



**KOMPARASI METODE *K-MEANS* DAN *FUZZY C-MEANS* DALAM
KLASTERISASI KEPADATAN PENDUDUK KABUPATEN KOTA DI
PROVINSI KALIMANTAN SELATAN**

SKRIPSI

**Untuk memenuhi persyaratan
dalam menyelesaikan program sarjana Strata-1 Matematika**

**Oleh:
Nur Aulia Awwaliyah
NIM 1911011120005**

**PROGRAM STUDI MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU**

MARET 2023

SKRIPSI

KOMPARASI METODE *K-MEANS* DAN *FUZZY C-MEANS* DALAM KLASTERISASI KEPADATAN PENDUDUK KABUPATEN KOTA DI PROVINSI KALIMANTAN SELATAN


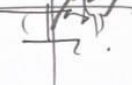
Oleh:
NUR AULIA AWWALIYAH
NIM 1911011120005

telah dipertahankan di depan Dosen Penguji pada tanggal 6 Maret 2023.
Susunan Dosen Penguji:


Pembimbing I


Akhmad Yusuf, S.Si., M.Kom.
NIP. 198004022005011001

Dosen Penguji:

1. Oni Soesanto, S.Si., M.Si. ()
2. Pardi Affandi, S.Si., M.Sc. ()

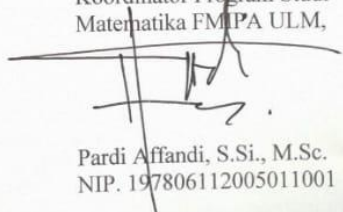
Pembimbing II


Yuana Sukmawaty, S.Si., M.Si.
NIP. 198810152015042002

Banjarbaru, Maret 2023



Koordinator Program Studi
Matematika FMIPA ULM,


Pardi Affandi, S.Si., M.Sc.
NIP. 197806112005011001

PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam Daftar Pustaka.

Banjarbaru, Maret 2023



Nur Aulia Awwaliyah
NIM. 1911011120005

ABSTRAK

Komparasi Metode *K-Means* Dan *Fuzzy C-Means* Dalam eva Kepadatan Penduduk Kabupaten Kota di Provinsi Kalimantan Selatan (Oleh: Nur Aulia Awwaliyah; Pembimbing: Akhmad Yusuf, Yuana Sukmawaty, 2022, 49 halaman)

Pertumbuhan penduduk semakin meningkat dari tahun ke tahun sedangkan luas wilayah tidak mengalami peluasan signifikan. Kepadatan penduduk merupakan rasio penduduk dengan luas suatu wilayah administratif. Klasterisasi kepadatan penduduk kabupaten kota di Provinsi Kalimantan Selatan dilakukan untuk mengetahui konsentrasi penduduk suatu wilayah menggunakan metode klasterisasi. *K-Means* merupakan metode klasterisasi yaitu memecah data menjadi beberapa klaster berdasarkan jarak data terdekat terhadap pusat klaster. *Fuzzy C-Means* ialah metode klasterisasi pengembangan dari *K-Means* yaitu mengklasterisasi data ke dalam beberapa klaster berdasarkan fungsi derajat keanggotaan. Data dalam penelitian yaitu kepadatan penduduk wilayah Provinsi Kalimantan Selatan tahun 2003-2019. Metode *elbow* merupakan ukuran menentukan jumlah klaster optimum dengan perhitungan *Sum of Square Error* (SSE) dari tiap klaster untuk menentukan perubahan nilai SSE terbesar antar klaster Berdasarkan hasil penelitian, jumlah klaster optimum adalah 3 karena perubahan nilai SSE sebesar 2, 28. Nilai SSE *Fuzzy C-Means* lebih besar dibanding *K-Means* sehingga metode *K-Means* lebih cocok digunakan dalam penelitian ini. Klasterisasi kepadatan penduduk kabupaten kota di Provinsi Kalimantan Selatan yaitu klaster kepadatan penduduk tinggi, kepadatan penduduk sedang, dan kepadatan penduduk rendah.

Kata kunci : *Klasterisasi, K-Means, Fuzzy C-Means, Kepadatan Penduduk, Elbow*

ABSTRACT

Comparison of K-Means And *Fuzzy C-Means* Methods In Clustering The Population Density Of Urban Districts In South Kalimantan Province (By: Nur Aulia Awwaliyah; Supervisors: Akhmad Yusuf, Yuana Sukmawaty, 2022, 49 pages)

Population growth is increasing from year to year while the area has not experienced significant expansion. Population density is the ratio of population to the area of an administrative area. Clustering of population density of city districts in South Kalimantan Province was carried out to determine the population concentration of an area using the clustering method. K-Means is a clustering method, which breaks data into clusters based on the closest data distance to the cluster center. Fuzzy C-Means is a clustering method developed from K-Means, which is to cluster data into several clusters based on the membership degree function. The data in this study was the population density of South Kalimantan Province in 2003-2019. The elbow method is a measure of determining the optimal number of clusters by calculating the Sum of Square Error (SSE) of each cluster to determine the largest change in SSE value between clusters. Based on the results of the study, the optimum number of clusters was 3 due to the change in SSE value of 2.28. The value of SSE Fuzzy C-Means was greater than K-Means so the K-Means method was more suitable for use in this study. The clustering of population density of urban districts in South Kalimantan Province is a cluster of high population density, medium population density, and low population density.

