

**OPTIMALISASI ALGORITMA *FUZZY GUSTAFSON-KESSEL*
DENGAN VARIASI PARAMETER *WEIGHTING EXPONENT*
UNTUK *CLUSTERING* DATA RISIKO DEPRESI PELAJAR**

SKRIPSI

**OLEH
AMIRA ADELIA PUTRI
NIM. 220601110051**



**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM MALANG
2026**

**OPTIMALISASI ALGORITMA *FUZZY GUSTAFSON-KESSEL*
DENGAN VARIASI PARAMETER *WEIGHTING EXPONENT*
UNTUK *CLUSTERING* DATA RISIKO DEPRESI PELAJAR**

SKRIPSI

**Diajukan Kepada
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk Memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Matematika (S.Mat)**

**Oleh
Amira Adelia Putri
NIM. 220601110051**

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM MALANG
2026**

**OPTIMALISASI ALGORITMA *FUZZY GUSTAFSON-KESSEL*
DENGAN VARIASI PARAMETER *WEIGHTING EXPONENT*
UNTUK *CLUSTERING* DATA RISIKO DEPRESI PELAJAR**

SKRIPSI

Oleh
Amira Adelia Putri
NIM. 220601110051

Telah Disetujui untuk Diuji

Malang, 20 Februari 2026

Pembimbing I



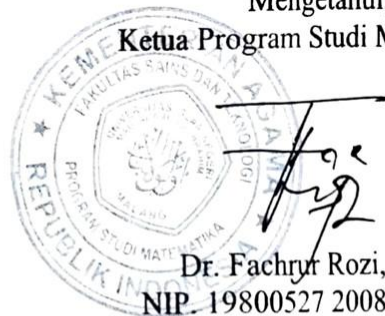
Intan Nisfulaila, M.Si
NIP. 19900215 201903 2 015

Pembimbing II



Erna Herawati, M.Pd
NIPPPK. 1976072320232120006

Mengetahui,
Ketua Program Studi Matematika



Dr. Fachrut Rozi, M.Si
NIP. 19800527 200801 1 012

**OPTIMALISASI ALGORITMA *FUZZY GUSTAFSON-KESSEL*
DENGAN VARIASI PARAMETER *WEIGHTING EXPONENT*
UNTUK *CLUSTERING* DATA RISIKO DEPRESI PELAJAR**

SKRIPSI

**Oleh
Amira Adelia Putri
NIM. 220601110051**

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima sebagai Salah Satu Persyaratan
untuk Memperoleh Gelar Sarjana Matematika (S.Mat)

Tanggal 27 April 2026

Ketua Penguji : Prof. Dr. H. Turmudi, M.Si, Ph.D

Anggota Penguji 1 : Mohammad Nafie Jauhari, M.Si

Anggota Penguji 2 : Intan Nisfulaila, M.Si

Anggota Penguji 3 : Erna Herawati, M.Pd

Mengetahui,
Ketua Program Studi Matematika



Dr. Fachrur Rozi, M.Si
NIP. 19800527 200801 1 012



PERNYATAAN KEASLIAN TULISAN

Saya bertanda tangan di bawah ini

Nama : Amira Adelia Putri

NIM : 220601110051

Program Studi : Matematika

Fakultas : Sains dan Teknologi

Judul Skripsi : Optimalisasi Algoritma *Fuzzy Gustafson-Kessel* dengan
Variasi Parameter *Weighting Exponent* untuk *Clustering Data*
Risiko Depresi Pelajar

Menyatakan bahwa skripsi yang saya susun merupakan hasil karya sendiri dan tidak mengambil atau menyalin tulisan maupun pemikiran orang lain. Seluruh isi skripsi ini adalah hasil pemikiran saya, kecuali bagian yang telah disebutkan sumbernya dalam daftar rujukan pada halaman terakhir. Apabila di kemudian hari terbukti bahwa skripsi ini merupakan hasil plagiasi atau meniru karya orang lain, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 27 April 2026



Amira Adelia Putri
NIM. 220601110051

MOTO

“Manusia tidak akan memperoleh selain dari apa yang diusahakannya. Usahamu akan diperlihatkan, lalu dibalas dengan balasan yang sempurna”

(Q.S An-Najm: 39-41)

PERSEMBAHAN

Skripsi ini penulis persembahkan kepada:

Terkhusus untuk Alm. Bapak, sosok yang paling dirindukan, yang semasa hidupnya telah memberikan kasih sayang, doa, serta nasihat yang menjadi kekuatan bagi penulis hingga sampai pada titik ini. Meskipun tidak lagi bersama, pengorbanan beliau selalu menjadi semangat dalam setiap langkah penulis. Ibu tercinta, yang senantiasa memberikan kasih sayang, doa, dan dukungan tanpa henti, yang dengan segala ketulusan dan kekuatannya telah mengantarkan penulis sampai pada pencapaian ini.

Kakak dan keluarga besar, yang selalu hadir dengan dukungan dan perhatian yang menjadi bagian penting dalam perjalanan penulis hingga mampu menyelesaikan skripsi ini.

KATA PENGANTAR

Puji syukur ke hadirat Allah SWT atas limpahan rahmat, taufik, dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi ini yang berjudul “Optimalisasi Algoritma *Fuzzy Gustafson-Kessel* dengan Variasi Parameter *Weighting Exponent* untuk *Clustering* Data Risiko Depresi Pelajar” sebagai salah satu syarat untuk memperoleh gelar sarjana pada Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Penulis menyadari bahwa penyusunan skripsi ini tidak terlepas dari bantuan, bimbingan, serta dukungan dari berbagai pihak, baik secara langsung maupun tidak langsung. Oleh karena itu, penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Prof. Dr. Hj. Ilfi Nur Diana, M.Si., CAHRM., CRMP selaku Rektor Universitas Islam Negeri Maulana Malik Ibrahim.
2. Dr. Agus Mulyono, M.Kes, selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim.
3. Dr. Fachrur Rozi, M.Si, selaku Ketua Program Studi Matematika, Universitas Islam Negeri Maulana Malik Ibrahim.
4. Intan Nisfulaila, M.Si, selaku Dosen Pembimbing I, yang telah memberikan bimbingan, arahan, serta masukan yang sangat berharga selama proses penyusunan skripsi ini.
5. Erna Herawati, M.Pd, selaku Dosen Pembimbing II, yang telah memberikan bimbingan dan arahan dalam pengintegrasian nilai-nilai keislaman pada penelitian ini.
6. Prof. Dr. H. Turmudi, M.Si, Ph.D, selaku Ketua Penguji dalam Ujian Skripsi, yang telah berkenan memberikan arahan serta saran yang membangun bagi penyusunan skripsi.
7. Mohammad Nafie Jauhari, M.Si, selaku Anggota Penguji I dalam Ujian Skripsi, yang telah memberikan arahan dan saran yang sangat mendukung dalam penyusunan skripsi ini.
8. Seluruh dosen Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim.

9. Kedua orang tua tercinta, Alm. Bapak yang semasa hidupnya dengan penuh perhatian telah memberikan bimbingan dan nasihat kepada penulis untuk terus melanjutkan pendidikan, serta Ibu yang telah memberikan kasih sayang, doa, dukungan, serta pengorbanan yang besar sehingga penulis dapat menyelesaikan pendidikan hingga tahap ini.
10. Kakak Penulis Devi Irawati dan seluruh keluarga besar penulis yang senantiasa memberikan doa, dukungan, dan bantuan kepada penulis hingga dapat menyelesaikan pendidikan ini.
11. Teman-teman terdekat penulis, Do'aul Isma, Ayu Habibatun, Nurus Syafi'ah, Rossima Eva, dan Mutiara Alfi, yang senantiasa memberikan dukungan, semangat, serta motivasi kepada penulis dalam menyelesaikan skripsi ini. Terima kasih atas kebersamaan, bantuan, serta waktu yang telah diberikan, baik dalam suka maupun duka, sehingga penulis dapat melalui setiap proses penyusunan skripsi ini dengan lebih baik.
12. Seluruh Mahasiswa Prodi Matematika angkatan 2022 yang telah memberikan bantuan dan kerja samanya selama menjalani perkuliahan ini.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna, baik dari segi penyajian maupun substansi pembahasan. Oleh karena itu, penulis dengan terbuka mengharapkan kritik dan saran yang bersifat membangun dari berbagai pihak agar penelitian ini dapat disempurnakan dan memberikan manfaat yang lebih optimal.

Malang, 27 April 2026

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGANTAR	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
PERNYATAAN KEASLIAN TULISAN	v
MOTO	vi
PERSEMBAHAN	vii
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR SIMBOL	xiv
DAFTAR LAMPIRAN	xv
ABSTRAK	xvi
ABSTRACT	xvii
مستخلص البحث	xviii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	6
1.3 Tujuan Penelitian	6
1.4 Manfaat Penelitian	6
1.5 Batasan Masalah	7
1.6 Definisi Istilah	7
BAB II KAJIAN TEORI	9
2.1 Logika <i>Fuzzy</i>	9
2.2 <i>Clustering</i>	10
2.3 <i>Fuzzy Clustering</i>	12
2.4 <i>Weighting Exponent (m)</i>	15
2.5 <i>Fuzzy Gustafson-Kessel (FGK)</i>	16
2.6 <i>Xie-Beni Index</i>	23
2.7 <i>Preprocessing Data</i>	26
2.7.1 <i>Label Encoding</i>	26
2.7.2 <i>Normalisasi Min-Max</i>	27
2.8 <i>Principal Component Analysis (PCA)</i>	28
2.9 Depresi	29
2.10 Kajian Integrasi Topik dengan Al-Qur'an	31
2.11 Kajian Topik dengan Teori Pendukung	34
BAB III METODE PENELITIAN	37
3.1 Jenis Penelitian	37
3.2 Data dan Sumber Data	37
3.3 Teknik Analisis Data	38
BAB IV HASIL DAN PEMBAHASAN	43
4.1 Analisis Deskriptif Data Risiko Depresi Pelajar	43
4.2 <i>Preprocessing Data Risiko Depresi Pelajar</i>	45
4.2.1 <i>Label Encoding</i>	46
4.2.2 <i>Normalisasi Data</i>	47
4.3 Implementasi Algoritma <i>Fuzzy Gustafson-Kessel</i>	50

4.3.1 Perhitungan untuk Variasi Nilai $m = 1,1$	50
4.3.2 Perhitungan untuk Variasi Nilai $m = 1,3$	64
4.3.3 Perhitungan untuk Variasi Nilai $m = 1,5$	71
4.3.4 Perhitungan untuk Variasi Nilai $m = 1,7$	77
4.3.5 Perhitungan untuk Variasi Nilai $m = 1,9$	84
4.3.6 Perhitungan untuk Variasi Nilai $m = 2,1$	90
4.4 Evaluasi Hasil <i>Clustering</i> dengan <i>Xie-Beni Index</i>	97
4.5 Visualisasi dan Interpretasi Hasil <i>Clustering</i> dengan Parameter Optimal.....	104
4.6 Kajian Integrasi Al-Qur'an terhadap Risiko Depresi Pelajar.....	107
BAB V PENUTUP	111
5.1 Kesimpulan	111
5.2 Saran untuk Penelitian Lanjutan	111
DAFTAR PUSTAKA	113
LAMPIRAN	117
RIWAYAT HIDUP	124

DAFTAR TABEL

Tabel 3.1	Variabel <i>Dataset</i>	38
Tabel 4.1	Analisis Deskriptif Variabel Numerik.....	43
Tabel 4.2	Analisis Deskriptif Variabel Kategori	44
Tabel 4.3	Tampilan <i>Head Dataset</i>	45
Tabel 4.4	Skema Label <i>Encoding</i>	47
Tabel 4.5	Contoh <i>Dataset</i> setelah Label <i>Encoding</i>	47
Tabel 4.6	Contoh <i>Dataset</i> setelah Normalisasi <i>Min-Max</i>	50
Tabel 4.7	Hasil Pusat Dua <i>Cluster</i> ($m = 1,1$).....	54
Tabel 4.8	Hasil Jarak <i>Mahalanobis</i> Dua <i>Cluster</i> ($m = 1,1$).....	59
Tabel 4.9	Hasil Matriks Keanggotaan Baru dengan Dua <i>Cluster</i> ($m = 1,1$)	62
Tabel 4.10	Hasil <i>Clustering</i> Dua <i>Cluster</i> ($m = 1,1$).....	64
Tabel 4.11	Hasil Pusat Dua <i>Cluster</i> ($m = 1,3$).....	66
Tabel 4.12	Hasil Jarak <i>Mahalanobis</i> Dua <i>Cluster</i> ($m = 1,3$).....	67
Tabel 4.13	Hasil Matriks Keanggotaan Baru dengan Dua <i>Cluster</i> ($m = 1,3$)	69
Tabel 4.14	Hasil <i>Clustering</i> Dua <i>Cluster</i> ($m = 1,3$).....	71
Tabel 4.15	Hasil Pusat Dua <i>Cluster</i> ($m = 1,5$).....	73
Tabel 4.16	Hasil Jarak <i>Mahalanobis</i> Dua <i>Cluster</i> ($m = 1,5$).....	74
Tabel 4.17	Hasil Matriks Keanggotaan Baru dengan Dua <i>Cluster</i> ($m = 1,5$)	75
Tabel 4.18	Hasil <i>Clustering</i> Dua <i>Cluster</i> ($m = 1,5$).....	77
Tabel 4.19	Hasil Pusat Dua <i>Cluster</i> ($m = 1,7$).....	79
Tabel 4.20	Hasil Jarak <i>Mahalanobis</i> Dua <i>Cluster</i> ($m = 1,7$).....	80
Tabel 4.21	Hasil Matriks Keanggotaan Baru dengan Dua <i>Cluster</i> ($m = 1,7$)	82
Tabel 4.22	Hasil <i>Clustering</i> Dua <i>Cluster</i> ($m = 1,7$).....	83
Tabel 4.23	Hasil Pusat Dua <i>Cluster</i> ($m = 1,9$).....	85
Tabel 4.24	Hasil Jarak <i>Mahalanobis</i> Dua <i>Cluster</i> ($m = 1,9$).....	87
Tabel 4.25	Hasil Matriks Keanggotaan Baru dengan Dua <i>Cluster</i> ($m = 1,9$)	89
Tabel 4.26	Hasil <i>Clustering</i> Dua <i>Cluster</i> ($m = 1,9$).....	90
Tabel 4.27	Hasil Pusat Dua <i>Cluster</i> ($m = 2,1$).....	92
Tabel 4.28	Hasil Jarak <i>Mahalanobis</i> Dua <i>Cluster</i> ($m = 2,1$).....	93
Tabel 4.29	Hasil Matriks Keanggotaan Baru dengan Dua <i>Cluster</i> ($m = 2,1$)	95
Tabel 4.30	Hasil <i>Clustering</i> Dua <i>Cluster</i> ($m = 2,1$).....	97
Tabel 4.31	Matriks Keanggotaan untuk Evaluasi <i>Xie-Beni Index</i>	98
Tabel 4.32	Pusat <i>Cluster</i> untuk Evaluasi <i>Xie-Beni Index</i>	99
Tabel 4.33	Hasil Evaluasi <i>Clustering</i> dengan <i>Xie-Beni Index</i>	102
Tabel 4.34	Nilai Rata-rata Variabel Tiap <i>Cluster</i>	105

DAFTAR GAMBAR

Gambar 3.1 Diagram Alir Penelitian.....	41
Gambar 4.1 Grafik <i>Xie-Beni Index</i> terhadap Variasi <i>Weighting Exponent</i>	103
Gambar 4.2 Visualisasi Hasil <i>Clustering</i>	104

DAFTAR SIMBOL

X_1	: Variabel Umur
X_2	: Variabel Tekanan Akademik
X_3	: Variabel Kepuasan Belajar
X_4	: Variabel Jam Belajar
X_5	: Variabel Stres Finansial
X_6	: Variabel Durasi Tidur
X_7	: Variabel Pola Makan
X	: Himpunan data
$\mathbb{R}^s$	: Ruang berdimensi s
s	: Dimensi data
c	: Banyaknya <i>cluster</i>
n	: Jumlah data
m	: <i>Weighting exponent</i> (Pangkat pembobot)
t	: Indeks iterasi
x_i	: Data ke- i
x_j	: Data ke- j
x_{min}	: Nilai minimum data
x_{max}	: Nilai maksimum data
x'	: Nilai data normalisasi
$\mu_i(x)$	: Derajat keanggotaan data terhadap <i>cluster</i> ke- i
$\mu_{i,k}$	: Derajat keanggotaan data ke- i pada <i>cluster</i> ke- k
$\mu_{i,k}^{(t)}$	: Derajat keanggotaan pada iterasi ke- t
U	: Matriks keanggotaan
U_t	: Matriks keanggotaan pada iterasi ke- t
U_{t+1}	: Matriks keanggotaan hasil pembaruan
ε	: Ambang batas konvergensi
v_k	: Vektor pusat <i>cluster</i> ke- k
$v_{k,j}$	: Pusat <i>cluster</i> ke- k pada variabel ke- j
p	: Jumlah variabel
j	: Indeks variabel
F_k	: Matriks <i>fuzzy covariance cluster</i> ke- k
A_k	: <i>Norm inducing matrix cluster</i> ke- k
ρ_k	: Volume <i>cluster</i> ke- k
$D_{i,k,Ak}^2$	: Jarak <i>Mahalanobis</i> data ke- i terhadap <i>cluster</i> ke- k
$J_R(U)$	: Fungsi objektif <i>fuzzy clustering</i>
$J_{GJ}^{(0)}$	: Fungsi objektif awal
$J_{GK}^{(t)}$	: Fungsi objektif pada iterasi ke- t
v_{XB}	: Nilai <i>Xie-Beni Index</i>
S	: Tingkat kekompakan internal <i>cluster</i>
d_{min}	: Jarak minimum antar pusat <i>cluster</i>
q	: Indeks pembandingan jarak antar <i>cluster</i>
$d_{i,k}^2(x_i, v_k)$	: Jarak data ke- i terhadap pusat <i>cluster</i> ke- k
$d_{i,k}^2(v_i, v_k)$	: Jarak antar pusat <i>cluster</i>

DAFTAR LAMPIRAN

Lampiran 1	<i>Dataset Penelitian</i>	117
Lampiran 2	<i>Dataset setelah Preprocessing</i>	117
Lampiran 3	Perhitungan Hasil <i>Clustering</i> dengan $m = 1,1$	117
Lampiran 4	Perhitungan Hasil <i>Clustering</i> dengan $m = 1,3$	117
Lampiran 5	Perhitungan Hasil <i>Clustering</i> dengan $m = 1,5$	117
Lampiran 6	Perhitungan Hasil <i>Clustering</i> dengan $m = 1,7$	117
Lampiran 7	Perhitungan Hasil <i>Clustering</i> dengan $m = 1,9$	117
Lampiran 8	Perhitungan Hasil <i>Clustering</i> dengan $m = 2,1$	118
Lampiran 9	<i>Source Code</i> Algoritma <i>Fuzzy Gustafson-Kessel</i>	118
Lampiran 10	<i>Source Code</i> <i>Xie-Beni Index</i>	121
Lampiran 11	<i>Source Code</i> Visualisasi Hasil <i>Clustering</i> dengan <i>Principal Component Analysis (PCA)</i>	122

ABSTRAK

Putri, Amira Adelia. 2026. **Optimalisasi Algoritma *Fuzzy Gustafson-Kessel* dengan Variasi Parameter *Weighting Exponent* untuk *Clustering* Data Risiko Depresi Pelajar**. Skripsi. Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Intan Nisfulaila, M.Si. (II) Erna Herawati, M.Pd.

