



**PENENTUAN KLASTER OPTIMAL PADA ALGORITMA K-MEANS
MENGUNAKAN QUANTITATIVE DISCRIMINANT METHODS**

SKRIPSI

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

Oleh:

RAIHANA ZULFAH

NIM. 1711011320013

**PROGRAM STUDI MATEMATIKA
FAKULTAS MATEMATIKADAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU
2024**

SKRIPSI

PENENTUAN KLASTER OPTIMAL PADA ALGORITMA K-MEANS MENGUNAKAN QUANTITATIVE DISCRIMINANT METHODS

Oleh:
RAIHANA ZULFAH
1711011320013



telah dipertahankan di depan Dosen Penguji pada tanggal 5 Juli 2024
Susunan Dosen Penguji:

Pembimbing I



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

Dosen Penguji:

1. Oni Soesanto, S.Si., M.Si ()
2. Saman Abdurrahman, S.Si, M.Sc ()

Pembimbing II



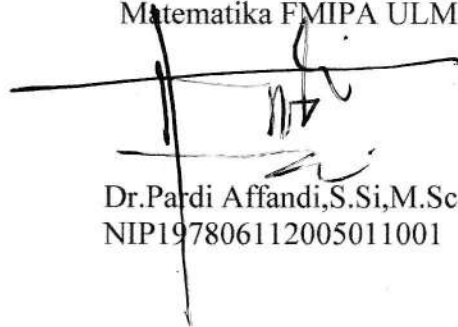
Drs. Faisal, S.Si
NIP.196309021992031001

Banjarbaru, 8 Juli 2024
Wakil Dekan Bidang Akademik



Dr. Gunawan, S.Si., M.Si
NIP 197911012005011002

Koordinator Program Studi
Matematika FMIPA ULM



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

SKRIPSI

**PENENTUAN KLASTER OPTIMAL PADA ALGORITMA K-MEANS
MENGUNAKAN QUANTITATIVE DISCRIMINANT METHODS**

Oleh:
RAIHANA ZULFAH
1711011320013



telah dipertahankan di depan Dosen Penguji pada tanggal 5 Juli 2024
Susunan Dosen Penguji:

Pembimbing Utama



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

Dosen Penguji:

1. Oni Soesanto, S.Si., M.Si ()
2. Saman Abdurrahman, S.Si, M.Sc ()

Pembimbing Pendamping



Drs. Faisal, S.Si
NIP.196309021992031001

Banjarbaru, 05 Juli 2024
Program Studi Matematika FMIPA ULM
Ketua



Dr. Pardi Affandi, S.Si, M.Sc
NIP197806112005011001

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 orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam Daftar Pustaka.

Banjarbaru,

Raihana Zulfah

NIM. 1711011320013

ABSTRAK

PENENTUAN KLASTER OPTIMAL PADA ALGORITMA K-MEANS MMENGGUNAKAN QUANTITATIVE DISCRIMINANT METHODS

(Oleh : Raihana Zulfah ; Pembimbing: Akhmad Yusuf, Faisal,2024; 44 halaman)

Provinsi Kalimantan Selatan termasuk salah satu provinsi yang menyokong pangan tingkat Nasional hal ini dapat dilihat dari hasil produksi padi yang besar menyebar di 13 kabupaten Kota. Untuk mendukung ketahanan pangan maka perlu dilakukan klasterisasi produki padi di Kalimantan Selatan. Maka Untuk menentukan jumlah klaster yang optimal sering kali menggunakan metode elbow akan tetapi metode elbow memiliki kelemahan karena penentuannya dapat menjadi multitafsir. Tujuan penelitian ini adalah mendiskripsikan distribusi produksi padi, menerapkan metode k-means untuk menentukan klasterisasi dan menentukan jumlah klaster yang optimal. Metode penelitian yang digunakan yaitu K-Means dalam menyelesaikan proses klasterisasi dan quantitative untuk menentukan klaster yang optimal. Penentuan titik klaster dilakukan secara acak dan di lakukan beberapa kali iterasi hingga stabil. Dengan menggunakan Aquantitative didapatkan bahwa klaster yang optimal berada pada 5 klaster yaitu klaster 1 Kabupaten Kotabaru, Hulu Sungai Utara, Tabalong , Tanah Bumbu dan Balangan. Klaster 2 Kabupaten Banjar , Tapin dan Hulu Sungai Tengah , Klaster 3 Kota Banjarmasin dan Banjarbaru , Klaster 4 terdiri dari Barito Kuala dan Klaster 5 Kabupaten Tanah laut dan Hulu Sungai Selatan.