Keywords : *Clustering, K-Means, Fuzzy C-Means, Population Density, Elbow*

PRAKATA

Alhamdulillah puji dan syukur dipanjatkan atas rahmat dan karunia yang diberikan Allah swt kepada penulis berupa kesehatan, kelancaran, dan kekuatan sehingga mampu menyelesaikan penelitian tepat waktu. Shalawat dan salam tidak lupa selalu dihaturkan kepada Nabi Muhammad SAW, kepada keluarga, sahabat hingga pengikut beliau di akhir zaman.

Penelitian ini berjudul “Komparasi Metode *K-Means* dan *Fuzzy C-Means* Dalam Klasterisasi Kepadatan Penduduk Kabupaten Kota di Provinsi Kalimantan Selatan” telah diselesaikan guna memenuhi syarat kelulusan di Program Studi Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lambung Mangkurat.

Penelitian hingga penulisan skripsi ini tidak akan terjadi tanpa bantuan berbagai pihak. Pada kesempatan ini, penulis ingin mengucapkan terima kasih sebesar-besarnya kepada semua pihak yang terlibat, yaitu :

1. Bapak Drs. Abdul Gafur, M.Si., M.Sc, Ph.d selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lambung Mangkurat .
2. Bapak Pardi Affandi, S.Si., M. Sc. selaku Koordinator Program Studi dan seluruh dosen serta staf Program Studi Matematika.
3. Bapak Akhmad Yusuf S.Si., M.Kom. selaku dosen penasehat akademik sekaligus dosen pembimbing utama bersama Ibu Yuana Sukmawaty S.Si., M.Si. selaku dosen pembimbing pendamping dalam menyelesaikan tugas akhir. Atas segala bimbingan dan arahan yang diberikan kepada penulis sehingga tugas akhir dapat diselesaikan tepat waktu.
4. Bapak Oni Soesanto, S.Si., M.Si dan Bapak Pardi Affandi, S.Si., M.Sc. selaku dosen penguji I dan II yang telah meluangkan waktu untuk memberikan masukan perbaikan terhadap tugas akhir penulis.

5. Kedua orang tua serta keluarga besar yang selalu mendoakan dan memberi dukungan kepada penulis.
6. Keluarga besar Program Studi Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lambung Mangkurat terutama rekan-rekan mahasiswa angkatan 2019 serta seluruh pihak yang tak dapat disebut satu persatu dalam membantu penelitian hingga penulisan tugas akhir selesai dilakukan.

Penulis menyadari dalam penulisan tugas akhir ini masih jauh dari sempurna dan terdapat banyak kekurangan. Kritik dan saran yang membangun untuk tugas akhir ini dengan tujuan perbaikan sangat diperlukan. Semoga tugas akhir ini dapat bermanfaat bagi pembaca.

Banjarbaru, Maret 2023

Nur Aulia Awwaliyah
NIM. 1911011120005

ARTI LAMBANG DAN SINGKATAN

k	: Jumlah kluster
μ_{ik}	: Bilangan acak/ random
ε	: <i>Error</i>
$MaxIter$: Maksimum iterasi
Min	: Jarak terdekat data dengan pusat kluster
P_0	: Fungsi obyektif awal
P_t	: Fungsi obyektif iterasi –t
x_i	: Data atribut
y_j	: Pusat kluster (<i>K-Means</i>)
v_{kj}	: Pusat kluster (<i>Fuzzy C-Means</i>)
w	: Pangkat/ bobot
i	: Setiap data
j	: Periode/ tahun
c	: jarak atribut ke pusat kluster
σ	: Varian
\bar{x}	: Rata-rata
n	: Banyak data
$d(x, y)$: Jarak <i>euclidean</i> x dan y
SSE	: <i>Sum of Squarer Error</i>
$d(x_i, c_k)$: Jarak <i>euclidean</i> x ke pusat kluster
$d(x_{ik}v_{kj})$: Jarak <i>euclidean</i> x ke pusat kluster