Kata Kunci: *Fuzzy Gustafson-Kessel*, *Weighting Exponent*, *Clustering*, *Xie-Beni Index*, Depresi Pelajar.

Clustering digunakan untuk mengelompokkan data berdasarkan kesamaan karakteristik, namun metode konvensional memiliki keterbatasan dalam menangani kompleksitas bentuk data. Algoritma *Fuzzy Gustafson-Kessel* (FGK) sebagai metode *fuzzy clustering* mampu menangani bentuk *cluster* yang beragam, tetapi kinerjanya dipengaruhi oleh pemilihan parameter *weighting exponent*. Penelitian ini bertujuan untuk mengoptimalkan parameter *weighting exponent* pada algoritma FGK dalam menghasilkan struktur *cluster* yang lebih mencerminkan karakteristik data. Penelitian ini menggunakan pendekatan kuantitatif deskriptif dengan data sekunder dari *Depression Student Dataset*. Analisis dilakukan melalui implementasi algoritma FGK dan evaluasi *cluster* menggunakan *Xie-Beni Index* pada berbagai variasi parameter dan banyaknya *cluster*. Hasil penelitian menunjukkan bahwa kualitas *clustering* mencapai kondisi paling optimal pada penggunaan nilai *weighting exponent* sebesar 1,1 dengan banyaknya *cluster* sebanyak tiga. Nilai tersebut memberikan performa terbaik dibandingkan variasi parameter lainnya. Peningkatan nilai parameter menyebabkan penurunan kualitas *cluster* akibat tingkat kekaburan yang lebih tinggi. Hasil *clustering* mengelompokkan data ke dalam tiga tingkat risiko depresi, yaitu rendah, sedang, dan tinggi.

ABSTRACT

Putri, Amira Adelia. 2026. **Optimisation of Fuzzy Gustafson-Kessel with Variations in the Weighting Exponent Parameter for Clustering Data on Students' Risk of Depression.** Undergraduate Thesis. Department of Mathematics, Faculty of Science and Technology, Universitas Islam Negeri Maulana Malik Ibrahim, Malang. Supervisors: (I) Intan Nisfulaila, M.Si. (II) Erna Herawati, M.Pd.

Keywords: Fuzzy Gustafson-Kessel, Weighting Exponent, Clustering, Xie-Beni Index, Student Depression.

Clustering is used to group data based on similar characteristics; however, conventional methods have limitations in handling the complexity of data structures. The Fuzzy Gustafson-Kessel (FGK) algorithm, as a fuzzy clustering method, is capable of handling diverse cluster structures, but its performance is influenced by the choice of the weighting exponent parameter. This study aims to optimise the weighting exponent parameter in the FGK algorithm to produce a cluster structure that better reflects the data characteristics. This study employs a descriptive quantitative approach using secondary data from the Depression Student Dataset. Analysis was conducted through the implementation of the FGK algorithm and cluster evaluation using the Xie-Beni Index across various parameter variations and the number of clusters. The results indicate that the optimal clustering quality was achieved at a weighting exponent value of 1.1 with three clusters, yielding the best performance among all tested parameter variations. Increasing the parameter value leads to a decline in clustering quality due to higher levels of ambiguity. The clustering results group the data into three levels of depression risk: low, moderate, and high.

مستخلص البحث

بوتري، أميرة أديلية ٢٠٢٦. تحسن خوارزمية *Fuzzy Gustafson-Kessel* اضبابية مع تغيير معامل *Weighting Exponent* لتجميع بيانات مخاطر الاكتتاب لدى الطلاب. البحث الجامعي. قسم الرياضيات، كلية العلوم والتكنولوجيا، جامعة مولان ملك إبراهيم الإسلامية الحكومية ملانج. المشرفة: (١) إينتان نسف الليلة، الماجستير في العلوم. (٢) إيرنا هيراواي، الماجستير في التربية

الكلمات الأساسية: اكتتاب الطلاب، التجميع، *Weighting Exponent*، *Fuzzy Gustafson Kessel*، *Xie-Beni Index*،

استخدم التجميع لتصنيف البيانات بناءً على تشابه الخصائص، إلا أن الطرق التقليدية تواجه قيودًا في التعامل مع تعقيد أشكال البيانات. وتُعد خوارزمية *Fuzzy Gustafson-Kessel (FGK)* إحدى طرق التجميع الضبابي قادرةً على التعامل مع الأشكال المختلفة للمجموعات، غير أن أدائها يتأثر باختيار معامل *weighting exponent*. هذفت هذه الدراسة إلى تحسين معامل *weighting exponent* في خوارزمية *FGK* وتقييم المجموعات باستخدام مؤشر *Xie Beni Index* عبر عدة قيم للمعاملات وأعداد مختلفة من المجموعات. أظهرت النتائج أن أفضل جودة للتجميع تحققت عند قيمة *weighting exponent* مقدارها 1.1 مع ثلاث مجموعات، حيث أعطت هذه القيمة أفضل أداء مقارنةً بقيم المستخدمة في الدراسة. كما أدى ارتفاع قيمة المعامل إلى انخفاض جودة المجموعات بسبب زيادة مستوى الغموض. وأسفرت نتائج التجميع عن تصنيف البيانات إلى ثلاثة مستويات من مخاطر الاكتتاب، وهي منخفضة ومتوسطة وعالية.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Clustering merupakan salah satu teknik dalam analisis data yang mengelompokkan sekumpulan objek ke dalam kelompok atau *cluster* berdasarkan kesamaan karakteristiknya (Rahayu dkk, 2024). Dalam statistik, hal ini dikenal sebagai analisis *cluster*, yaitu teknik multivariat untuk membentuk kelompok berdasarkan karakteristik yang serupa (Talakua dkk, 2017). Di era *big data*, analisis *cluster* semakin penting karena mampu mengorganisir data berskala besar ke dalam kelompok yang lebih seragam sehingga karakteristik masing-masing kelompok dapat diidentifikasi dengan lebih jelas (Suhaeni dkk, 2018).

Algoritma *clustering* yang paling umum adalah *K-Means*, yang mengelompokkan data berdasarkan jarak terdekat ke *centroid*. Namun, *K-Means* menerapkan pendekatan *hard clustering*, di mana setiap objek hanya dapat menjadi anggota satu *cluster* secara tegas (Wagner dkk, 2022). Selain itu, *K-Means* sensitif terhadap inisialisasi *centroid* (Duan dkk, 2018). Untuk mengatasi keterbatasan tersebut, diperkenalkan pendekatan *fuzzy clustering* yang memungkinkan setiap objek memiliki derajat keanggotaan pada lebih dari satu *cluster*.

Pendekatan *fuzzy clustering* berlandaskan teori *fuzzy set* yang dikembangkan oleh Zadeh (1965). Salah satu algoritma paling dikenal dalam kelompok ini adalah *Fuzzy C-Means* (FCM). FCM lebih fleksibel dibanding *K-Means*, namun masih menggunakan jarak Euclidean sehingga hanya dapat mengenali *cluster* berbentuk sferis, yang berarti sebaran data diasumsikan memiliki variansi yang sama ke segala arah (isotropik) (Bezdek, 1981). Ketika data memiliki orientasi elips atau

memanjang performa FCM menjadi kurang optimal (Cebeci & Yildiz, 2015). Untuk meningkatkan kemampuan adaptasi bentuk *cluster*, Gustafson dan Kessel (1978) memperkenalkan algoritma *Fuzzy Gustafson-Kessel* (FGK).

Algoritma FGK memungkinkan setiap *cluster* memiliki bentuk, ukuran, dan orientasi yang berbeda melalui penggunaan matriks kovarian adaptif. Dengan pendekatan ini, FGK tidak lagi membatasi *cluster* pada bentuk bulat seperti pada FCM, tetapi mampu mengikuti pola penyebaran data yang elips, memanjang, atau miring sesuai struktur sebenarnya. Fleksibilitas tersebut diperoleh karena FGK menggunakan jarak *Mahalanobis* yang disesuaikan untuk setiap *cluster*, sehingga hubungan antarvariabel dan arah penyebaran data juga diperhitungkan dalam proses pengelompokan (Gustafson & Kessel, 1978).

Meskipun demikian, FGK memiliki beberapa kelemahan. Salah satunya terletak pada penentuan nilai *weighting exponent* (m). Parameter ini mengontrol tingkat kekaburan derajat keanggotaan dalam *fuzzy clustering*, yaitu sejauh mana suatu data dapat berbagi keanggotaan pada lebih dari satu *cluster* (Bezdek, 1981). Penelitian *fuzzy clustering* oleh Pal & Bezdek (1995) menunjukkan bahwa pemilihan nilai m sangat berpengaruh terhadap kualitas hasil *clustering*. Nilai m yang terlalu kecil ($< 1,5$) menyebabkan *cluster* menjadi sangat tegas sehingga mendekati *hard clustering*, sedangkan nilai m yang terlalu besar ($> 2,5$) membuat struktur *cluster* menjadi terlalu kabur dan sulit diinterpretasikan. Pemilihan nilai m di luar kisaran tersebut dapat menghasilkan pola *cluster* yang tidak stabil atau tidak representatif, sehingga tidak dapat ditentukan secara sembarangan.

Selain itu, FGK juga dipengaruhi oleh penentuan banyak *cluster* (c), karena nilai tersebut berkaitan dengan proses validitas *cluster*. Banyaknya *cluster* tidak

dapat ditetapkan secara langsung pada satu nilai tertentu, mengingat setiap pilihan c akan membentuk struktur *cluster* yang berbeda dan perlu dievaluasi kesesuaiannya dengan pola distribusi data (Pal & Bezdek, 1995). Oleh karena itu, penetapan c umumnya dilakukan dengan mempertimbangkan beberapa kandidat alternatif untuk memperoleh struktur *cluster* yang paling sesuai dengan karakteristik data yang dianalisis.

Kemampuan FGK dalam mengenali pola penyebaran data yang kompleks membuatnya relevan untuk menganalisis data psikologis, termasuk data risiko depresi pelajar. Data psikologis bersifat multidimensional, tidak homogen, dan mengandung ketidakpastian karena menggambarkan beragam kondisi emosional (Han, 2023). Oleh karena itu, pendekatan *fuzzy clustering* dipandang lebih sesuai untuk memodelkan variasi perilaku dan ketidakpastian tersebut.

Depresi merupakan gangguan mental yang ditandai suasana hati yang terus-menerus tertekan serta hilangnya minat atau kesenangan dalam aktivitas. Kondisi ini berbeda dari perubahan *mood* biasa karena dapat berlangsung lama dan memengaruhi berbagai aspek kehidupan, termasuk hubungan sosial, sekolah, dan pekerjaan (World Health Organization, 2025). Pada pelajar, kondisi ini sering muncul karena tekanan akademik, tuntutan sosial, maupun perubahan emosional yang kompleks. Dalam perspektif islam, kecenderungan emosional semacam ini juga disinggung dalam Q.S. Al-Ma‘arij ayat 19-21:

إِنَّ الْإِنْسَانَ خُلِقَ هَلُوعًا ﴿١٩﴾ إِذَا مَسَّهُ الشَّرُّ جَزُوعًا ﴿٢٠﴾ وَإِذَا مَسَّهُ الْخَيْرُ مَنُوعًا ﴿٢١﴾

Artinya: “Sungguh, manusia diciptakan bersifat suka mengeluh. Apabila dia ditimpa kesusahan dia berkeluh kesah, dan apabila mendapat kebaikan (harta) dia jadi kikir.” (Q.S. Al-Ma‘arij ayat 19-21) (Kementerian Agama RI, 2022)

Menurut Ibn Katsir (dalam Hasanah, 2024), ayat ini menggambarkan sifat manusia yang secara alami cenderung mengeluh dan merasa gusar ketika menghadapi kesulitan. Manusia cenderung gelisah saat menghadapi kesulitan dan enggan berbagi ketika memperoleh kebaikan, yang secara konseptual mencerminkan ketidakstabilan emosi sebagai sifat dasar manusia. Secara psikologis, kecenderungan tersebut dapat meningkatkan kerentanan terhadap depresi, terutama pada pelajar yang belum memiliki mekanisme pengelolaan stres yang efektif. Dengan demikian, ayat ini memberikan landasan konseptual bahwa depresi pada pelajar merupakan fenomena manusiawi yang perlu ditangani melalui pendekatan psikologis yang tepat.

Ada beberapa penelitian yang berkaitan dengan penggunaan algoritma *Fuzzy Gustafson-Kessel* (FGK). Putri & Rochmawati (2021) menerapkan FGK untuk *clustering tweets* akun Go-Jek dan Grab. Dalam penelitian tersebut, *weighting exponent* hanya dicantumkan sebagai parameter tanpa dijelaskan nilai yang digunakan. Banyak *cluster* ditentukan dengan menguji 2-6 *cluster* dan memilih hasil dengan nilai *Xie-Beni Index* terendah, yang menghasilkan dua *cluster* optimum pada kedua data Twitter tersebut.

Sementara itu, penelitian oleh Fajri dkk (2023) menerapkan FGK untuk mengelompokkan wilayah di Sulawesi Tengah berdasarkan indikator penyakit menular. Pada penelitian tersebut, *weighting exponent (m)* ditetapkan sebesar 2. Penentuan banyak *cluster* dilakukan menggunakan indeks validitas *Xie-Beni*, dan dari pengujian diperoleh lima *cluster* sebagai hasil optimum. Penelitian Hidayat (2022) menerapkan FGK untuk mengelompokkan wilayah di Pulau Kalimantan berdasarkan indikator kesejahteraan rakyat. Pada penelitian tersebut, *weighting*

exponent (m) ditetapkan sebesar 2 sebagai nilai standar dan tidak dibandingkan dengan nilai m lainnya. Penentuan banyak *cluster* dilakukan dengan menguji beberapa opsi dan memilih hasil dengan nilai *Xie-Beni Index* terkecil, yang menghasilkan empat *cluster* optimum.

Beberapa penelitian terdahulu yang menerapkan algoritma *Fuzzy Gustafson-Kessel* umumnya telah menguji sejumlah alternatif banyaknya *cluster* untuk mendapatkan struktur pengelompokan yang paling sesuai dengan karakteristik data. Tetapi, penetapan nilai *weighting exponent* pada penelitian tersebut dilakukan secara langsung menggunakan satu nilai standar tanpa dilakukan optimalisasi lebih lanjut. Padahal nilai m memiliki peran penting dalam mengatur tingkat kekaburan dan sensitivitas algoritma terhadap variasi data. Oleh karena itu, diperlukan optimalisasi terhadap nilai m untuk memperoleh nilai yang paling optimal sesuai karakteristik data berdasarkan indeks validitas *cluster*.

Berdasarkan kondisi tersebut, penelitian ini berfokus pada optimalisasi nilai *weighting exponent* guna memperoleh konfigurasi parameter terbaik yang menghasilkan struktur *cluster* paling optimal pada algoritma FGK. Pengujian banyaknya *cluster* (c) tetap dilakukan sebagai bagian dari prosedur validitas untuk memastikan model bekerja secara optimal, namun c tidak menjadi parameter yang diteliti. Dengan demikian, fokus utama penelitian sepenuhnya diarahkan pada optimalisasi nilai m agar algoritma menghasilkan struktur *cluster* yang tidak terlalu tegas (*hard*) maupun terlalu kabur (*fuzzy*), sehingga karakteristik kelompok data risiko depresi pelajar dapat dikenali dengan lebih tepat.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka rumusan masalah dalam penelitian ini adalah bagaimana pengaruh variasi nilai parameter *weighting exponent* pada algoritma *Fuzzy Gustafson-Kessel* terhadap hasil *clustering* data risiko depresi pelajar dalam menentukan parameter yang optimal?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah di atas maka tujuan penelitian ini adalah untuk mengetahui pengaruh variasi nilai parameter *weighting exponent* pada algoritma *Fuzzy Gustafson-Kessel* terhadap hasil *clustering* data risiko depresi pelajar dalam menentukan parameter yang optimal.

1.4 Manfaat Penelitian

Adapun beberapa manfaat yang ada di penelitian ini, adalah sebagai berikut:

1. Bagi Penulis

Menambah pengetahuan dalam menerapkan algoritma *Fuzzy Gustafson-Kessel* serta memahami pengaruh optimalisasi parameter *weighting exponent* terhadap kualitas hasil *clustering*, khususnya pada data risiko depresi pelajar.

2. Bagi Instansi

Memberikan informasi mengenai pengelompokan data risiko depresi pelajar berdasarkan karakteristik masing-masing cluster, sehingga dapat dijadikan bahan pertimbangan dalam merancang program penanganan dan pencegahan depresi yang lebih tepat sasaran.

3. Bagi Pembaca

Dapat dijadikan referensi dan bahan kajian dalam penelitian selanjutnya, khususnya yang berkaitan dengan penerapan algoritma *Fuzzy Gustafson-*

Kessel, optimalisasi parameter *clustering*, maupun analisis data risiko depresi pelajar.