Kata kunci: Klasterisasi, K-Means, Elbow , *Quantitative* ,Hasil produksi padi

ABSTARCT

DETERMINING OPTIMAL CLUSTERS IN THE K-MEANS ALGORITHM USING QUANTITATIVE DISCRIMINANT METHODS (By : Raihana Zulfah ; Supervisor : Akhmad Yusuf, Faisal,2024; 44 pages)

South Kalimantan Province is one of the provinces that supports food at the national level. This can be seen from the large rice production spread across 13 city/districts. To support food security, it is necessary to cluster rice production in South Kalimantan. So, to determine the optimal number of clusters, the elbow method is often used, but the elbow method has a weakness because the determination can be subject to multiple interpretations. The aim of this research was to describe the distribution of rice production, apply the k-means method to determine clustering and determine the optimal number of clusters. The research method used was K-Means to complete the clustering process and quantitative to determine the optimal cluster. Determination of cluster points was carried out randomly and in several iterations until stable. By using Aquantitative, it was found that the optimal clusters were in 5 clusters, namely Kotabaru Regency, Hulu Sungai Utara, Tabalong, Tanah Bumbu, and Balangan for cluster 1. Cluster 2 consisted of Banjar, Tapin, and Hulu Sungai Tengah Regencies, Cluster 3 consisted of Banjarmasin and Banjarbaru Cities, Cluster 4 consisted of Barito Kuala, and Cluster 5 consisted of Tanah Laut and Hulu Sungai Selatan Regencies.

Keywords: Clusterization, K-Means, Elbow , Quantitative ,Rice production results.

PRAKATA

Assalamualaikum. Wr. Wb

Puji syukur kehadiran Allah SWT, atas segala rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan penyusunan skripsi ini yang berjudul. **PENENTUAN KLASSTER OPTIMAL PADA ALGORITMA K-MEANS MMENGGUNAKAN QUANTITATIVE DISCRIMINANT METHODS.** Sholawat serta salam semoga selalu tercurah kepada junjungan Nabi Muhammad SAW yang telah menyampaikan risalah dan wahyu dari Allah SWT untuk menuju kebahagiaan hidup dunia dan akhirat.

Banyak pihak yang telah membantu penulis dalam penyusunan hingga terwujudnya skripsi ini. Oleh karena itu, dengan segala kerendahan hati penulis sampaikan ungkapan terima kasih dan penghargaan yang tulus kepada:

1. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lambung Mangkurat Banjarbaru.
2. Ketua Program Studi Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lambung Mangkurat Banjarbaru.
3. Bapak Akhmad Yusuf, S.Si.,M.Kom. selaku dosen pembimbing pertama skripsi yang telah bersedia meluangkan waktu, memberikan bantuan, bimbingan, dan motivasi yang sangat berharga dalam pelaksanaan penelitian serta penyusunan skripsi ini.
4. Bapak Drs. Faisal, M.Si. selaku dosen pembimbing kedua skripsi yang telah bersedia meluangkan waktu, memberikan bantuan, bimbingan, dan motivasi yang sangat berharga dalam pelaksanaan penelitian serta penyusunan skripsi ini.
5. Bapak Saman Abdurrahman, S.Si., M.Sc selaku dosen pembimbing akademik yang telah memberikan bimbingan akademik berupa saran dan motivasi selama penulis kuliah.

6. Dosen-dosen pengajar Program Studi Matematika atas bantuan dan bimbingan, serta kepercayaan dan motivasi yang besar dalam pelaksanaan penelitian dan penyusunan skripsi.
7. Orangtua dan keluarga tercinta yang selalu ada untuk menemani, memberi dukungan, motivasi, kasih sayang, pengertian dan doa yang tiada henti untuk kelancaran semuanya.
8. Seluruh teman terdekat dan rekan mahasiswa Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lambung Mangkurat Banjarbaru. Kepada teman-teman Calculus 2017, serta seluruh pihak yang telah memberikan bantuan, baik berupa masukan, saran, maupun nasehat kepada penulis selama proses penulisan skripsi ini.
9. Terakhir, untuk diri sendiri yang sudah bertahan dan berjuang sejauh ini demi menyelesaikan pendidikan. Tidak apa-apa terlambat, dari pada tidak sama sekali.