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
LEMBAR PENGESAHAN	ii
PERNYATAAN	iii
ABSTRAK	iv
<i>ABSTRACT</i>	v
PRAKATA	vi
ARTI LAMBANG DAN SINGKATAN	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Tujuan Penelitian	2
1.3. Sistematika Penulisan.....	2
BAB II TINJAUAN PUSTAKA	4
2.1. Provinsi Kalimantan Selatan	4
2.2. Kepadatan Penduduk (KP)	4
2.3. CRISP-DM	5
2.4. <i>Outlier</i>	7
2.5. Transformasi Data.....	8
2.6. Normalisasi Data.....	9
2.7. Klasterisasi.....	10
2.7.1. <i>K-Means</i>	11
2.7.2. <i>Fuzzy C-Means</i>	12
2.8. Metode <i>Elbow</i>	14

BAB III METODE PENELITIAN	16
3.1. Sumber Data	16
3.2. Prosedur Penelitian	16
3.2. Diagram Alur Penelitian.....	18
BAB IV HASIL DAN PEMBAHASAN	19
4.1. Hasil	19
4.1.1. Persiapan Data	19
4.1.2. Klasterisasi Dengan Metode <i>K-Means</i>	22
4.1.3. Klasterisasi Dengan Metode <i>Fuzzy C-Means</i>	28
4.1.4. Evaluasi	41
4.2. Pembahasan	47
BAB V PENUTUP	49
5.1. Kesimpulan.....	49
5.2. Saran.....	49
DAFTAR PUSTAKA	50
LAMPIRAN	

DAFTAR TABEL

Tabel	Halaman
1. Data Kepadatan Penduduk Kabupaten Kota di Provinsi Kalimantan Selatan	19
2. Transformasi Data.....	21
3. Normalisasi Data.....	22
4. Iterasi 1	23
5. Pusat Klaster Iterasi 1.....	24
6. Iterasi 2.....	25
7. Pusat Klaster Iterasi 2.....	25
8. Klasterisasi Kepadatan Penduduk Metode <i>K-Means</i>	28
9. Data Observasi.....	29
10. Derajat Keanggotaan Pangkat W	31
11. Perkalian antara Kolom μ_{i1}^w dengan Setiap Kolom Matriks X.....	32
12. Perkalian antara Kolom μ_{i2}^w dengan Setiap Kolom Matriks X.....	33
13. Perkalian antara Kolom μ_{i3}^w dengan Setiap Kolom Matriks X.....	34
14. Pusat Klaster Iterasi 1.....	34
15. Hasil Perhitungan C_1	36
16. Hasil Perhitungan C_2	37
17. Hasil Perhitungan C_3	37
18. Hasil Perhitungan Fungsi Objektif (P_1)	38
19. Hasil Klasterisasi Metode <i>Fuzzy C-Means</i> Tahun 2012-2013	40
20. Klasterisasi Kepadatan Penduduk Metode <i>Fuzzy C-Means</i>	41
21. Perhitungan SSE pada 3 Klaster Tahun 2012-2013 <i>K-Means</i>	42
22. Perhitungan SSE Jumlah Klaster 1-8 <i>K-Means</i>	43
23. Perhitungan SSE pada 3 Klaster Tahun 2012-2013 <i>FCM</i>	45
24. Perhitungan SSE Jumlah Klaster 1-8 <i>Fuzzy C-Means</i>	45

25. Klasterisasi Kepadatan Penduduk Kabupaten Kota di Provinsi Kalimantan Selatan	47
---	----

DAFTAR GAMBAR

Gambar	Halaman
1. Tahapan dalam CRISP-DM.....	7
2. <i>K-Means</i>	11
3. <i>Flowchart</i> Metode <i>K-Means</i> dan <i>Fuzzy C-Means</i>	18
4. Peta kepadatan penduduk	20
5. Klasterisasi Dengan RapidMiner Studio Versi 10.0	27
6. Jumlah klaster 3	27
7. Metode <i>Elbow K-Means</i>	43
8. Metode <i>Elbow FCM</i>	46

DAFTAR LAMPIRAN

Lampiran

1. Jumlah Penduduk Kabupaten Kota di Provinsi Kalimantan Selatan
2. Luas Wilayah Kabupaten Kota di Provinsi Kalimantan Selatan
3. Kepadatan Penduduk Kabupaten Kota di Provinsi Kalimantan Selatan
4. Transformasi Data Kepadatan Penduduk Kabupaten Kota di Provinsi Kalimantan Selatan Tahun 2003-2019
5. Normalisasi Data Kepadatan Penduduk Kabupaten Kota di Provinsi Kalimantan Selatan Tahun 2003-2019
6. Pusat Klaster Untuk Jumlah Klaster 1-8 Metode *K-Means*
7. Derajat Keanggotaan Untuk Jumlah Klaster 1-8 Metode *Fuzzy C-Means*
8. Input *Python Fuzzy C-Means* Untuk Klasterisasi