1.5 Batasan Masalah

1. Penelitian ini menggunakan data sekunder *Depression Student Dataset* yang diambil dari platform *Kaggle*, merupakan subset dari *Depression Survey/Dataset for Analysis*, dengan responden pelajar berusia 18-34 tahun.
2. Variasi parameter yang diuji difokuskan pada *weighting exponent* dengan rentang nilai tertentu yang ditentukan berdasarkan literatur, untuk menganalisis pengaruhnya terhadap kualitas hasil *clustering*.
3. Evaluasi kualitas *cluster* dilakukan menggunakan *Xie-Beni Index* sebagai dasar dalam menentukan nilai parameter *weighting exponent* yang optimal.
4. Penentuan banyak *cluster* dilakukan sebatas pengujian beberapa kandidat nilai sebagai prosedur validitas standar, bukan sebagai fokus penelitian.

1.6 Definisi Istilah

Optimalisasi : Suatu tindakan, proses, atau metodologi untuk membuat sesuatu (sebagai sebuah desain, sistem, atau keputusan) menjadi lebih/sepenuhnya sempurna, fungsional, atau lebih efektif.

Logika *Fuzzy* : Metode yang memodelkan suatu masalah dengan logika yang mendekati cara berpikir manusia, di mana kebenaran tidak hanya bernilai 0 atau 1, tetapi dapat berada di antaranya, sehingga mampu merepresentasikan

kondisi yang tidak pasti.

- Clustering* : Metode dalam analisis data yang mengelompokkan sekumpulan objek ke dalam kelompok atau *cluster* berdasarkan kesamaan karakteristiknya.
- Fuzzy clustering* : Metode pengelompokan data di mana setiap titik dapat menjadi anggota beberapa *cluster* dengan derajat keanggotaan antara 0 dan 1.
- Fuzzy Gustafson-Kessel* : Algoritma *fuzzy clustering* yang dikembangkan oleh Gustafson dan Kessel (1978) yang memanfaatkan *fuzzy covariance matrix* dan jarak *Mahalanobis*.
- Xie-Beni Index* : Indeks validitas *clustering* yang mengukur rasio variasi dalam *cluster* terhadap jarak antar *cluster*, di mana nilai lebih kecil menunjukkan kualitas *cluster* yang lebih baik.
- Weighting Exponent* : Parameter pada *fuzzy clustering* yang mengatur tingkat pembagian derajat keanggotaan suatu data ke dalam beberapa *cluster*.

BAB II

KAJIAN TEORI

2.1 Logika *Fuzzy*

Logika *fuzzy* merupakan metode pemecahan masalah yang menggambarkan persoalan dalam bentuk pengetahuan menggunakan logika yang mirip dengan pola pikir manusia. Konsep logika *fuzzy* pertama kali di perkenalkan oleh Prof. Lotfi Asker Zadeh pada tahun 1965 melalui publikasinya yang berjudul “*Fuzzy sets*”. Logika ini dirancang sebagai perkembangan dari logika *boolean* yang bersifat tegas (*crisp*). Pada logika *boolean*, konsep kebenaran hanya dinyatakan dalam dua nilai, yaitu 0 (salah) dan 1 (benar), sehingga sering kali kurang mampu merepresentasikan kondisi nyata yang bersifat tidak pasti (*uncertain*) (Rosita dkk, 2021).

Berbeda dengan logika *boolean*, nilai kebenaran dalam logika *fuzzy* dinyatakan dalam derajat kebenaran dengan derajat keanggotaan antara 0 dan 1. Sehingga, suatu elemen bisa mempunyai dua nilai kebenaran secara bersamaan, namun derajat keanggotaan yang berbeda. Konsep ini membuat keanggotaan dalam himpunan *fuzzy* tidak lagi bersifat mutlak, melainkan bersifat berderajat (Rosita dkk, 2021). Perluasan tersebut memberikan fleksibilitas dalam memodelkan fenomena nyata yang tidak dapat didefinisikan secara tegas, sekaligus menjadi dasar dari berbagai metode berbasis *fuzzy*, termasuk dalam analisis data.

Salah satu konsep dasar dalam logika *fuzzy* adalah himpunan *fuzzy*. Himpunan *fuzzy* merupakan perluasan dari himpunan *crisp*. Pada himpunan *crisp*, nilai keanggotaan suatu elemen x dalam himpunan A , yang biasa ditulis $\mu_A(x)$, hanya memiliki dua kemungkinan, yaitu menjadi anggota ($\mu_A(x) = 1$) atau tidak

menjadi anggota ($\mu_A(x) = 0$). Sebaliknya, pada himpunan *fuzzy*, nilai keanggotaan suatu elemen dalam himpunan A , yang biasa ditulis $\mu_A(x)$ terletak pada interval nilai $[0,1]$. Secara matematis hal ini dinyatakan dengan $\mu_A: X \rightarrow [0,1]$. Himpunan fuzzy A dalam semesta pembicaraan U dinyatakan sebagai himpunan pasangan terurut dari elemen $x(x \in U)$ dengan derajat keanggotaan, yaitu $A = \{(x, \mu_A(x)) \mid x \in U\}$. Selain dapat direpresentasikan secara matematis, himpunan fuzzy juga memiliki dua atribut utama, yaitu linguistik dan numeris. Atribut linguistik adalah label atau istilah dalam bahasa alami yang digunakan untuk menggambarkan suatu kondisi atau keadaan tertentu, seperti muda, parobaya dan tua, sedangkan atribut numeris adalah nilai berupa angka yang menunjukkan ukuran kuantitatif dari suatu variabel, seperti 40, 25, 60, dsb (Rindengan & Langi, 2019).

2.2 Clustering

Clustering merupakan metode penting dalam analisis data yang mengelompokkan sekumpulan objek ke dalam kelompok atau *cluster* berdasarkan kesamaan karakteristik mereka. Dalam *clustering*, objek dalam *cluster* yang sama memiliki kesamaan yang lebih tinggi dibandingkan dengan objek di *cluster* lain. Definisi ini mencerminkan prinsip dasar *clustering*, yaitu homogenitas internal dan heterogenitas eksternal dalam kelompok (Rahayu dkk, 2024). Dalam data mining, *clustering* digunakan untuk mendapatkan informasi dan pengetahuan mengenai distribusi data serta mengamati karakteristik setiap *cluster* untuk dianalisis lebih lanjut (Muhima dkk, 2021). *Clustering* bersifat tanpa arahan (*unsupervised*), artinya metode ini diterapkan tanpa adanya latihan (*training*) dan tanpa ada guru (*teacher*) serta tidak menemukan target *output* (Ulfah & Irtwaty, 2022).

Tujuan dari *clustering* adalah untuk menemukan struktur tersembunyi dalam data yang tidak terstruktur atau semi-terstruktur. Dengan mengelompokkan data berdasarkan kesamaan fitur, *clustering* membantu mengidentifikasi pola dan hubungan yang tidak terlihat secara kasual (Rahayu dkk, 2024). Cara kerja teknik ini adalah dengan mengelompokkan sekumpulan data ke dalam beberapa *cluster*, di mana setiap objek dalam suatu *cluster* memiliki tingkat kesamaan yang tinggi satu sama lain dibandingkan dengan objek di luar *cluster* tersebut, tetapi menunjukkan tingkat kesamaan yang rendah ketika dibandingkan dengan objek di *cluster* lain (Hermansyah dkk, 2020).

Selain itu, *clustering* dapat dibedakan menjadi beberapa jenis, yaitu *hierarchical clustering* dan *non-hierarchical clustering* (Ulfah & Irtwaty, 2022). Secara umum, metode *clustering* juga dapat dikategorikan menjadi *hard clustering*, yang mencakup metode *hierarchical* dan *non-hierarchical*, serta *soft clustering*, seperti *Fuzzy C-Means*, yang memungkinkan data memiliki derajat keanggotaan pada lebih dari satu *cluster* (Thamrin & Wijayanto, 2021). *Hierarchical clustering* adalah metode pengelompokan data yang dimulai dengan menggabungkan dua atau lebih objek yang memiliki kesamaan paling dekat, kemudian secara bertahap menambahkan objek lain berdasarkan tingkat kedekatan, hingga terbentuk struktur hierarki atau pohon yang jelas antar objek, biasanya divisualisasikan dengan dendogram. *Non-hierarchical clustering*, di sisi lain, dimulai dengan menetapkan jumlah *cluster* tertentu terlebih dahulu, kemudian melakukan pengelompokan tanpa mengikuti proses hierarki (Ulfah & Irtwaty, 2022). Lebih lanjut, *hard clustering* seperti *K-Means* dan *K-Medoids* menempatkan setiap data secara tegas ke satu *cluster*, sedangkan *soft clustering* seperti *Fuzzy C-Means* lebih fleksibel karena

memungkinkan data memiliki keanggotaan pada beberapa *cluster* sekaligus (Prastanika & Wijayanto, 2023).

Analisis *cluster* merupakan metode analitis yang digunakan untuk mengelompokkan sekumpulan data yang tidak memiliki target atau label spesifik (*unsupervised*) (Muhima dkk, 2021). Analisis ini bertujuan untuk mengelompokkan individu atau objek ke dalam kelompok yang lebih kecil, sehingga objek dalam suatu kelompok memiliki karakteristik yang serupa (Ayuningtias dkk, 2019). Analisis *cluster* mampu mengelompokkan data hasil observasi dalam jumlah besar dengan jumlah variabel yang relatif banyak. Data yang telah direduksi ke dalam kelompok-kelompok akan lebih mudah untuk dianalisis. Namun, pada data yang bersifat heterogen antarobjek penelitian, sering kali mengalami kesulitan dalam menentukan jumlah kelompok yang tepat untuk dibentuk (Telaumbanua dkk, 2025).

2.3 Fuzzy Clustering

Fuzzy Clustering adalah metode pengelompokan data dalam ruang vektor yang menggunakan ukuran jarak berbasis norma Euclidean untuk mengukur kedekatan antara vektor. Metode ini berperan penting dalam pemodelan sistem *fuzzy*, terutama pada identifikasi aturan *fuzzy* yang kompleks, meskipun penentuan jumlah *cluster* optimal umumnya memerlukan metode evaluasi tambahan (Rahakbauw dkk, 2017). Selain itu, *fuzzy clustering* juga dikategorikan sebagai *soft clustering*, yaitu metode pengelompokan di mana setiap titik data dapat menjadi anggota dari dua atau lebih *cluster* secara bersamaan. *Fuzzy clustering* menerapkan prinsip logika *fuzzy*, di mana nilai keanggotaan suatu elemen tidak mutlak benar atau salah, tetapi dapat berada dalam rentang nilai di antara keduanya (*intermediate*). Dengan demikian,

metode ini tidak memaksa titik data untuk masuk hanya ke dalam satu *cluster*, melainkan memberikan nilai derajat keanggotaan yang menunjukkan tingkat keterkaitan data terhadap setiap *cluster*. Nilai derajat keanggotaan tersebut berada pada rentang $[0,1]$ (Nurzida dkk, 2024).

Teori himpunan *fuzzy* yang diperkenalkan oleh Lotfi A. Zadeh pada tahun 1965 kemudian memberikan dasar teoretis bagi perkembangan metode *fuzzy clustering*. Konsep ini membawa gagasan mengenai ketidakpastian keanggotaan (*uncertainty of belonging*) yang direpresentasikan melalui fungsi keanggotaan (*membership function*). Pendekatan ini memungkinkan adanya informasi keanggotaan yang tidak bersifat mutlak, sehingga suatu objek dapat memiliki tingkat keterkaitan yang berbeda terhadap beberapa *cluster*. Aplikasi awal teori ini dalam analisis *cluster* dilakukan oleh Bellman, Kalaba, Zadeh, dan Ruspini, yang membuka jalan bagi penelitian lanjutan dalam bidang *fuzzy clustering*. Saat ini, metode *fuzzy clustering* telah banyak dikembangkan dan diaplikasikan secara luas di berbagai bidang, serta menjadi salah satu alat penting dalam analisis *cluster* modern (Yang, 1993).

Menurut Yang (1993), konsep dasar *fuzzy clustering* diperkenalkan oleh Ruspini melalui perluasan dari partisi tegas (*hard c-partition*). Misalkan $X \subseteq \mathbb{R}^s$ adalah himpunan data dan $c > 1$ jumlah *cluster*. Dalam *hard partition*, setiap data hanya menjadi anggota satu *cluster* dengan fungsi indikator:

$$p_i(x) = \begin{cases} 1, & \text{jika } x \in B_i \\ 0, & \text{jika } x \notin B_i \end{cases}$$

dengan:

$p_i(x)$: Fungsi indikator,

B_i : *Cluster* ke- i .

Teori himpunan *fuzzy* yang diperkenalkan oleh Zadeh (1965) memperluas fungsi ini menjadi fungsi keanggotaan $\mu_i(x)$ yang bernilai pada interval $[0,1]$. Berdasarkan hal tersebut, Ruspini memperkenalkan *fuzzy c-partition* $P = (\mu_1, \mu_2, \dots, \mu_c)$ dengan $\sum_{i=1}^c \mu_i(x) = 1$. Konsep ini memungkinkan setiap data memiliki derajat keanggotaan terhadap lebih dari satu *cluster* secara bersamaan.

Berdasarkan prinsip ini, berbagai metode *fuzzy clustering* dikembangkan, Menurut (Yang, 1993), diklasifikasikan menjadi tiga pendekatan utama:

1. Pendekatan Berbasis Relasi *Fuzzy* (*Fuzzy Relation-Based Approach*)

Metode awal ini menggunakan relasi *fuzzy* $R: X \times X \rightarrow [0,1]$ untuk mengukur kemiripan antar data, dengan sifat refleksivitas, simetri, dan transitivitas. Komposisi relasi berulang menghasilkan relasi kesamaan R^∞ yang stabil, menjadi dasar pembentukan *cluster* berbasis hierarki atau hubungan antar data.

2. Pendekatan Berbasis Fungsi Objektif (*Objective Function-Based Approach*)

Pendekatan berbasis fungsi objektif merupakan metode yang paling banyak digunakan dalam *fuzzy clustering* modern. Pendekatan ini berfokus pada proses optimasi, yaitu dengan meminimalkan fungsi objektif yang mewakili tingkat ketidakmiripan (*dissimilarity*) antara data dengan pusat *cluster*. Misalkan $d(x, y)$ adalah fungsi jarak antara dua titik data x dan y yang memenuhi dua syarat dasar:

$$d(x, y) = d(y, x), \quad d(x, x) = 0$$

Fungsi objektif pertama kali diperkenalkan oleh Ruspini (1969) dan dinyatakan sebagai berikut:

$$J_R(U) = \sum_{i=1}^c \sum_{j=1}^n \mu_{i,j} d(x_j, x_i) \quad (2.1)$$

dengan:

$J_R(U)$: Fungsi objektif,

U : Matriks keanggotaan,

$\mu_{i,j}$: Derajat keanggotaan data x_j terhadap *cluster* ke- i ,

$d(x_j, x_i)$: Jarak antara data x_j dengan pusat *cluster* x_i ,

c : Banyaknya *cluster*,

n : Banyaknya data.

Tujuan dari pendekatan ini adalah mencari partisi *fuzzy* U^* yang meminimalkan fungsi objektif $J_R(U)$ dengan U^* sebagai matriks keanggotaan optimal hasil proses optimasi.

3. Pendekatan *Fuzzy Generalized k-Nearest Neighbor (Fuzzy Generalized k-NN Approach)*

Metode ini mengembangkan *k-NN* konvensional dengan memberikan derajat keanggotaan setiap data terhadap semua *cluster* berdasarkan kedekatannya. Kelebihannya adalah informasi keanggotaan yang lebih kaya, konsistensi terhadap risiko posterior, dan konvergensi cepat. Meski kurang populer dibanding fungsi objektif, pendekatan ini tetap signifikan dalam pengembangan *fuzzy clustering*.

2.4 *Weighting Exponent (m)*

Pada *fuzzy clustering*, *weighting exponent (m)* adalah parameter yang mengatur tingkat ketidak-tegasan (*fuzziness*) dalam proses pengelompokan. Semakin besar nilai m , semakin besar ketidak-tegasan pada derajat keanggotaan suatu data terhadap *cluster* (Bezdek, 1981). Bezdek juga menjelaskan bahwa Bezdek (1981) menjelaskan bahwa nilai m menentukan tingkat ketegasan suatu partisi, di mana

nilai m yang mendekati 1 membuat hasil pengelompokan semakin menyerupai *hard clustering*, sedangkan nilai m yang lebih besar menghasilkan keanggotaan yang lebih kabur.

Penelitian oleh Pal & Bezdek (1995), menunjukkan bahwa nilai m umumnya digunakan pada kisaran 1,5 hingga 2,5 dengan $m = 2$ sering diterapkan sebagai nilai standar karena dianggap memberikan keseimbangan antara stabilitas dan tingkat *fuzziness*. Variasi nilai m berpengaruh terhadap kecenderungan data untuk lebih berpihak pada *cluster* tertentu, kejelasan batas antarklaster, serta sensitivitas model terhadap *noise* dan *outlier*. Pada algoritma *Fuzzy Gustafson-Kessel* (FGK), konsep m tetap digunakan sebagaimana dalam *Fuzzy C-Means* karena pengaturan derajat keanggotaan oleh m merupakan dasar dari seluruh proses perhitungan, termasuk ketika FGK membentuk *cluster* berbentuk elips melalui pemanfaatan matriks kovarian adaptif. Oleh karena itu, *weighting exponent* (m) menjadi elemen teoritis penting yang mendasari mekanisme *fuzzy clustering*, termasuk penerapannya pada FGK.

Secara matematis, nilai *weighting exponent* harus memenuhi $m > 1$. Hal ini dikarenakan pada proses pembaruan matriks keanggotaan terdapat operasi perpangkatan yang melibatkan bentuk $\frac{2}{m-1}$. Jika $m = 1$, maka akan terjadi pembagian dengan nol sehingga perhitungan menjadi tidak terdefinisi. Selain itu, nilai $m = 1$ akan menghilangkan sifat *fuzzy* dan mengubah metode menjadi *hard clustering*.

2.5 *Fuzzy Gustafson-Kessel* (FGK)

Metode *Fuzzy Gustafson-Kessel* (FGK) diperkenalkan oleh Gustafson & Kessel (1978) sebagai pengembangan dari algoritma *fuzzy clustering*, untuk

mengatasi keterbatasan metode sebelumnya yang hanya mampu membentuk *cluster* bulat karena menggunakan jarak Euclidean. FGK memperkenalkan *fuzzy covariance matrix* berbeda pada setiap *cluster*, sehingga setiap *cluster* dapat memiliki bentuk, ukuran, dan orientasi yang berbeda, serta mampu menyesuaikan diri dengan struktur data yang tidak simetris. Dengan pendekatan ini, FGK dapat mendeteksi *cluster* berbentuk kompleks, tidak hanya bulat, dan secara konseptual dapat dipandang sebagai perluasan alami dari *maximum likelihood estimation* (MLE) pada *mixture decomposition* (Gustafson & Kessel, 1978).