Semoga segala kebaikan kalian mendapat balasan yang setimpal dari Allah SWT. Penulis sadar bahwa tulisan ini masih jauh dari sempurna. Oleh karena itu, kritik dan saran selalu penulis harapkan demi perbaikan tulisan ini. Akhir kata, penulis berharap semoga skripsi ini dapat bermanfaat bagi para pembaca.

Banjarbaru, Juli 2024

Raihana Zulfah

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
PERNYATAAN	iii
ABSTRAK	iv
<i>ABSTRACT</i>	v
PRAKATA	vi
DAFTAR ISI	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xii
ARTI LAMBANG DAN SINGKATAN	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Tujuan Penelitian	3
1.3 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA	4
2.1 Produksi Padi	4
2.2 Luas Lahan Panen	5
2.3 Hubungan Luas Lahan panen Dengan Produksi Padi	5
2.4 Data Mining	6
2.5 CRISP-DM.....	8
2.6 Normalisa Data	10
2.7 Klasterisasi	10
2.8 K-Means.....	11
2.9 Metode Elbow	13
2.10 Metode Quantitative Discriminant	14

BAB III PROSEDUR PENELITIAN	18
3.1 Sumber data	18
3.2 Prosedur Penelitian.....	18
3.3 Diagram Alur Penelitian	19
BAB IV HASIL DAN PEMBAHASAN	21
4.1 Deskripsi Data Hasil Panen	21
4.2 Persiapan Data	26
4.3 Normalisasi Data	27
4.4 Klasterisasi Dengan Metode K-Means	28
4.5 Evaluasi (<i>Evaluation</i>)	34
4.5.1 Elbow	34
4.5.2 Metode Quantitative Discriminant.....	36
BAB V PENUTUP	40
5.1 Kesimpulan.....	40
5.2 Saran	40
DAFTAR PUSTAKA	41
LAMPIRAN	

DAFTAR TABEL

Tabel

Halaman

1. Data Hasil Produksi Padi di Provinsi Kalimantan Selatan	21
2. Data Luas Panen Padi di Provisi Kalimantan Selatan.....	24
3. Data Hasil Poduksi Padi Kabupaten Kota di Provinsi Kallimantan Selatan...27	
4. Normalisasi Data	28
5. Iterasi 1	29
6. Pusat Klaster Iterasi 1.....	30
7. Iterasi	
2.....	31
8. Pusat Klaster Iterasi 2.....	31
9. Klaterisasi hasil Produksi padi Provinsi Kalimantan Selatan	33
10. Perhitungan SSE pada 2 klaster tahun 2016-2017.....	34
11. Perhitungan SSE pada klaster 2-13.....	35
12. Perhitungan MQD hasil produksi padi di Privinsi Kalimantan Selatan.....	39

DAFTAR GAMBAR

Gambar	Halaman
1. $K=2$	33
2. Metode <i>Elbow</i>	35
3. MQD	39

DAFTAR LAMPIRAN

Lampiran

1. Data hasil Produksi Padi Provinsi Kalimantan Selatan
2. Data Luas Panen Padi Provinsi Kalimantan Selatan
3. Hasil Normalisasi data hasil Produksi padi Provinsi Kalimantan Selatan
4. Klasterisasi degan aplikasi Rapidminer Studio Versi 10.0

ARTI LAMBANG DAN SINGKATAN

x_{new}	: Data yang sudah dinormalisasi
x	: Data yang akan dinormalisasi
x_{min}	: Nilai terkecil dalam data
x_{max}	: Nilai terbesar dalam data
d	: Jarak antara x dan y
x_i	: data pada setiap i
y_i	: Data pada pusat klaster ke i
n	: jumlah data
K	: Jumlah Klaster
x_i	: Data ke i
y_k	: Pusat klaster
SSE	: Squared Euclidean Distance
MQD	: Metode Quantitative Discriminant
MD	: Mean distorsi
MD_i	: Nilai Mean distorsi
N	: Jumlah Klaster
n_i	: Nilai MD ke- i
n_k	: Nilai MD ke- k
k_i	: Klaster ke- i
k_j	: Klaster ke- j
α_j	: Nilai titik siku
E_{ij}	: Nilai Euclidean E_{ij}
E_{ik}	: Nilai Euclidean E_{ik}
E_{jk}	: Nilai Euclidean E_{jk}