
3.8.2 Algoritma K-Means dengan Silhouette Index	33
BAB IV HASIL DAN PEMBAHASAN	37
4.1 Data Pengujian	37
4.2 <i>Clustering</i> pada RStudio	38
4.2.1 Fuzzy C-Means	38
4.2.2 K-Means.....	42
4.3 Penentuan <i>Cluster Optimum</i>	47
4.4 <i>Profiling Cluster Optimum</i>	49
4.4.1 Fuzzy C-Means	49
4.4.2 K-Means.....	52
4.5 Uji Validitas Algoritma.....	54
4.6 Perbedaan <i>Fuzzy C-Means</i> dan <i>K-means</i> dalam <i>Clustering</i> Data	56
BAB V KESIMPULAN DAN SARAN	57
5.1 Kesimpulan.....	57
5.2 Saran.....	57
DAFTAR PUSTAKA	58

DAFTAR GAMBAR

Gambar 3.1 Diagram Alur Penelitian	15
Gambar 3.2 <i>Flowchart Fuzzy C-Means</i>	17
Gambar 3.3 <i>Flowchart K-Means</i>	24
Gambar 4.1 <i>Plot 3 Cluster Fuzzy C-Means</i> pada RStudio.....	38
Gambar 4.2 <i>Plot 4 Cluster Fuzzy C-Means</i> pada RStudio.....	39
Gambar 4.3 <i>Plot 5 Cluster Fuzzy C-Means</i> pada RStudio.....	39
Gambar 4.4 <i>Plot 6 Cluster Fuzzy C-Means</i> pada RStudio.....	40
Gambar 4.5 <i>Plot 7 Cluster Fuzzy C-Means</i> pada RStudio.....	40
Gambar 4.6 <i>Plot 8 Cluster Fuzzy C-Means</i> pada RStudio.....	41
Gambar 4.7 <i>Plot 9 Cluster Fuzzy C-Means</i> pada RStudio.....	41
Gambar 4.8 <i>Plot 10 Cluster Fuzzy C-Means</i> pada RStudio.....	42
Gambar 4.9 <i>Plot 3 Cluster K-Means</i> pada RStudio	43
Gambar 4.10 <i>Plot 4 Cluster K-Means</i> pada RStudio	43
Gambar 4.11 <i>Plot 5 Cluster K-Means</i> pada RStudio	44
Gambar 4.12 <i>Plot 6 Cluster K-Means</i> pada RStudio	44
Gambar 4.13 <i>Plot 7 Cluster K-Means</i> pada RStudio	45
Gambar 4.14 <i>Plot 8 Cluster K-Means</i> pada RStudio	45
Gambar 4.15 <i>Plot 9 Cluster K-Means</i> pada RStudio	46
Gambar 4.16 <i>Plot 10 Cluster K-Means</i> pada RStudio	46
Gambar 4.17 Diagram <i>Cluster Optimum</i>	48

DAFTAR TABEL

Tabel 3.1 Data Penduduk Buta Huruf Berdasarkan Provinsi di Indonesia	18
Tabel 3.2 Parameter Awal.....	18
Tabel 3.3 Bilangan Random Iterasi 1	19
Tabel 3.4 Pusat <i>Cluster</i> Iterasi 1	20
Tabel 3.5 Fungsi Objektif (P) Iterasi 1	20
Tabel 3.6 Derajat Keanggotaan Iterasi 1	21
Tabel 3.7 Fungsi Objektif (P) Iterasi 2	22
Tabel 3.8 Fungsi Objektif (P) Iterasi 2	23
Tabel 3.9 Data Penduduk Buta Huruf Berdasarkan Provinsi di Indonesia	25
Tabel 3.10 <i>Centroid</i> Awal Iterasi 1.....	25
Tabel 3.11 Jarak Data Ke Titik Pusat (<i>Centroid</i>).....	26
Tabel 3.12 Pengelompokan Data Iterasi 1	27
Tabel 3.13 <i>Centroid</i> Awal (CA) Iterasi 1	29
Tabel 3.14 <i>Centroid</i> Baru (CB) Iterasi 1	29
Tabel 3.15 Kondisi Berhenti	29
Tabel 3.16 Pengelompokan Data Iterasi 2	29
Tabel 3.17 <i>Centroid</i> Awal (CA) Iterasi 2	30
Tabel 3.18 <i>Centroid</i> Baru (CA) Iterasi 2	30
Tabel 3.19 Kondisi Berhenti	30
Tabel 3.20 Nilai <i>Dunn Index</i> pada 3 sampai 10 <i>cluster</i>	31
Tabel 3.21 Uji Validitas <i>Partition Coefficient Index</i>	33
Tabel 3.22 Jarak Data Ke-1 dengan semua data	35
Tabel 3.23 Nilai <i>Silhouette Index</i> pada setiap data	36
Tabel 4.1 Data Tingkat Buta Huruf	37
Tabel 4.2 Nilai <i>Dunn Index</i>	47
Tabel 4.3 Anggota pada <i>Cluster</i> 1	49
Tabel 4.4 Karakteristik Cluster 1	49
Tabel 4.5 Anggota pada <i>Cluster</i> 2	49
Tabel 4.6 Karakteristik pada Cluster 2	50

Tabel 4.7 Anggota pada <i>Cluster 3</i>	50
Tabel 4.8 Karakteristik pada <i>Cluster 3</i>	50
Tabel 4.11 Anggota pada <i>Cluster 1</i>	52
Tabel 4.12 Karakteristik pada <i>Cluster 1</i>	52
Tabel 4.13 Anggota pada <i>Cluster 2</i>	52
Tabel 4.14 Karakteristik pada <i>Cluster 2</i>	52
Tabel 4.15 Anggota pada <i>Cluster 3</i>	53
Tabel 4.16 Karakteristik pada <i>Cluster 3</i>	53
Tabel 4.17 Anggota pada <i>Cluster 4</i>	53
Tabel 4.18 Karakteristik pada <i>Cluster 4</i>	54
Tabel 4.23 Nilai Validitas <i>Fuzzy C-Means</i>	55
Tabel 4.24 Nilai Validitas <i>K-Means</i>	55