Algoritma dari pengelompokan *fuzzy Gustafson-Kessel* (Bezdek, 1981; Paradilla dkk, 2023):

1. Masukkan data yang akan dikelompokkan.

Misalkan data dengan bentuk matriks X ukuran 4×2 :

$$X = \begin{bmatrix} 2 & 3 \\ 4 & 2 \\ 6 & 7 \\ 7 & 6 \end{bmatrix}$$

2. Menentukan banyaknya *cluster* yang akan dibentuk ($c \geq 2$), *weighting exponent* ($m > 1$), maksimum iterasi (MaxIter), batas ambang konvergensi (ε), fungsi objektif awal ($J_{GK}^{(0)} = 0$), dan iterasi awal ($t = 1$).

Misalkan ditetapkan parameter awal sebagai berikut:

$$c = 2, m = 2, \text{MaxIter} = 100, \varepsilon = 0,001, J_{GK}^{(0)} = 0, t = 1$$

3. Membangkitkan nilai derajat keanggotaan awal $\mu_{i,k}$ secara acak untuk $i = 1, 2, 3, \dots, n; 1 \leq k \leq c$ sebagai elemen matriks keanggotaan awal U_0 .

$$U_0 = \begin{bmatrix} \mu_{1,1} & \mu_{1,2} & \cdots & \mu_{1,c} \\ \mu_{2,1} & \mu_{2,2} & \cdots & \mu_{2,c} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n,1} & \mu_{n,2} & \cdots & \mu_{n,c} \end{bmatrix} \text{ dengan } \sum_{k=1}^c \mu_{i,k} = 1 \quad (2.2)$$

Sebagai contoh, untuk data dengan $n = 4$ dan $c = 2$, matriks keanggotaan awal U_0 dapat berupa:

$$U_0 = \begin{bmatrix} 0,6 & 0,4 \\ 0,3 & 0,7 \\ 0,5 & 0,5 \\ 0,8 & 0,2 \end{bmatrix}$$

Masing-masing baris menunjukkan derajat keanggotaan setiap data terhadap 2 *cluster*, dan jumlah pada tiap baris sama dengan 1.

4. Menentukan pusat *cluster* ke- k , dinotasikan sebagai v_{kj} , dengan $k = 1, 2, \dots, c$ sebagai banyaknya *cluster*, $j = 1, 2, \dots, p$ sebagai banyaknya variabel, dan $x_{i,j}$ merupakan nilai data ke- i pada variabel ke- j :

$$v_{k,j} = \frac{\sum_{i=1}^n (\mu_{i,k})^m x_{i,j}}{\sum_{i=1}^n (\mu_{i,k})^m} \quad (2.3)$$

Hasil perhitungan $v_{k,j}$ untuk seluruh variabel kemudian membentuk sebuah vektor pusat *cluster*, yaitu:

$$v_k = (v_{k1}, v_{k2}, \dots, v_{kp})$$

Vektor v_k inilah yang digunakan dalam perhitungan matriks kovarian pada algoritma *Fuzzy Gustafson-Kessel*.

Contoh Perhitungan:

Untuk *cluster* 1, variabel pertama ($j = 1$):

$$\begin{aligned} v_{1,1} &= \frac{(0,6)^2 \cdot 2 + (0,3)^2 \cdot 4 + (0,5)^2 \cdot 6 + (0,8)^2 \cdot 7}{(0,36 + 0,09 + 0,25 + 0,64)} \\ &= \frac{0,72 + 0,36 + 1,5 + 4,48}{1,34} = \frac{7,06}{1,34} \approx 5,27 \end{aligned}$$

Untuk *cluster* 1, variabel kedua ($j = 2$):

$$v_{1,2} = \frac{(0,6)^2 \cdot 3 + (0,3)^2 \cdot 2 + (0,5)^2 \cdot 7 + (0,8)^2 \cdot 6}{(0,36 + 0,09 + 0,25 + 0,64)}$$

$$= \frac{1,08 + 0,18 + 1,75 + 3,84}{1,34} = \frac{6,85}{1,34} \approx 5,11$$

Sehingga pusat *cluster* 1 adalah:

$$v_1 = (5,27, 5,11)$$

Perhitungan untuk *cluster* 2 dilakukan dengan cara yang sama seperti pada *cluster* 1.

5. Menentukan matriks kovarian

$$F_k = \frac{\sum_{i=1}^n (\mu_{i,k})^m (x_i - v_k)(x_i - v_k)^T}{\sum_{i=1}^n (\mu_{i,k})^m} \quad (2.4)$$

dengan:

F_k : Matriks *fuzzy covariance cluster* ke- k ,

x_i : Vektor data ke- i ,

v_k : Pusat *cluster* ke- k ,

$\mu_{i,k}$: Derajat keanggotaan data ke- i pada *cluster* ke- k ,

m : *Weighting exponent* atau pangkat pembobot untuk fungsi keanggotaan *fuzzy*,

n : Banyaknya data.

Contoh Perhitungan:

Untuk data ke-1 = [2,3]

$$x_1 - v_1 = [2 - 5,27; 3 - 5,11] = [-3,27 \quad -2,11]$$

$$(x_1 - v_1)(x_1 - v_1)^T = [-3,27 \quad -2,11] \cdot \begin{bmatrix} -3,27 \\ -2,11 \end{bmatrix} = \begin{bmatrix} 10,70 & 6,90 \\ 6,90 & 4,45 \end{bmatrix}$$

$$(\mu_{1,1})^2 (x_1 - v_1)(x_1 - v_1)^T = 0,36 \cdot \begin{bmatrix} 10,70 & 6,90 \\ 6,90 & 4,45 \end{bmatrix} = \begin{bmatrix} 3,85 & 2,48 \\ 2,48 & 1,60 \end{bmatrix}$$

Data ke-2 sampai data ke-4:

$$\sum_{i=2}^4 (\mu_{i,1})^2 (x_i - v_1)(x_i - v_1)^T = \begin{bmatrix} 2,69 & 2,55 \\ 2,55 & 5,67 \end{bmatrix}$$

$$\sum_{i=1}^4 (\mu_{i,1})^2 (x_i - v_1)(x_i - v_1)^T = \begin{bmatrix} 6,54 & 5,03 \\ 5,03 & 7,27 \end{bmatrix}$$

$$F_1 = \frac{1}{1,34} \begin{bmatrix} 6,54 & 5,03 \\ 5,03 & 7,27 \end{bmatrix} = \begin{bmatrix} 4,88 & 3,75 \\ 3,75 & 5,43 \end{bmatrix}$$

Perhitungan F_2 dilakukan dengan cara yang sama seperti pada F_1 .

6. Menentukan jarak *Mahalanobis* dengan persamaan berikut:

$$D_{i,kA_k}^2 = (x_i - v_k)^T A_k (x_i - v_k) \quad (2.5)$$

$$A_k = [\rho_k \det(F_k)^{\frac{1}{p}} F_k^{-1}]$$

dengan:

D_{i,kA_k}^2 : Jarak data ke- i terhadap pusat *cluster* ke- k dengan *norm inducing matrix* A_k ,

x_i : Vektor data ke- i ,

v_k : Pusat *cluster* ke- k ,

F_k : Matriks *fuzzy covariance cluster* ke- k ,

ρ_k : Volume *cluster* ke- k , $\rho_k = 1$,

p : Banyak variabel.

Parameter ρ_k merupakan parameter volume *cluster* pada algoritma *Fuzzy Gustafson-Kessel*. Dalam penelitian ini, nilai ρ_k ditetapkan sebesar 1 dengan mengikuti formulasi standar algoritma FGK dan tidak diestimasi dari data.

Contoh Perhitungan:

Data ke-1, *cluster* 1:

$$\det(F_1) = (4,88)(5,43) - (3,75)^2 = 12,43$$

$$\det(F_1)^{1/2} = \sqrt{12,43} \approx 3,52$$

$$F_1^{-1} = \frac{1}{12,43} \begin{bmatrix} 5,43 & -3,75 \\ -3,75 & 4,88 \end{bmatrix} \approx \begin{bmatrix} 0,43 & -0,30 \\ -0,30 & 0,39 \end{bmatrix}$$

$$\begin{aligned} A_1 &= \rho_i \det(F_1)^{\frac{1}{2}} \cdot F_1^{-1} \\ &= 3,52 \cdot \begin{bmatrix} 0,43 & -0,30 \\ -0,30 & 0,39 \end{bmatrix} \approx \begin{bmatrix} 1,51 & -1,05 \\ -1,05 & 1,37 \end{bmatrix} \end{aligned}$$

$$D_{1,1A_1}^2 = \begin{bmatrix} -3,27 \\ -2,11 \end{bmatrix} \cdot \begin{bmatrix} 1,51 & -1,05 \\ -1,05 & 1,37 \end{bmatrix} \cdot \begin{bmatrix} -3,27 & -2,11 \end{bmatrix} \approx 7,75$$

$$D_{1,1A_1} = \sqrt{7,75} \approx 2,78$$

Perhitungan jarak *Mahalanobis* untuk data 2-4 dan *cluster* 2 dilakukan dengan cara yang sama seperti pada data 1 dan *cluster* 1.

7. Menentukan fungsi objektif pada iterasi ke- t dengan rumus:

$$J_{GK}^{(t)} = \sum_{i=1}^n \sum_{k=1}^c (\mu_{i,k})^m D_{i,kA_k}^2 \quad (2.6)$$

dengan:

$J_{GK}^{(t)}$: Fungsi objektif pada iterasi ke- t ,

$\mu_{i,k}$: Derajat keanggotaan data ke- i terhadap cluster ke- k ,

D_{i,kA_k}^2 : Jarak antara data ke- i dan pusat *cluster* ke- k berdasarkan matriks

A_k .

Sebagai contoh, fungsi objektif pada iterasi pertama dihitung untuk data 1 terhadap *cluster* 1:

$$J_{GK}^{(1)} = (\mu_{1,1})^m \cdot D_{1,1A_1}^2 = (0,6)^2 \cdot 7,75 \approx 2,79$$

Fungsi objektif iterasi pertama $J_{GK}^{(1)}$ diperoleh dengan menjumlahkan kontribusi dari seluruh data dan *cluster*, sehingga menghasilkan satu nilai

skalar, misal 0,001. Perhitungan serupa diterapkan untuk semua data dan *cluster* lainnya.

8. Memperbarui matriks keanggotaan U_{t+1}

Matriks U_t merupakan matriks keanggotaan pada iterasi ke- t , sedangkan U_{t+1} adalah matriks keanggotaan baru yang dihasilkan setelah proses pembaruan. Karena matriks U tersusun dari elemen-elemen derajat keanggotaan $\mu_{i,k}$, maka pembaruan U_{t+1} dilakukan dengan menghitung ulang setiap $\mu_{i,k}$ menggunakan persamaan berikut:

$$\mu_{i,k}^{(t+1)} = \left[\sum_{q=1}^c \left(\frac{D_{i,kA_k}}{D_{i,qA_q}} \right)^{\frac{2}{m-1}} \right]^{-1} \quad (2.7)$$

dengan:

$\mu_{i,k}^{(t+1)}$: Derajat keanggotaan data x_i pada *cluster* ke- k pada iterasi ke-
($t + 1$),

D_{i,qA_q} : Jarak antara data ke- i dan pusat *cluster* ke- q berdasarkan matriks
 A_q ,

q : Indeks pembandingan jarak terhadap semua *cluster*.

Misalkan:

Cluster 1: $D_{1,1A_1} = 2,78$ dan *cluster* 2: $D_{1,2A_1} = 2,49$

$$\mu_{1,1}^{(t+1)} = \left[\left(\frac{D_{1,1A_1}}{D_{1,1A_1}} \right)^{\frac{2}{2-1}} + \left(\frac{D_{1,1A_2}}{D_{1,2A_2}} \right)^{\frac{2}{2-1}} \right]^{-1} = \left[1 + \left(\frac{2,78}{2,49} \right)^2 \right]^{-1}$$

$$\mu_{1,1}^{(t+1)} = [1 + 1,24]^{-1} = 0,44$$

Jadi derajat keanggotaan data ke-1 terhadap *cluster* 1 diperbarui menjadi 0,44. Secara otomatis, $\mu_{1,2}^{(t+1)} = 1 - 0,44 = 0,56$ (karena jumlah

keanggotaan per data sama dengan 1). Dengan perhitungan yang sama diterapkan untuk data dan *cluster* lainnya.

9. Membandingkan matriks keanggotaan $U = [\mu_{i,k}]$, jika $\|U_{t+1} - U_t\| < \varepsilon$ atau ($t > \text{MaxIter}$) maka sudah konvergen. Nilai *threshold* (ε) adalah suatu nilai yang sangat kecil mendekati nol. Nilai $\|U_{t+1} - U_t\|$ dihitung menggunakan norma *Frobenius*, yaitu norma matriks yang mengukur besarnya perbedaan dua matriks melalui akumulasi seluruh perubahan elemen di dalamnya (Herzog dkk, 2023). Jika $\|U_{t+1} - U_t\| \geq \varepsilon$ kembali ke langkah 4.

Misalnya, untuk data 1 terhadap *cluster* 1, μ berubah dari 0,6 menjadi 0,445; perubahan ini masih cukup besar, sehingga iterasi dilanjutkan hingga seluruh μ_{ik} stabil mendekati 0,001.

2.6 Xie-Beni Index

Xie-Beni Index (XB) merupakan salah satu ukuran validitas *clustering* dengan membandingkan variasi internal *cluster* dan jarak antarpusat *cluster*. Semakin kecil nilai XB, semakin kompak dan terpisah *cluster* tersebut, sehingga banyaknya *cluster* optimal ditentukan dari nilai XB minimum. Persamaan *Xie-Beni Index* dapat dituliskan sebagai berikut (Feng & Fan, 2016):

$$v_{XB} = \frac{S}{d_{min}} \quad (2.9)$$

di mana,

$$S = \frac{\sum_{k=1}^c \sum_{i=1}^n (\mu_{i,k})^m d_{i,k}^2(x_i, v_k)}{n}$$

$$d_{i,k}^2(x_i, v_k) = \|x_i - v_k\|^2$$

$$d_{min} = \min_{i \neq k} d_{i,k}^2(v_i, v_k) = \min_{i \neq k} \|v_i - v_k\|^2$$

Norm yang digunakan adalah norm Euclidean (2-norm), yang didefinisikan sebagai

$$\|a\| = \sqrt{a^T a}. \text{ Dengan demikian, untuk vektor selisih } x_i - v_k \text{ berlaku } \|x_i - v_k\|^2 = (x_i - v_k)^T (x_i - v_k).$$

dengan:

v_{XB} : Nilai *Xie-Beni Index*,

S : Tingkat kekompakan internal *cluster*,

$\mu_{i,k}$: Derajat keanggotaan data ke- i terhadap *cluster* ke- k ,

d_{min} : Jarak minimum antarpusat *cluster*,

$d_{i,k}^2$: Jarak kuadrat antara data ke- i dan pusat *cluster* ke- k ,

x_i : Data ke- i ,

v_k : Pusat *cluster* ke- k ,

m : Pangkat pembobot (*weighting exponent*).

Sebagai contoh:

Data:

$$x_1 = 1, \quad x_2 = 4$$

Pusat *cluster*:

$$v_1 = 2,5 \quad v_2 = 3,5$$

Matriks keanggotaan:

$$U = \begin{bmatrix} \mu_{1,1} & \mu_{1,2} \\ \mu_{2,1} & \mu_{2,2} \end{bmatrix} = \begin{bmatrix} 0,8 & 0,2 \\ 0,6 & 0,4 \end{bmatrix}$$

Parameter *fuzzy*:

$$m = 2, \quad n = 2$$

Perhitungan S :

1. Data 1 ke *cluster* 1

$$\|x_1 - v_1\|^2 = (1 - 2,5)^2 = 2,25$$

$$\mu_{1,1}^2 = (0,8)^2 = 0,64$$

$$\mu_{1,1}^2 \cdot \|x_1 - v_1\|^2 = 0,64 \cdot 2,25 = 1,44$$

2. Data 1 ke *cluster* 2

$$\|x_1 - v_2\|^2 = (1 - 3,5)^2 = 6,25$$

$$\mu_{1,2}^2 = (0,2)^2 = 0,04$$

$$\mu_{1,2}^2 \cdot \|x_1 - v_2\|^2 = 0,04 \cdot 6,25 = 0,25$$

3. Data 2 ke *cluster* 1

$$\|x_2 - v_1\|^2 = (4 - 2,5)^2 = 2,25$$

$$\mu_{2,1}^2 = (0,6)^2 = 0,36$$

$$\mu_{2,1}^2 \cdot \|x_2 - v_1\|^2 = 0,36 \cdot 2,25 = 0,81$$

4. Data 2 ke *cluster* 2

$$\|x_2 - v_2\|^2 = (2 - 3,5)^2 = 2,25$$

$$\mu_{2,2}^2 = (0,4)^2 = 0,16$$

$$\mu_{2,2}^2 \cdot \|x_2 - v_2\|^2 = 0,16 \cdot 2,25 = 0,36$$

$$\sum_{k=1}^2 \sum_{i=1}^2 (\mu_{i,k})^m \|x_i - v_k\|^2 = 1,44 + 0,25 + 0,81 + 0,36 = 2,86$$

sehingga:

$$S = \frac{2,86}{2} = 1,43$$

Perhitungan d_{min}

$$\|v_1 - v_2\|^2 = (2,5 - 3,5)^2 = 1$$

$$d_{min} = 1$$

Perhitungan *Xie-Beni Index*:

$$XB = \frac{S}{d_{min}} = \frac{1,43}{1} = 1,43$$

Untuk banyak *cluster* 3 dan seterusnya, proses perhitungannya dilakukan dengan cara yang sama. Setelah semua nilai XB diperoleh, nilai-nilai tersebut dibandingkan dan dipilih banyaknya *cluster* dengan nilai XB paling kecil.

2.7 Preprocessing Data

Preprocessing data merupakan tahap awal untuk menyiapkan data mentah agar siap dianalisis. Proses ini diperlukan karena data sering memiliki skala atau format yang berbeda, sehingga perlu disesuaikan agar hasil analisis lebih akurat. Penyesuaian tersebut umumnya mencakup proses normalisasi dan label *encoding*.

2.7.1 Label Encoding

Label *encoding* merupakan proses transformasi data kategorikal menjadi bentuk numerik dengan mengubah setiap nilai kategori menjadi bilangan bulat. Setiap kategori diberi label angka berdasarkan urutan kemunculannya dalam data atau secara alfabetis, sehingga data kategorikal dapat direpresentasikan dalam bentuk numerik (Kiwasi Wororomi dkk, 2024). Teknik ini sederhana dan tidak menambahkan dimensi pada data, sehingga sering digunakan pada tahap *preprocessing* data. Namun, perlu diperhatikan bahwa proses label *encoding* dapat menciptakan hubungan ordinal antara kategori, yang menunjukkan urutan tertentu yang tidak ada dalam data asli. Kondisi ini berpotensi menyebabkan algoritma salah menafsirkan hubungan antara kategori. Oleh karena itu, label *encoding* lebih sesuai diterapkan pada variabel kategorikal yang bersifat ordinal (Herdian dkk, 2024).

Sebagai contoh penerapan label *encoding* pada kategori ordinal, misalkan variabel tingkat pendidikan memiliki tiga kategori: SD, SMP, dan SMA. Dengan label encoding, setiap kategori diberikan angka sesuai urutan tingkatannya, yaitu SD bernilai 0, SMP bernilai 1, dan SMA bernilai 2. Pemberian angka ini memungkinkan data kategorikal diubah menjadi numerik sehingga dapat diproses oleh algoritma, sekaligus mempertahankan urutan logis antar kategori.

2.7.2 Normalisasi *Min-Max*

Normalisasi merupakan salah satu tahap dalam *preprocessing* data yang bertujuan untuk menyamakan skala nilai atribut numerik dalam suatu *dataset* tanpa menghilangkan perbedaan relatif antara rentang nilainya. Secara umum, proses ini memastikan bahwa semua nilai data diperlakukan secara setara, sehingga analisis dapat dilakukan secara optimal tanpa bias yang disebabkan oleh perbedaan skala nilai antara atribut (Hapsari dkk, 2024). Salah satu metode normalisasi yang paling sering digunakan adalah normalisasi *Min-Max*, sebuah metode yang bertujuan untuk mengubah skala nilai dalam suatu *dataset* menjadi rentang tertentu, biasanya antara 0 dan 1. Rumus dari normalisasi *min-max* adalah:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (2.8)$$

dengan:

x' : Nilai hasil normalisasi,

x : Nilai persatuan data,

x_{max} : Nilai maksimum dari seluruh data,

x_{min} : Nilai minimum dari seluruh data.

Sebagai contoh:

Data tekanan darah dari tiga orang dewasa: [110,90,120]

$Min = 90$

$Max = 120$

$$x' = \frac{110-90}{120-90} = 0,666, \quad x' = \frac{90-90}{120-90} = 0, \quad x' = \frac{120-90}{120-90} = 1$$

Maka hasil untuk normalisasi data tekanan darah adalah [0,666, 0, 1].

2.8 *Principal Component Analysis (PCA)*

Principal Component Analysis (PCA) adalah metode reduksi dimensi yang bertujuan untuk menyederhanakan data berdimensi tinggi dengan mengubah variabel awal menjadi sejumlah komponen utama yang saling ortogonal. Komponen utama tersebut dibentuk sedemikian rupa sehingga mampu merepresentasikan sebagian besar variasi data meskipun jumlahnya lebih sedikit dibandingkan dengan jumlah variabel asal. Selain itu, PCA banyak digunakan sebagai alat bantu visualisasi karena mampu menyajikan informasi penting secara lebih ringkas dan terstruktur, sehingga pola hasil pengelompokan data dapat diamati dengan lebih jelas (Firdawanti dkk, 2025). Adapun tahapan PCA dapat dijelaskan berikut ini (Syahreni dkk, 2025):

1. Normalisasi Data

Normalisasi dilakukan agar seluruh variabel berada pada skala yang sebanding sehingga tidak ada variabel yang mendominasi pembentukan komponen utama.

2. Pembentukan Komponen Utama

Setelah data dinormalisasi, dibentuk matriks kovarians untuk menggambarkan hubungan antarvariabel. Selanjutnya dilakukan perhitungan nilai eigen dan

vektor eigen dengan persamaan:

$$C \cdot v = \lambda \cdot v$$

di mana C adalah matriks kovarians dari data yang telah dinormalisasi, v merupakan vektor eigen dari matriks kovarians, dan λ adalah nilai eigen yang bersesuaian.

3. Reduksi Dimensi dan Visualisasi

Komponen utama dipilih berdasarkan nilai eigen terbesar, kemudian data diproyeksikan ke dalam ruang berdimensi lebih rendah. Hasil proyeksi tersebut divisualisasikan untuk menampilkan sebaran data dan pola pemisahan antarklaster.

2.9 Depresi

Depresi adalah kondisi yang ditandai dengan perubahan *mood* pada seseorang, seperti kesedihan yang mendalam, kekecewaan, perasaan kehilangan, dan kegagalan, yang mengakibatkan ketidakmampuan individu untuk beradaptasi dengan lingkungannya. Perasaan sedih yang bersifat sementara tidak dapat dijadikan cerminan dari depresi, karena kondisi ini dikategorikan sebagai gangguan medis-psikiatrik apabila berlangsung secara persisten hingga aktivitas sosial sehari-hari terganggu. Kondisi ini memengaruhi berbagai aspek kehidupan seseorang, termasuk aspek psikologis, fisik, dan sosial (Dillasamola, 2021).

Anak-anak usia sekolah dikategorikan sebagai kelompok rentan yang memiliki potensi tinggi untuk mengalami berbagai masalah psikososial. Anak-anak yang mengalami depresi menunjukkan sejumlah gejala, termasuk tidak pernah merasa puas, merasa bosan dengan kehidupan, selalu merasa putus asa, diam dalam waktu yang lama dan enggan berbicara, mengalami kesedihan mendalam disertai rasa

takut dan cemas, mengalami penurunan kesadaran, memiliki pikiran yang kacau, dan tidak mengenali batas atau larangan (Saputri & Nurrahima, 2020).

Selain anak-anak, perhatian global telah tertuju pada depresi yang dialami oleh mahasiswa perguruan tinggi dalam beberapa dekade terakhir. Prevalensi depresi pada kelompok ini dilaporkan lebih tinggi dibandingkan dengan populasi umum, dan beberapa individu menunjukkan risiko bunuh diri yang signifikan. Tekanan akademik yang tinggi, isolasi sosial, kesulitan finansial, dan transisi dari masa remaja ke dewasa telah diidentifikasi sebagai faktor yang memicu dan memperparah kondisi ini. Fenomena ini menunjukkan bahwa depresi di kalangan mahasiswa bukan hanya masalah kesehatan mental, tetapi juga mewakili tantangan berat yang dihadapi generasi muda dalam lingkungan akademik (Ratu dkk, 2025).

Gejala depresi dapat muncul dalam berbagai bentuk yang memengaruhi aspek emosional, kognitif, fisik, dan sosial individu. Secara emosional, penderita sering merasakan kesedihan mendalam, kehilangan minat terhadap aktivitas yang sebelumnya menyenangkan, rasa putus asa, dan perasaan tidak berharga. Dari sisi kognitif, depresi ditandai dengan kesulitan berkonsentrasi, pengambilan keputusan yang terganggu, serta munculnya pikiran negatif atau bahkan keinginan untuk menyakiti diri. Secara fisik, individu dapat mengalami perubahan pola tidur dan nafsu makan, kelelahan berlebih, serta penurunan energi tanpa sebab yang jelas. Sementara itu, secara sosial, penderita cenderung menarik diri dari lingkungan, menghindari interaksi, dan kehilangan motivasi untuk menjalankan aktivitas sehari-hari (Ramadani dkk, 2024).

Faktor penyebab depresi sangat beragam dan saling berkaitan, mencakup aspek psikososial, kepribadian, pengalaman hidup, dan pola pikir. Menurut Freud (dalam

Ramadani dkk, 2024) kehilangan objek yang dicintai dapat memicu depresi karena individu menginternalisasi rasa kehilangan tersebut secara negatif ke dalam diri. Faktor psikososial seperti kehilangan peran sosial, kematian orang terdekat, penurunan kesehatan, isolasi, masalah ekonomi, dan penurunan fungsi kognitif turut berkontribusi. Dari sisi kepribadian, individu yang bergantung secara emosional lebih rentan mengalami depresi. Pengalaman kegagalan berulang juga dapat menimbulkan rasa tidak berdaya, sedangkan pola pikir negatif, distorsi penilaian, pesimisme, dan penilaian diri yang buruk memperkuat perasaan putus asa yang menjadi ciri khas depresi (Ramadani dkk, 2024).

2.10 Kajian Integrasi Topik dengan Al-Qur'an

Dalam perspektif Islam, depresi tidak hanya dipandang sebagai gangguan psikologis semata, tetapi juga sebagai bagian dari dinamika spiritual manusia dalam menghadapi ujian kehidupan. Al-Qur'an dan hadis memberikan landasan komprehensif tentang bagaimana seorang Muslim seharusnya menyikapi kesedihan, kegelisahan, dan keputusasaan yang sering kali menyertai kondisi depresi. Meskipun istilah “depresi” tidak disebutkan secara eksplisit dalam Al-Qur'an, konsep-konsep yang berkaitan dengan kesedihan (*huzn*), kesulitan/kesusahan (*dlaiq*), serta ujian (*ibtilā'*) dibahas secara mendalam dan menjadi pedoman bagi umat Islam dalam mengelola tekanan psikologis (Hasanah, 2024).

Salah satu ayat yang sangat relevan dalam konteks ini adalah firman Allah SWT dalam QS. Al-Baqarah ayat 155 yang berbunyi:

وَلَنَبْلُوَنَّكُمْ بِشَيْءٍ مِّنَ الْخَوْفِ وَالْجُوعِ وَنَقْصٍ مِّنَ الْأَمْوَالِ وَالْأَنْفُسِ وَالتَّمَرُّتِ وَبَشِيرِ الصَّابِرِينَ ﴿١٥٥﴾

Artinya: ”Kami pasti akan mengujimu dengan sedikit ketakutan dan kelaparan, kekurangan harta, jiwa, dan buah-buahan. Sampaikanlah (wahai Nabi

Muhammad,) kabar gembira kepada orang-orang sabar.” (QS. Al-Baqarah ayat 155) (Kementerian Agama RI, 2022)

Menurut Tafsir Tahlili, Allah SWT pasti menguji kaum Muslimin dengan berbagai kesulitan seperti rasa takut, kelaparan, kekurangan harta, kehilangan jiwa, dan berkurangnya hasil panen. Ujian ini bukan untuk memberatkan, tetapi untuk membentuk kepribadian yang kuat, meneguhkan iman, serta melatih ketabahan dan kesabaran. Melalui ujian tersebut, seseorang menjadi lebih matang secara spiritual dan mental, hingga layak disebut sebagai orang sabar yang akan menerima kabar gembira dari Allah SWT (Kementerian Agama RI, 2022). Dalam konteks pelajar, tekanan akademik, kegagalan, atau kecemasan akan masa depan dapat dipandang sebagai bentuk ujian yang menuntut kesabaran serta keteguhan iman untuk menguatkan kondisi psikologis mereka.

Selain kesabaran dalam menghadapi ujian, Al-Qur'an juga memberikan penguatan spiritual melalui janji bahwa setiap kesulitan pasti akan diiringi oleh kemudahan. Hal ini sebagaimana difirmankan Allah SWT dalam QS. Al-Insyirah ayat 5-6:

فَإِنَّ مَعَ الْعُسْرِ يُسْرًا ۖ إِنَّ مَعَ الْعُسْرِ يُسْرًا ۚ

Artinya: *”Maka, sesungguhnya beserta kesulitan ada kemudahan. Sesungguhnya beserta kesulitan ada kemudahan.”* (QS. Al-Insyirah ayat 5-6) (Kementerian Agama RI, 2022)

Menurut Tafsir Tahlili, ayat ini menegaskan bahwa setiap kesulitan selalu disertai kemudahan. Manusia diajarkan untuk tetap optimis, sabar, dan bertawakal kepada Allah dalam menghadapi rintangan. Kesulitan bukan akhir, tetapi jalan menuju kemudahan yang telah Allah siapkan. Dalam dakwah Nabi Muhammad SAW, meskipun beliau mendapat penentangan, dengan sabar dan ikhlas Allah memberi pertolongan, pengikut setia, dan kemenangan. Pengulangan ayat ini menegaskan

bahwa kemudahan pasti datang bagi siapa pun yang menghadapi kesulitan dengan tekad, kesabaran, dan kepercayaan kepada Allah (Kementerian Agama RI, 2022). Pesan ini relevan bagi pelajar agar tidak terjebak dalam keputusan saat menghadapi tekanan mental atau kesulitan akademik, karena setiap tantangan selalu menyimpan peluang dan kemudahan jika dihadapi dengan sabar.

Al-Qur'an juga menunjukkan bahwa ketenangan batin sejati bersumber dari kedekatan spiritual dengan Allah SWT. Hal ini sebagaimana dijelaskan dalam QS. Ar-Ra'd ayat 28:

الَّذِينَ آمَنُوا وَتَطْمَئِنُّ قُلُوبُهُمْ بِذِكْرِ اللَّهِ أَلَا بِذِكْرِ اللَّهِ تَطْمَئِنُّ الْقُلُوبُ ﴿٢٨﴾

Artinya: *"(Yaitu) orang-orang yang beriman dan hati mereka menjadi tenteram dengan mengingat Allah. Ingatlah, bahwa hanya dengan mengingat Allah hati akan selalu tenteram."* (QS. Ar-Ra'd ayat 28) (Kementerian Agama RI, 2022)

Ayat ini mengajarkan bahwa salah satu cara meredakan tekanan psikologis adalah dengan memperkuat hubungan spiritual melalui zikir, doa, dan keikhlasan. Ketenangan batin yang lahir dari kedekatan dengan Allah menjadi kekuatan penting dalam menghadapi stres dan depresi.

Selanjutnya, Allah SWT menegaskan bahwa orang-orang beriman yang dekat kepada-Nya akan terbebas dari rasa takut dan kesedihan, sebagaimana firman-Nya dalam QS. Yunus ayat 62-63:

أَلَا إِنَّ أَوْلِيَاءَ اللَّهِ لَا خَوْفٌ عَلَيْهِمْ وَلَا هُمْ يَحْزَنُونَ ﴿٦٢﴾ الَّذِينَ ءَامَنُوا وَكَانُوا يَتَّقُونَ ﴿٦٣﴾

Artinya: *"Ingatlah! Wali-wali Allah itu, tidak ada rasa takut pada mereka dan mereka tidak bersedih hati. (Yaitu) orang-orang yang beriman dan senantiasa bertakwa."* (QS. Yunus ayat 62-63) (Kementerian Agama RI, 2022)

Menurut Tafsir Al-Mishbah, ayat ini menggambarkan ketenangan batin para wali Allah yang senantiasa beriman dan bertakwa, sehingga mereka terbebas dari kecemasan terhadap masa depan maupun kesedihan atas masa lalu (Shihab, 2002).

Nilai ini menegaskan bahwa keteguhan iman dan ketakwaan dapat menjadi sumber ketenangan dan daya tahan mental, sesuatu yang sangat dibutuhkan pelajar dalam menghadapi tekanan akademik maupun kehidupan modern yang penuh tuntutan.

Sebagai bentuk keseimbangan antara konsep *khauf* (rasa takut) dan *rajā'* (harapan), Allah SWT berfirman dalam QS. Az-Zumar ayat 53:

قُلْ يٰٓعِبَادِىَ الَّذِيْنَ اَسْرَفُوْا عَلٰٓى اَنْفُسِهِمْ لَا تَقْنَطُوْا مِنْ رَّحْمَةِ اللّٰهِ ۚ اِنَّ اللّٰهَ يَغْفِرُ الذُّنُوْبَ جَمِيْعًا ۚ اِنَّهٗ هُوَ الْغَفُوْرُ ﴿٥٣﴾

Artinya: "Katakanlah, "Wahai hamba-hamba-Ku yang melampaui batas terhadap diri mereka sendiri! Janganlah kamu berputus asa dari rahmat Allah. Sesungguhnya Allah mengampuni dosa-dosa semuanya. Sungguh, Dialah Yang Maha Pengampun, Maha Penyayang." (QS. Az-Zumar ayat 53) (Kementerian Agama RI, 2022)

Menurut Tafsir Al-Azhar jilid 8, ayat ini menunjukkan betapa luasnya rahmat Allah, yang tetap membuka pintu ampunan bagi siapa pun yang bertaubat dan berusaha memperbaiki diri, meskipun dosa dan kesalahannya sangat besar. Ayat ini menjadi simbol nilai *rajā'* (harapan kepada Allah) yang mengajarkan manusia untuk tidak kehilangan optimisme spiritual dalam menghadapi tekanan dan kesalahan (Amrullah, n.d.). Dalam konteks pelajar, pesan ini dapat dimaknai sebagai dorongan untuk tidak tenggelam dalam rasa bersalah, stres, atau keputusasaan, melainkan untuk bangkit dengan semangat baru melalui taubat, usaha, dan keyakinan akan kasih sayang Allah SWT.

2.11 Kajian Topik dengan Teori Pendukung

Algoritma *Fuzzy Gustafson-Kessel* (FGK) merupakan salah satu metode *fuzzy clustering* yang dikembangkan untuk mengatasi keterbatasan *Fuzzy C-Means* (FCM), terutama ketika data memiliki pola sebaran yang tidak berbentuk sferis. Melalui penerapan matriks kovarian adaptif, FGK mampu membentuk *cluster* dengan bentuk elips yang memiliki ukuran, orientasi, dan arah penyebaran yang

berbeda. Fleksibilitas ini menjadikan FGK lebih sesuai untuk data yang distribusinya tidak seragam, termasuk data sosial dan psikologis yang sering kali tidak memiliki pemisahan *cluster* yang tegas.

Dalam proses *clustering* menggunakan FGK, langkah pertama adalah penentuan parameter awal. Parameter yang ditentukan mencakup jumlah *cluster* yang akan dibentuk, nilai pangkat pembobot (*weighting exponent*), nilai ambang konvergensi, batas maksimum iterasi, serta nilai awal fungsi objektif. Penentuan parameter awal ini menjadi langkah penting karena akan menentukan bagaimana algoritma berjalan, arah pencarian solusinya, serta kestabilan proses iterasinya.

Setelah parameter ditetapkan, FGK memulai proses *clustering* dengan melakukan inisialisasi matriks derajat keanggotaan awal yang dibentuk secara acak. Nilai ini kemudian dinormalisasi sehingga setiap data memiliki total keanggotaan bernilai satu sebagaimana dijelaskan pada persamaan (2.1). Langkah ini menjadi dasar utama dalam membentuk struktur *cluster* selanjutnya. Tahap berikutnya adalah pembaruan pusat *cluster*, yang dihitung sebagai rata-rata berbobot berdasarkan derajat keanggotaan setiap data. Perhitungan pusat *cluster* mengikuti ketentuan pada persamaan (2.2), di mana data yang lebih dekat dengan suatu *cluster* akan memberikan kontribusi lebih besar terhadap posisi pusat *cluster* tersebut.

Setelah pusat *cluster* diperoleh, algoritma membentuk matriks kovarian untuk masing-masing *cluster* sesuai pada persamaan (2.3). Matriks kovarian ini menggambarkan pola penyebaran data di sekitar pusat *cluster*, sehingga FGK dapat mengenali apakah *cluster* berbentuk memanjang, melebar, atau mengarah pada orientasi tertentu. Tahap ini merupakan keunggulan utama FGK dibandingkan metode *fuzzy clustering* lainnya. Dari matriks kovarian tersebut, FGK menyusun

matriks pembentuk *norm* (*norm-inducing matrix*) sebagaimana ditetapkan dalam persamaan (2.4). Matriks ini digunakan untuk menghitung jarak antar data dan pusat *cluster*. Jarak tersebut dievaluasi menggunakan pendekatan *Mahalanobis* yang dimodifikasi, sesuai dengan proses yang dirumuskan pada persamaan (2.5), sehingga pengukuran kedekatan tidak lagi bergantung pada asumsi bahwa *cluster* harus berbentuk sferis.

Berdasarkan jarak yang diperoleh, FGK melakukan pembaruan derajat keanggotaan untuk setiap data sebagaimana dalam persamaan (2.6). Pada tahap ini proses *clustering* berlangsung secara sistematis, data yang berada lebih dekat dengan suatu *cluster* akan memiliki keanggotaan yang lebih tinggi, sedangkan data yang lebih jauh akan memperoleh nilai keanggotaan yang lebih kecil. Siklus ini secara bertahap menyempurnakan bentuk *cluster* dari iterasi ke iterasi. Seluruh tahapan pembaruan pusat *cluster*, pembentukan kovarian, perhitungan jarak, dan pembaruan keanggotaan dilakukan secara iteratif hingga perubahan nilai keanggotaan antar iterasi berada di bawah ambang konvergensi atau hingga batas iterasi tercapai, seperti dijelaskan oleh kriteria penghentian pada persamaan (2.7). Ketika kondisi tersebut terpenuhi, proses *clustering* dianggap telah stabil dan hasil *cluster* merupakan struktur akhir yang diperoleh FGK.

BAB III

METODE PENELITIAN

3.1 Jenis Penelitian

Jenis penelitian yang digunakan adalah penelitian kuantitatif deskriptif. Penelitian ini mengolah data numerik berupa variabel risiko depresi pelajar dan hasil implementasi algoritma *Fuzzy Gustafson-Kessel*. Analisis dilakukan secara deskriptif menggunakan *Xie-Beni Index* untuk menggambarkan kualitas hasil *clustering* dan penentuan parameter *weighting exponent* yang optimal. Penelitian ini juga didukung oleh studi literatur terkait algoritma *fuzzy clustering* dan evaluasi *cluster*.

3.2 Data dan Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari platform *Kaggle*, yaitu *Depression Student Dataset* yang disusun oleh Nurrizky Arum Jatmiko. *Dataset* ini merupakan subset dari *Depression Survey/Dataset for Analysis* yang dikembangkan oleh Suman Sharma, dengan pemilihan variabel yang relevan untuk pelajar serta data yang telah melalui proses pembersihan dari duplikasi dan *missing values*. Survei asli dilakukan secara daring dan anonim pada periode Januari hingga Juni 2023. Variabel-variabel yang digunakan dalam penelitian ini beserta keterangan skala pengukuran dan kategorinya mengacu pada dokumentasi dan deskripsi *dataset* yang disediakan pada platform *Kaggle*, sehingga peneliti menggunakan definisi variabel tersebut tanpa melakukan perubahan. Adapun variabel data yang digunakan dalam penelitian ini disajikan Tabel 3.1.

Tabel 3.1 Variabel *Dataset*

No.	Variabel		Tipe Data	Keterangan
1.	Umur	X_1	Numerik	Rentang umur 18-34 tahun
2.	Tekanan Akademik	X_2	Numerik	Skala 1 (rendah) sampai 5 (tinggi)
3.	Kepuasan Belajar	X_3	Numerik	Skala 1 (rendah) sampai 5 (tinggi)
4.	Jam Belajar	X_4	Numerik	Jumlah jam belajar per hari
5.	Stres Finansial	X_5	Numerik	Skala 1 (rendah) sampai 5 (tinggi)
6.	Durasi Tidur	X_6	Kategori	Kurang dari 5 jam, 5-6 jam, 7-8 jam, Lebih dari 8 jam
7.	Pola Makan	X_7	Kategori	Sehat, Sedang, Tidak Sehat

3.3 Teknik Analisis Data

Teknik analisis data dalam penelitian ini meliputi beberapa tahapan sebagai berikut:

1. Analisis Deskriptif Data Risiko Depresi Pelajar

Data dianalisis secara deskriptif untuk memahami karakteristik awal sebelum dilakukan *preprocessing*.

2. *Preprocessing* Data Risiko Depresi Pelajar

Tahapan *preprocessing* yang dilakukan meliputi:

a. Label *Encoding*

Beberapa variabel kategori diubah menjadi bentuk numerik agar dapat dianalisis secara komputasional.

b. Normalisasi Data

Variabel numerik yang digunakan untuk *clustering* dinormalisasi menggunakan metode *Min-Max* agar berada pada skala yang sama, sehingga tidak ada variabel yang mendominasi proses *clustering*.

4. Implementasi Algoritma *Fuzzy Gustafson-Kessel* (FGK)

Tahapan implementasi algoritma FGK sebagai berikut:

a. Memasukkan data yang akan dikelompokkan ke dalam matriks X . Matriks X memiliki ukuran $n \times p$, di mana n adalah jumlah data keseluruhan, yaitu 502 dan p adalah jumlah variabel numerik yang digunakan, yaitu 7 yang terdiri atas 5 variabel numerik dan 2 variabel kategori yang telah ditransformasikan ke dalam bentuk numerik menggunakan metode label *encoding*. Seluruh variabel tersebut kemudian dinormalisasi sebelum dilakukan proses *clustering*.

b. Menentukan parameter awal:

1) Banyaknya *cluster* (c)

Dipilih dalam rentang 2 hingga 5 untuk menguji beberapa kemungkinan banyaknya *cluster*. Hal ini dilakukan karena dalam analisis *clustering* banyaknya kelompok tidak diketahui sejak awal, sehingga perlu dilakukan percobaan pada beberapa nilai c untuk memperoleh struktur pengelompokan yang paling sesuai dengan karakteristik data.

2) *Weighting exponent* atau pangkat pembobot (m)

Weighting exponent atau pangkat pembobot (m) digunakan untuk mengatur tingkat kefuzzian dalam proses *clustering*. Pada penelitian ini, nilai m divariasikan pada rentang $m > 1$, yaitu $m = 1,1; 1,3; 1,5; 1,7; 1,9$; dan 2,1. Pemilihan nilai awal $m = 1,1$ dilakukan karena nilai tersebut merupakan nilai yang mendekati 1 dan masih memenuhi syarat perhitungan algoritma. Variasi nilai m kemudian disusun secara bertahap dengan interval yang konsisten untuk memudahkan analisis pengaruh perubahan *weighting exponent* terhadap kualitas hasil *clustering*.

- 3) Ambang batas konvergensi (ε) = 0,0001
- 4) Maksimum iterasi (MaxIter) = 1000
- 5) Fungsi objektif ($J_{GK}^{(0)}$) = 0

Nilai tersebut merupakan inisialisasi awal sebelum proses iterasi dilakukan. Selanjutnya, nilai fungsi objektif akan dihitung dan diperbarui pada setiap iterasi hingga memenuhi kriteria konvergensi.

- 6) Iterasi awal (t) = 1

- c. Inisialisasi matriks keanggotaan awal (U_0), di mana derajat keanggotaan awal $\mu_{i,k}$ dibangkitkan secara acak dan dinormalisasi per baris sehingga setiap baris berjumlah 1 menggunakan persamaan (2.2).
- d. Menentukan pusat *cluster* dengan persamaan (2.3).
- e. Menentukan matriks kovarian untuk setiap *cluster* dengan persamaan (2.4).
- f. Menentukan jarak *Mahalanobis* dengan *norm inducing matrix* (A_k) menggunakan persamaan (2.5).
- g. Menentukan fungsi objektif pada setiap iterasi menggunakan persamaan (2.6).
- h. Memperbarui matriks keanggotaan dengan persamaan (2.7).
- i. Menentukan kriteria konvergensi, jika $\|U_{t+1} - U_t\| < \varepsilon$ atau ($t > \text{MaxIter}$) maka sudah konvergen. Namun, Jika $\|U_{t+1} - U_t\| \geq \varepsilon$ kembali ke langkah d.

5. Evaluasi Hasil *Clustering* dengan *Xie-Beni Index*

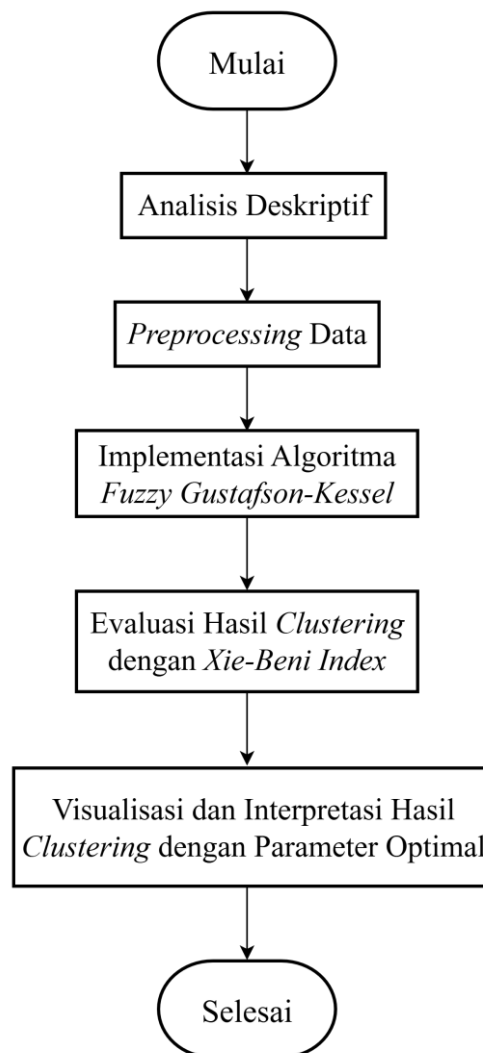
Evaluasi hasil *clustering* dilakukan menggunakan *Xie-Beni Index* (XB) untuk menilai kualitas *cluster* dan penentuan nilai m yang optimal dilakukan

berdasarkan XB, sedangkan variasi c hanya digunakan sebagai prosedur validitas.

6. Visualisasi dan Interpretasi Hasil *Clustering* dengan Parameter Optimal

Parameter optimal yang diperoleh digunakan untuk visualisasi hasil *clustering* menggunakan metode *Principal Component Analysis* (PCA) guna melihat sebaran dan keterpisahan antarklaster. Selanjutnya, interpretasi dilakukan melalui analisis nilai rata-rata variabel pada tiap *cluster* untuk menentukan karakteristik dan tingkat risiko depresi pelajar.

Adapun Gambar 3.1 menunjukkan diagram alir dari penelitian ini.



Gambar 3.1 Diagram Alir Penelitian

Keterangan:



: Simbol terminator digunakan untuk menunjukkan awal dan akhir dari suatu alur proses atau algoritma.



: Simbol proses digunakan untuk menunjukkan tahapan atau aktivitas pengolahan data yang dilakukan secara berurutan dalam penelitian.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Analisis Deskriptif Data Risiko Depresi Pelajar

Analisis deskriptif dilakukan untuk memberikan gambaran umum mengenai karakteristik data yang digunakan dalam penelitian ini. Jumlah data yang digunakan sebanyak 502 responden dengan total tujuh variabel, yang terdiri dari lima variabel numerik, yaitu Umur (X_1), Tekanan Akademik (X_2), Kepuasan Belajar (X_3), Jam Belajar (X_4), dan Stres Finansial (X_5). Adapun untuk dua variabel kategori yaitu, Durasi Tidur (X_6) dan Pola Makan (X_7).

Analisis deskriptif pada variabel numerik disajikan menggunakan ukuran pemusatan dan penyebaran data, yaitu *mean*, *median*, nilai minimum, maksimum, serta standar deviasi. Sementara itu, variabel kategori dianalisis menggunakan distribusi frekuensi dan persentase untuk menunjukkan proporsi masing-masing kategori. Berikut ini disajikan analisis deskriptif untuk variabel numerik pada Tabel 4.1.

Tabel 4.1 Analisis Deskriptif Variabel Numerik

Variabel	<i>Mean</i>	<i>Median</i>	Standar Deviasi	<i>Minimum</i>	<i>Maximum</i>
X_1	26,24	26,5	4,90	18	34
X_2	3,00	3,00	1,39	1	5
X_3	3,08	3,00	1,37	1	5
X_4	6,40	7,00	3,74	0	12
X_5	2,93	3,00	1,43	1	5

Berdasarkan hasil analisis deskriptif pada Tabel 4.1, variabel Umur (X_1) memiliki nilai rata-rata sebesar 26,24 tahun dengan rentang usia 18 hingga 34 tahun. Nilai standar deviasi sebesar 4,90 variasi usia responden berada pada tingkat yang cukup beragam. Variabel Tekanan Akademik (X_2) memiliki rata-rata 3,00

yang berada di sekitar nilai tengah skala penilaian 1-5. Hal ini menunjukkan bahwa sebagian besar responden mengalami tingkat tekanan akademik yang sedang. Variabel Kepuasan Belajar (X_3) memiliki nilai rata-rata sebesar 3,08 yang mengindikasikan bahwa tingkat kepuasan belajar responden berada pada tingkat menengah.

Pada variabel Jam Belajar (X_4), diperoleh rata-rata waktu belajar sebesar 6,40 jam per hari dengan standar deviasi 3,74 menandakan adanya waktu variasi belajar yang cukup besar antar responden, mulai dari 0-12 jam per hari. Sementara itu, variabel Stres Finansial (X_5) memiliki rata-rata sebesar 2,93 dengan standar deviasi 1,43 menunjukkan bahwa tingkat stres finansial berada pada kategori sedang dengan keragaman jawaban yang cukup tinggi.

Selain variabel numerik, penelitian ini juga memiliki dua variabel kategori, yaitu Durasi Tidur (X_6) dan Pola Makan (X_7). Hasil analisis ditampilkan pada Tabel 4.2.

Tabel 4.2 Analisis Deskriptif Variabel Kategori

Variabel	Kategori	Frekuensi	Persentase
X_6	Kurang dari 5 jam	123	25%
	5-6 jam	123	25%
	7-8 jam	128	25%
	Lebih dari 8 jam	128	25%
Total		502	100%
X_7	Sehat	161	32%
	Sedang	172	34%
	Tidak Sehat	169	34%
Total		502	100%

Berdasarkan hasil distribusi frekuensi pada Tabel 4.2, variabel Durasi Tidur (X_6) menunjukkan pola distribusi yang seimbang. Setiap kategori memiliki persentase yang hampir sama, yaitu sekitar 25% dari total responden. Hal ini mengindikasikan bahwa pola durasi tidur responden cenderung beragam dan tidak

terpusat pada kategori tertentu. Dengan kata lain, tidak terdapat kecenderungan mayoritas responden berada dalam kategori tidur terlalu sedikit maupun terlalu lama.

Pada variabel Pola Makan (X_7), proporsi responden menunjukkan pola yang sedikit berbeda. Sebanyak 34% responden berada pada kategori sedang, diikuti kategori tidak sehat sebesar 34%, dan sehat sebanyak 32%. Meskipun selisih tiap kategori tidak terlalu jauh, temuan ini menunjukkan bahwa sebagian besar responden belum sepenuhnya menerapkan pola makan sehat, dengan lebih dari setengah responden berada pada kategori sedang dan tidak sehat.

4.2 *Preprocessing* Data Risiko Depresi Pelajar

Tahap *preprocessing* data dilakukan sebagai langkah awal untuk mempersiapkan data mentah sebelum dianalisis. Proses ini penting karena data biasanya memiliki perbedaan skala maupun format, sehingga perlu diseragamkan agar analisis yang dihasilkan lebih tepat. Sebagai bagian dari tahap ini, sebagian data awal ditampilkan untuk memberikan gambaran mengenai struktur serta karakteristik *dataset* sebelum dilakukan proses transformasi lebih lanjut. Data tersebut disajikan pada Tabel 4.3.

Tabel 4.3 Tampilan *Head Dataset*

Umur	Tekanan Akademik	Kepuasan Belajar	Jam Belajar	Stres Finansial	Durasi Tidur	Pola Makan
28	2	4	9	2	7-8 jam	Sedang
28	4	5	7	1	5-6 jam	Sehat
25	1	3	10	4	5-6 jam	Tidak Sehat
23	1	4	7	2	Lebih dari 8 jam	Tidak Sehat
31	1	5	4	2	Lebih dari 8 jam	Sehat

Berdasarkan Tabel 4.3, beberapa variabel seperti Umur, Tekanan Akademik, Kepuasan Belajar, Jam Belajar, dan Stres Finansial merupakan data numerik dengan skala dan rentang nilai yang berbeda antarvariabel. Sementara itu, variabel Durasi Tidur dan Pola Makan masih berbentuk data kategorikal, sehingga memerlukan proses transformasi ke bentuk numerik sebelum dilakukan analisis lebih lanjut. Kondisi ini menunjukkan bahwa *dataset* belum sepenuhnya siap untuk proses *clustering* dan menegaskan pentingnya tahap *preprocessing*, khususnya normalisasi data numerik dan *label encoding* pada data kategorikal.

4.2.1 Label Encoding

Pada tahap ini dilakukan transformasi data terhadap dua variabel kategori, yaitu Durasi Tidur dan Pola Makan agar keduanya dapat digunakan dalam proses *clustering*. Kedua variabel tersebut awalnya berbentuk data teks, sehingga tidak dapat langsung diproses karena algoritma *Fuzzy Gustafson-Kessel* hanya menerima data dalam bentuk numerik. Oleh karena itu, penelitian ini menerapkan metode *label encoding*, yaitu teknik yang mengubah kategori menjadi nilai bilangan bulat tanpa menambah jumlah variabel atau fitur baru.

Pada variabel Durasi Tidur, kategori disusun berdasarkan urutan dari durasi tidur terpendek hingga yang terpanjang. Variabel Pola Makan juga diberikan label berdasarkan urutan kualitas pola makan, dimulai dari kategori terbaik hingga yang kurang baik. Oleh karena itu, pemberian label dilakukan secara berurutan mulai dari 0, seperti ditunjukkan pada Tabel 4.4.

Tabel 4.4 Skema Label *Encoding*

Variabel	Kategori	Kode
Durasi Tidur	Kurang dari 5 jam	0
	5-6 jam	1
	7-8 jam	2
	Lebih dari 8 jam	3
Pola Makan	Sehat	0
	Sedang	1
	Tidak Sehat	2

Untuk memberikan gambaran perubahan format data, contoh *dataset* sebelum proses label *encoding* dapat dilihat pada Tabel 4.3, sedangkan hasil transformasi setelah *encoding* dapat dilihat pada Tabel 4.5. Melalui proses ini, seluruh variabel dalam *dataset* telah berada dalam format numerik dan siap digunakan pada tahap selanjutnya.

Tabel 4.5 Contoh *Dataset* setelah Label *Encoding*

X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇
28	2	4	9	2	2	1
28	4	5	7	1	1	2
25	1	3	10	4	1	0
23	1	4	7	2	3	0
31	1	5	4	2	3	2

4.2.2 Normalisasi Data

Pada tahap ini dilakukan normalisasi terhadap seluruh variabel numerik dalam *dataset* agar memiliki rentang nilai yang seragam sebelum diproses ke tahap *clustering*. Normalisasi diperlukan karena setiap variabel memiliki satuan dan skala yang berbeda sehingga dapat menyebabkan bias pada perhitungan jarak dalam algoritma *Fuzzy Gustafson-Kessel*. Metode yang digunakan adalah *Min-Max Normalization*, yaitu teknik yang mengubah nilai setiap variabel ke dalam rentang antara 0 hingga 1 berdasarkan nilai minimum dan maksimum dari variabel tersebut. Proses transformasi nilai dilakukan merujuk pada persamaan (2.8).

Untuk memberikan gambaran mengenai penerapan metode *Min-Max Normalization*, berikut ditunjukkan contoh perhitungan normalisasi pada satu responden (baris data pertama) berdasarkan Tabel 4.5.

1. Variabel Umur

Nilai maksimum = 34

Nilai minimum = 18

Nilai data = 28

$$x' = \frac{28 - 18}{34 - 18} = 0,625$$

2. Variabel Tekanan Akademik

Nilai maksimum = 5

Nilai minimum = 1

Nilai data = 2

$$x' = \frac{2 - 1}{5 - 1} = 0,25$$

3. Variabel Kepuasan Belajar

Nilai maksimum = 5

Nilai minimum = 1

Nilai data = 4

$$x' = \frac{4 - 1}{5 - 1} = 0,75$$

4. Variabel Jam Belajar

Nilai maksimum = 12

Nilai minimum = 0

Nilai data = 9

$$x' = \frac{9 - 0}{12 - 0} = 0,75$$

5. Variabel Stres Finansial

Nilai maksimum = 5

Nilai minimum = 1

Nilai data = 2

$$x' = \frac{2 - 1}{5 - 1} = 0,25$$

6. Variabel Durasi Tidur

Nilai maksimum = 3

Nilai minimum = 0

Nilai data = 2

$$x' = \frac{2 - 0}{3 - 0} \approx 0,6667$$

7. Variabel Pola Makan

Nilai maksimum = 2

Nilai minimum = 0

Nilai data = 1

$$x' = \frac{1 - 0}{2 - 0} = 0,5$$

Setelah perhitungan manual tersebut, metode normalisasi diterapkan pada seluruh *dataset*. Hasil normalisasi dapat dilihat pada Tabel 4.6, yang menampilkan lima data pertama setelah melalui proses *Min-Max Normalization*.

Tabel 4.6 Contoh *Dataset* setelah Normalisasi *Min-Max*

X_1	X_2	X_3	X_4	X_5	X_6	X_7
0,625	0,25	0,75	0,75	0,25	0,6667	0,5
0,625	0,75	1	0,5833	0	0,3333	1
0,4375	0	0,5	0,8333	0,75	0,3333	0
0,3125	0	0,75	0,5833	0,25	1	0
0,8125	0	1	0,3333	0,25	1	1

4.3 Implementasi Algoritma *Fuzzy Gustafson-Kessel*

4.3.1 Perhitungan untuk Variasi Nilai $m = 1, 1$

Data yang digunakan dalam penelitian ini merupakan data risiko depresi pelajar yang telah melalui tahap *preprocessing*, yaitu label *encoding* dan normalisasi *Min-Max*. Setelah melalui tahap normalisasi, data penelitian direpresentasikan ke dalam sebuah matriks X berukuran $n \times p$, di mana n adalah banyak data dan p merupakan banyak variabel yang digunakan dalam analisis.

Pada representasi tersebut, setiap baris menunjukkan satu responden, sedangkan setiap kolom berisi nilai variabel hasil normalisasi. Matriks inilah yang menjadi *input* utama dalam proses *clustering* menggunakan algoritma *Fuzzy Gustafson-Kessel*. Bentuk umum matriks X dinyatakan sebagai berikut:

$$X = \begin{bmatrix} 0,625 & 0,25 & 0,75 & 0,75 & \dots & 0,6667 & 0,5 \\ 0,625 & 0,75 & 1 & 0,5833 & \dots & 0,3333 & 1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0,9375 & 0,75 & 0,75 & 0,6667 & \dots & 1 & 1 \\ 0 & 1 & 0,5 & 0,5 & \dots & 1 & 0 \end{bmatrix}$$

Setelah data direpresentasikan dalam bentuk matriks, proses *clustering* dilanjutkan dengan menentukan parameter algoritma. Pada tahap ini, perhitungan dilakukan menggunakan nilai $m = 1,1$, sedangkan banyaknya *cluster* (c) mengacu pada rentang yang telah ditetapkan sebelumnya, yaitu 2 hingga 5. Seluruh proses perhitungan untuk setiap nilai c dilakukan dengan bantuan program *Python*. Namun demikian, perhitungan manual hanya disajikan untuk kasus $c = 2$ sebagai contoh

ilustratif.

Selepas parameter ditetapkan, proses perhitungan pada algoritma dimulai dan dilakukan secara bertahap sesuai tahapan iterasi dalam metode *Fuzzy Gustafson-Kessel*.

1. Inisialisasi Matriks Keanggotaan Awal

Tahap awal dalam proses *clustering* FGK adalah melakukan pembangkitan matriks keanggotaan awal (U_0). Matriks ini dihasilkan secara acak oleh program dengan syarat bahwa jumlah derajat keanggotaan pada setiap baris harus bernilai 1. Secara umum bentuk matriks keanggotaan awal dinyatakan sesuai pada persamaan (2.2). Pembentukan U_0 dilakukan sebagai nilai awal iterasi dalam algoritma. Dalam penelitian ini, banyaknya *cluster* digunakan sebagai salah satu unsur dalam proses validasi untuk mengevaluasi struktur hasil *clustering* yang diperoleh.

Hasil pembangkitan U_0 untuk beberapa nilai c ditunjukkan sebagai berikut:

Matriks U_0 untuk $c = 2$:

$$U_0 = \begin{bmatrix} 0,6831 & 0,3619 \\ 0,5518 & 0,4482 \\ \vdots & \vdots \\ 0,1193 & 0,8807 \\ 0,4586 & 0,5414 \end{bmatrix}$$

Matriks U_0 untuk $c = 3$:

$$U_0 = \begin{bmatrix} 0,3736 & 0,2119 & 0,4145 \\ 0,3946 & 0,0533 & 0,5521 \\ \vdots & \vdots & \vdots \\ 0,3354 & 0,4795 & 0,1851 \\ 0,2439 & 0,4659 & 0,2902 \end{bmatrix}$$

Matriks U_0 untuk $c = 4$:

$$U_0 = \begin{bmatrix} 0,2795 & 0,1585 & 0,3101 & 0,2519 \\ 0,0360 & 0,3728 & 0,2908 & 0,3004 \\ \vdots & \vdots & \vdots & \vdots \\ 0,2913 & 0,1543 & 0,3294 & 0,2250 \\ 0,0710 & 0,5424 & 0,2998 & 0,0868 \end{bmatrix}$$

Matriks U_0 untuk $c = 5$:

$$U_0 = \begin{bmatrix} 0,2703 & 0,1533 & 0,2999 & 0,2436 & 0,0329 \\ 0,3146 & 0,2454 & 0,2535 & 0,0413 & 0,1452 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0,0122 & 0,2462 & 0,3427 & 0,2989 & 0,0999 \\ 0,0172 & 0,3028 & 0,3017 & 0,1714 & 0,2068 \end{bmatrix}$$

2. Pusat *Cluster*

Setelah diperoleh U_0 dan ditetapkan nilai $m = 1,1$, tahap berikutnya adalah menghitung pusat *cluster* (v_k). Penghitungan pusat *cluster* dilakukan untuk memperoleh representasi setiap *cluster*, yang selanjutnya digunakan dalam proses pembaruan matriks keanggotaan pada iterasi berikutnya. Perhitungan pusat *cluster* mengacu pada rumus yang ditunjukkan pada persamaan (2.3). Berikut disajikan contoh perhitungan manual untuk pusat *cluster* 1 variabel Umur (X_1), yaitu $v_{1,1}$. Perhitungan untuk pusat *cluster* 1 dapat dituliskan sebagai berikut:

$$v_{1,1} = \frac{\sum_{i=1}^n (\mu_{i,1})^{1,1} x_{i,1}}{\sum_{i=1}^n (\mu_{i,1})^{1,1}}$$

$$v_{1,1} = \frac{(\mu_{1,1})^{1,1} x_{1,1} + (\mu_{2,1})^{1,1} x_{2,1} + \dots + (\mu_{502,1})^{1,1} x_{502,1}}{(\mu_{1,1})^{1,1} + (\mu_{2,1})^{1,1} + \dots + (\mu_{502,1})^{1,1}}$$

$$v_{1,1} = \frac{(0,6381)^{1,1} 0,625 + (0,5518)^{1,1} 0,625 + \dots + (0,4586)^{1,1} 0}{(0,6381)^{1,1} + (0,5518)^{1,1} + \dots + (0,4586)^{1,1}}$$

$$v_{1,1} = \frac{117,7563}{227,3088} = 0,5180$$

Perhitungan untuk pusat *cluster* 2:

$$v_{2,1} = \frac{\sum_{i=1}^n (\mu_{i,2})^{1,1} x_{i,1}}{\sum_{i=1}^n (\mu_{i,2})^{1,1}}$$

$$v_{2,1} = \frac{(\mu_{1,2})^{1,1} x_{1,1} + (\mu_{2,2})^{1,1} x_{2,1} + \dots + (\mu_{502,2})^{1,1} x_{502,1}}{(\mu_{1,2})^{1,1} + (\mu_{2,2})^{1,1} + \dots + (\mu_{502,2})^{1,1}}$$

$$v_{2,1} = \frac{(0,3619)^{1,1} 0,625 + (0,4482)^{1,1} 0,625 + \dots + (0,5414)^{1,1} 0}{(0,3619)^{1,1} + (0,4482)^{1,1} + \dots + (0,5414)^{1,1}}$$

$$v_{2,1} = \frac{126.8920}{247.5166} = 0,5127$$

Nilai tersebut merupakan pusat *cluster* untuk variabel Umur (X_1) pada *cluster* 1 dan *cluster* 2. Proses perhitungan yang sama juga diterapkan pada variabel lain dan pada *cluster* berikutnya, sehingga diperoleh pusat *cluster* untuk keseluruhan variabel pada kedua *cluster*. Karena jumlah variabel penelitian sebanyak tujuh, maka setiap pusat *cluster* berbentuk vektor berdimensi tujuh. Berdasarkan hasil pada Tabel 4.7, pusat *cluster* untuk $c = 2$ dapat dinyatakan sebagai berikut:

$$v_1 = (0,5180, 0,4916, 0,5063, 0,5393, 0,4829, 0,4990, 0,5049)$$

$$v_2 = (0,5127, 0,5095, 0,5297, 0,5286, 0,4814, 0,5133, 0,4791)$$

Setiap komponen pada vektor tersebut secara berurutan menyatakan nilai untuk variabel X_1 sampai dengan X_7 . Hasil lengkap pusat dua *cluster* yang terbentuk tetap ditampilkan pada Tabel 4.7 agar perbandingan nilai antarvariabel pada masing-masing *cluster* dapat terlihat dengan lebih jelas.

Tabel 4.7 Hasil Pusat Dua *Cluster* ($m = 1,1$)

Variabel	Pusat <i>Cluster</i>	
	<i>Cluster 1</i>	<i>Cluster 2</i>
X_1	0,5180	0,5127
X_2	0,4916	0,5095
X_3	0,5063	0,5297
X_4	0,5393	0,5286
X_5	0,4829	0,4814
X_6	0,4990	0,5133
X_7	0,5049	0,4791

Setelah proses perhitungan manual diperlihatkan pada contoh untuk $c = 2$, pusat *cluster* untuk banyaknya *cluster* lainnya, yaitu $c = 3$, $c = 4$, dan $c = 5$, dihitung dengan prosedur yang sama sesuai persamaan (2.3). Pada tahap ini, proses komputasi dilakukan dengan bantuan program *Python* dan perangkat lunak *Spreadsheet* untuk mempermudah perhitungan numerik yang bersifat repetitif yaitu, perhitungan yang dilakukan berulang dengan pola yang sama, khususnya pada bagian penjumlahan komponen keanggotaan yang berjumlah besar.

3. Matriks Kovarian

Matriks kovarian (F_k) digunakan untuk merepresentasikan bentuk, arah, dan tingkat penyebaran data di sekitar pusat *cluster*. Dengan demikian, perhitungan matriks kovarian bertujuan untuk menyesuaikan bentuk *cluster* berdasarkan karakteristik data, sehingga algoritma *Fuzzy Gustafson-Kessel* mampu membentuk *cluster* yang tidak terbatas pada bentuk bulat, tetapi dapat memanjang atau melebar sesuai dengan pola data. Perhitungan matriks kovarian dilakukan mengacu pada persamaan (2.4).

Contoh perhitungan manual F_k pada penelitian ini ditunjukkan menggunakan satu data untuk menggambarkan langkah-langkah dasar perhitungan yang

dilakukan dalam algoritma *Fuzzy Gustafson-Kessel*.

$$\begin{aligned}
 (x_1 - v_1) &= [0,625 - 0,5180 \quad 0,25 - 0,4916 \quad 0,75 - 0,5063 \\
 &\quad 0,75 - 0,5393 \quad 0,25 - 0,4829 \quad 0,667 - 0,4990 \quad 0,5 - 0,5049] \\
 &= [0,1070 \quad -0,2416 \quad 0,2437 \quad 0,2107 \quad -0,2329 \quad 0,1677 \quad -0,0049] \\
 (x_1 - v_1)(x_1 - v_1)^T &= [0,1070 \quad -0,2416 \quad 0,2437 \quad 0,2107 \quad -0,2329
 \end{aligned}$$

$$\begin{aligned}
 &\quad 0,1677 \quad -0,0049] \begin{bmatrix} 0,1070 \\ -0,2416 \\ 0,2437 \\ 0,2107 \\ -0,2329 \\ 0,1677 \\ -0,0049 \end{bmatrix}
 \end{aligned}$$

$$= \begin{bmatrix} 0,0114 & -0,0258 & 0,0261 & 0,0225 & -0,0249 & 0,0179 & -0,0005 \\ -0,0258 & 0,0584 & -0,0589 & -0,0509 & 0,0563 & -0,0405 & 0,0012 \\ 0,0261 & -0,0589 & 0,0594 & 0,0513 & -0,0568 & 0,0409 & -0,0012 \\ 0,0225 & -0,0509 & 0,0514 & 0,0443 & -0,0491 & 0,0353 & -0,0010 \\ -0,0249 & 0,0563 & -0,0568 & -0,0491 & 0,0543 & -0,0390 & 0,0011 \\ 0,0179 & -0,0405 & 0,0409 & 0,0353 & -0,0391 & 0,0281 & -0,0008 \\ -0,0005 & 0,0012 & -0,0012 & -0,0010 & 0,0011 & -0,0008 & 0,00002 \end{bmatrix}$$

$$(\mu_{1,1})^{1,1} (x_1 - v_1)(x_1 - v_1)^T = (0,6381)^{1,1} \begin{bmatrix} 0,0114 & \cdots & -0,0005 \\ \vdots & \ddots & \vdots \\ -0,0005 & \cdots & 0,00002 \end{bmatrix}$$

$$= (0,6101) \begin{bmatrix} 0,0114 & \cdots & -0,0005 \\ \vdots & \ddots & \vdots \\ -0,0005 & \cdots & 0,00002 \end{bmatrix}$$

$$= \begin{bmatrix} 0,0070 & -0,0158 & 0,0159 & 0,0138 & -0,0152 & 0,0109 & -0,0003 \\ -0,0158 & 0,0356 & -0,0359 & -0,0311 & 0,0343 & -0,0247 & 0,0007 \\ 0,0159 & -0,0359 & 0,037062 & 0,0313 & -0,0346 & 0,0249 & -0,0007 \\ 0,0137 & -0,0311 & 0,0313 & 0,0271 & -0,0299 & 0,0216 & -0,0006 \\ -0,0152 & 0,0343 & -0,0346 & -0,0299 & 0,0331 & -0,0238 & 0,0007 \\ 0,0109 & -0,0247 & 0,0249 & 0,0216 & -0,0238 & 0,0172 & -0,0005 \\ -0,0003 & 0,0007 & -0,0007 & -0,0006 & 0,0007 & -0,0005 & 0,00001 \end{bmatrix}$$

Matriks kovarian akhir *cluster* 1 kemudian diperoleh dengan menjumlahkan seluruh kontribusi data berdasarkan derajat keanggotaannya. Hasil akhirnya ditunjukkan pada matriks berikut.

$$\begin{aligned}
& \sum_{i=1}^{502} (\mu_{i,1})^{1,1} (x_i - v_1)(x_i - v_1)^T \\
&= \begin{bmatrix} 18,6987 & 0,6777 & -0,9997 & 1,3103 & -2,2135 & -1,1940 & 1,7229 \\ 0,6777 & 25,2073 & -2,8148 & 2,1982 & 0,5596 & -1,1984 & 1,4950 \\ -0,9997 & -2,8148 & 24,0392 & -0,9075 & -0,1293 & 0,2108 & -0,8167 \\ 1,3103 & 2,1982 & -0,9075 & 20,7480 & 0,8448 & -2,1802 & 1,0490 \\ -2,2135 & 0,5596 & -0,1293 & 0,8448 & 25,4174 & -0,8825 & -0,5464 \\ -1,1940 & -1,1984 & 0,2108 & -2,1802 & -0,8825 & 29,8016 & -4,7813 \\ 1,7229 & 1,4950 & -0,8167 & 1,0490 & -0,5464 & -4,7813 & 33,3260 \end{bmatrix} \\
& F_1 = \frac{1}{227,3088} \times \begin{bmatrix} 18,6987 & \dots & 1,7229 \\ \vdots & \ddots & \vdots \\ 1,7229 & \dots & 33,3260 \end{bmatrix} \\
&= \begin{bmatrix} 0,0922 & 0,0033 & -0,0042 & 0,0063 & -0,0111 & -0,0057 & 0,0085 \\ 0,0033 & 0,1236 & -0,0137 & 0,0100 & 0,0025 & -0,0049 & 0,0065 \\ -0,0042 & -0,0137 & 0,1183 & -0,0039 & -0,0011 & 0,0001 & -0,0030 \\ 0,0063 & 0,0100 & -0,0039 & 0,1017 & 0,0039 & -0,0103 & 0,0050 \\ -0,0111 & 0,0025 & -0,0011 & 0,0039 & 0,1254 & -0,0043 & -0,0028 \\ -0,0057 & -0,0049 & 0,0001 & -0,0103 & -0,0043 & 0,1457 & -0,0230 \\ 0,0085 & 0,0065 & -0,0030 & 0,0050 & -0,0028 & -0,0230 & 0,1637 \end{bmatrix}
\end{aligned}$$

Perhitungan manual pada data pertama untuk *cluster* 1 juga diterapkan dengan cara yang sama pada *cluster* 2, menggunakan pusat *cluster* masing-masing. Sehingga diperoleh matriks kovarian F_2 :

$$\begin{aligned}
& F_2 \\
&= \begin{bmatrix} 0,0946 & 0,0025 & 0,0026 & 0,0032 & -0,0095 & -0,0047 & 0,0069 \\ 0,0025 & 0,1176 & -0,0107 & 0,0021 & -0,0016 & 0,0068 & -0,0017 \\ 0,0026 & -0,0107 & 0,1168 & 0,0006 & -0,0080 & -0,0047 & 0,0090 \\ 0,0032 & 0,0021 & 0,0006 & 0,0926 & -0,0004 & -0,0055 & 0,0048 \\ -0,0095 & -0,0016 & -0,0080 & -0,0004 & 0,1277 & -0,0045 & -0,0069 \\ -0,0047 & 0,0068 & -0,0047 & -0,0055 & -0,0045 & 0,1325 & -0,0155 \\ 0,0069 & -0,0017 & 0,0090 & 0,0048 & -0,0069 & -0,0155 & 0,1647 \end{bmatrix}
\end{aligned}$$

Prosedur yang sama juga digunakan untuk *cluster* lainnya $c = 3$, $c = 4$, dan $c = 5$. Dengan langkah tersebut, setiap *cluster* memperoleh matriks kovarian final berdasarkan seluruh data yang termasuk dalam *cluster* tersebut.

4. Jarak *Mahalanobis*

Pada tahap ini dilakukan perhitungan jarak *Mahalanobis* antara setiap data dan pusat masing-masing *cluster*. Perhitungan jarak ini bertujuan untuk memperoleh ukuran kedekatan data terhadap *cluster* dengan mempertimbangkan penyebaran dan arah variansi data melalui matriks kovarian pada setiap *cluster*. Sebelum jarak dihitung, terlebih dahulu dibentuk *norm-inducing matrix* (A_k) menggunakan persamaan,

$$A_k = [\rho_k \det(F_k)^{\frac{1}{p}} F_k^{-1}]$$

yang berfungsi menentukan norma jarak yang sesuai dengan struktur penyebaran data pada *cluster* tersebut. Setelah matriks A_k diperoleh, jarak *Mahalanobis* setiap data dihitung menggunakan persamaan (2.5).

Untuk perhitungan A_k dapat diperlihatkan menggunakan data ke-1 terhadap *cluster* 1 sebagai contoh. Nilai determinan matriks kovarian F_1 yang diperoleh adalah

$$\det(F_1) = 5449094890,027897$$

Karena terdapat tujuh variabel, maka:

$$\det(F_1)^{1/7} = 24,59821$$

Pada tahap ini digunakan parameter $\rho_k = 1$. Pemilihan nilai $\rho_k = 1$ dilakukan karena nilai ini merupakan pilihan standar dalam metode *Gustafson-Kessel*, di mana tujuan utamanya adalah menjaga agar ukuran, bentuk, dan orientasi *cluster* tetap merepresentasikan struktur kovariannya secara alami.

$$F_1^{-1} = \begin{bmatrix} 0,0547 & -0,0009 & 0,0020 & -0,0032 & 0,0049 & 0,0017 & -0,0023 \\ -0,0009 & 0,0407 & 0,0045 & -0,0038 & -0,0008 & 0,0010 & -0,0014 \\ 0,0020 & 0,0045 & 0,0423 & 0,0012 & 0,0003 & 0,0002 & 0,0007 \\ -0,0032 & -0,0038 & 0,0012 & 0,0493 & -0,0017 & 0,0031 & -0,0008 \\ 0,0049 & -0,0008 & 0,0003 & -0,0017 & 0,0399 & 0,0013 & 0,0007 \\ 0,0017 & 0,0010 & 0,0002 & 0,0031 & 0,0013 & 0,0347 & 0,0048 \\ -0,0023 & -0,0014 & 0,0007 & -0,0008 & 0,0007 & 0,0048 & 0,0309 \end{bmatrix}$$

$$A_1 = \rho_i \det(F_1)^{\frac{1}{7}} F_1^{-1}$$

$$A_1 = 1(24,59821) \times \begin{bmatrix} 0,0547 & \dots & -0,0023 \\ \vdots & \ddots & \vdots \\ -0,0023 & \dots & 0,0309 \end{bmatrix}$$

$$= \begin{bmatrix} 1,3423 & -0,0226 & 0,0415 & -0,0777 & 0,1221 & 0,0412 & -0,0574 \\ -0,0226 & 1,0013 & 0,1113 & -0,0883 & -0,0182 & 0,0210 & -0,0311 \\ 0,0415 & 0,1113 & 1,0379 & 0,0260 & 0,0100 & 0,0092 & 0,0134 \\ -0,0777 & -0,0883 & 0,0260 & 1,2137 & -0,0407 & 0,0752 & -0,0195 \\ 0,1221 & -0,0182 & 0,0100 & -0,0407 & 0,9786 & 0,0331 & 0,0172 \\ 0,0412 & 0,0210 & 0,0092 & 0,0752 & 0,0331 & 0,8570 & 0,1158 \\ -0,0574 & -0,031 & 0,0134 & -0,0195 & 0,0172 & 0,1158 & 0,7602 \end{bmatrix}$$

Dengan menggunakan prosedur perhitungan yang sama seperti pada *cluster*

1, diperoleh A_k untuk *cluster* 2, yaitu A_2 sebagai berikut:

$$A_2 = \begin{bmatrix} 1,2643 & -0,0303 & -0,0194 & -0,0380 & 0,0911 & 0,0419 & -0,0436 \\ -0,0303 & 1,0162 & 0,0930 & -0,0255 & 0,0144 & -0,0499 & 0,0034 \\ -0,0194 & 0,0930 & 1,0281 & -0,0034 & 0,0626 & 0,0271 & -0,0490 \\ -0,0380 & -0,0255 & -0,0034 & 1,2808 & 0,0005 & 0,0498 & -0,0309 \\ 0,0911 & 0,0144 & 0,0626 & 0,0005 & 0,9378 & 0,0406 & 0,0362 \\ 0,0419 & -0,0499 & 0,0271 & 0,0498 & 0,0406 & 0,9084 & 0,0821 \\ -0,0436 & 0,0034 & -0,0490 & -0,0309 & 0,0362 & 0,0821 & 0,7310 \end{bmatrix}$$

Setelah memperoleh A_k untuk masing-masing *cluster*, tahap selanjutnya adalah menghitung jarak *Mahalanobis* antara setiap data dengan pusat *cluster*.

$$(x_1 - v_1) = [0,1070 \quad -0,2416 \quad 0,2437 \quad 0,2107 \quad -0,2329 \quad 0,1677 \\ -0,0049]$$

$$(x_1 - v_1)^T = \begin{bmatrix} 0,1070 \\ -0,2416 \\ 0,2437 \\ 0,2107 \\ -0,2329 \\ 0,1677 \\ -0,0049 \end{bmatrix}$$

$$D_{1,1A_1}^2 = \begin{bmatrix} 0,1070 \\ \vdots \\ -0,0049 \end{bmatrix} A_1 [0,1062 \quad \cdots \quad -0,0053]$$

$$D_{1,1A_1}^2 = \begin{bmatrix} 0,1070 \\ \vdots \\ -0,0049 \end{bmatrix} \begin{bmatrix} 1,3423 & \cdots & -0,0574 \\ \vdots & \ddots & \vdots \\ -0,0574 & \cdots & 0,7602 \end{bmatrix} [0,1070 \quad \cdots \quad -0,0049]$$

$$D_{1,1A_1}^2 = 0,2628$$

$$D_{1,1A_1} = 0,5126$$

Dengan cara perhitungan yang sama, berdasarkan matriks A_2 jarak *Mahalanobis* antara data pertama dan pusat *cluster* 2 diperoleh sebagai berikut:

$$D_{1,2A_2}^2 = 0,2574$$

$$D_{1,2A_2} = 0,5073$$

Kemudian untuk jarak *Mahalanobis* antara data pertama dan pusat *cluster* 1 maupun *cluster* 2, hasil tersebut kemudian dirangkum dalam Tabel 4.8.

Tabel 4.8 Hasil Jarak *Mahalanobis* Dua *Cluster* ($m = 1,1$)

Data	Cluster 1	Cluster 2
1	0,5126	0,5073
2	0,8708	0,8161
3	0,8116	0,8069
$\vdots$	$\vdots$	$\vdots$
502	1,0184	0,9648

Prosedur perhitungan jarak *Mahalanobis* untuk $c = 3$, $c = 4$, dan $c = 5$ dilakukan dengan langkah yang sama, yakni menggunakan selisih antara data

dengan pusat masing-masing *cluster* serta matriks A_k yang telah dihitung sebelumnya.

5. Fungsi Objektif

Perhitungan nilai fungsi objektif dilakukan untuk menilai kualitas pembentukan *cluster* pada iterasi ke- t . Nilai ini dihitung menggunakan persamaan (2.6) dan berfungsi sebagai indikator konvergensi dalam proses *clustering*. Hasil perhitungan fungsi objektif dua *cluster* disajikan berikut ini:

Perhitungan untuk *cluster* 1

$$\begin{aligned}\sum (\mu_{i,k})^m D_{i,kA_k}^2 &= (\mu_{1,1})^{1,1} D_{1,1A_1}^2 + \dots + (\mu_{502,1})^{1,1} D_{502,1A_1}^2 \\ &= (0,6381)^{1,1} (0,5126)^2 + \dots + (0,4586)^{1,1} (1,0184)^2 \\ &= 0,1603 + \dots + 0,4400 = 192,3949\end{aligned}$$

Perhitungan untuk *cluster* 2

$$\begin{aligned}\sum (\mu_{i,k})^m D_{i,kA_k}^2 &= (\mu_{1,2})^{1,1} D_{1,2A_2}^2 + \dots + (\mu_{502,2})^{1,1} D_{502,2A_2}^2 \\ &= (0,3619)^{1,1} (0,5073)^2 + \dots + (0,5414)^{1,1} (0,9648)^2 \\ &= 0,0841 + \dots + 0,4739 = 204,3948\end{aligned}$$

Maka untuk fungsi objektif pada iterasi ke-1 ini adalah:

$$\begin{aligned}J_{GK}^{(1)} &= \sum_{i=1}^{502} \sum_{k=1}^2 (\mu_{i,k})^{1,1} D_{i,kA_k}^2 \\ &= 192,3949 + 204,3948 = 396,7897\end{aligned}$$

Perhitungan fungsi objektif pada iterasi berikutnya dilakukan dengan langkah yang sama seperti pada iterasi awal, kemudian diperoleh nilai fungsi objektif baru untuk iterasi tersebut. Untuk $c = 3$, $c = 4$, dan $c = 5$, proses perhitungannya tetap mengikuti prosedur yang sama.

6. Matriks Keanggotaan Baru

Matriks keanggotaan baru diperoleh melalui pembaruan nilai derajat keanggotaan setiap data terhadap masing-masing *cluster*. Pembaruan ini dilakukan dengan mengacu pada persamaan (2.7), sehingga dihasilkan matriks U_{t+1} sebagai representasi keanggotaan pada iterasi berikutnya. Pada tahap ini disajikan contoh perhitungan pembaruan derajat keanggotaan untuk data pertama, dengan diketahui nilai jarak *Mahalanobis* data tersebut terhadap masing-masing *cluster* sebagai berikut:

$$D_{1,1A_1} = 0,5126 \quad D_{1,2A_2} = 0,5073$$

Dengan nilai $m = 1,1$, maka nilai pangkat:

$$\frac{2}{m-1} = \frac{2}{1,1-1} = 20$$

Derajat keanggotaan data pertama terhadap *cluster* 1

$$\begin{aligned} \mu_{1,1}^{(t+1)} &= \left[\left(\frac{D_{1,1A_1}}{D_{1,1A_1}} \right)^{20} + \left(\frac{D_{1,1A_1}}{D_{1,2A_2}} \right)^{20} \right]^{-1} \\ &= \left[1 + \left(\frac{0,5126}{0,5073} \right)^{20} \right]^{-1} = [1 + 1,2298]^{-1} \\ &= 0,4485 \end{aligned}$$

Derajat keanggotaan data pertama terhadap *cluster* 2

$$\begin{aligned} \mu_{1,2}^{(t+1)} &= \left[\left(\frac{D_{1,2A_2}}{D_{1,1A_1}} \right)^{20} + \left(\frac{D_{1,2A_2}}{D_{1,2A_2}} \right)^{20} \right]^{-1} \\ &= \left[\left(\frac{0,5073}{0,5126} \right)^{20} + 1 \right]^{-1} = [0,8131 + 1]^{-1} \\ &= 0,5515 \end{aligned}$$

Beberapa hasil pembaruan derajat keanggotaan untuk dua *cluster* ditampilkan pada Tabel 4.9.

Tabel 4.9 Hasil Matriks Keanggotaan Baru dengan Dua *Cluster* ($m = 1,1$)

Data	Cluster 1	Cluster 2
1	0,4485	0,5515
2	0,2144	0,7856
3	0,4713	0,5287
$\vdots$	$\vdots$	$\vdots$
502	0,2531	0,7469

Perhitungan untuk banyaknya *cluster* $c = 3$, $c = 4$, dan $c = 5$ dilakukan dengan prosedur yang sama seperti pada perhitungan dengan $c = 2$.

7. Kriteria Konvergensi

Pemeriksaan konvergensi dilakukan dengan membandingkan perubahan matriks keanggotaan pada dua iterasi berturutan, yaitu U_t dan U_{t+1} . Matriks keanggotaan sendiri dinotasikan sebagai $U = [\mu_{i,k}]$, yang berisi nilai derajat keanggotaan setiap data ke dalam masing-masing *cluster*. Konvergensi dicapai apabila perubahan antara U_t dan U_{t+1} sangat kecil, sehingga dianggap bahwa algoritma tidak lagi mengalami pembaruan yang signifikan. Oleh karena itu, perhitungan konvergensi dilakukan dengan menggunakan norma selisih matriks $\|U_{t+1} - U_t\|$, kemudian dibandingkan dengan nilai *threshold* (ϵ). Pada bagian ini ditunjukkan contoh pemeriksaan konvergensi menggunakan nilai matriks keanggotaan pada iterasi awal.

$$\|U_{t+1} - U_t\| = \sqrt{\sum_{i=1}^n \sum_{k=1}^c (U_{t+1} - U_t)^2}$$

Perhitungan untuk *cluster* 1:

$$\|U_{t+1} - U_t\| = \sqrt{(0,4485 - 0,6381)^2 + \dots + (0,2531 - 0,4586)^2}$$

$$= \sqrt{31,6428}$$

Perhitungan untuk *cluster* 2:

$$\begin{aligned} \|U_{t+1} - U_t\| &= \sqrt{(0,5515 - 0,3619)^2 + \dots + (0,7469 - 0,5414)^2} \\ &= \sqrt{31,6428} \end{aligned}$$

Karena perhitungan dilakukan dengan menjumlahkan hasil dari seluruh *cluster*, maka nilai akhirnya diperoleh sebagai berikut:

$$\begin{aligned} \|U_{t+1} - U_t\| &= \sqrt{31,6428 + 31,6428} = \sqrt{63,2856} \\ \|U_{t+1} - U_t\| &= 7,9552 \end{aligned}$$

Nilai tersebut kemudian dibandingkan dengan nilai *threshold* (ε) yang digunakan dalam penelitian, yaitu $\varepsilon = 0,0001$. Karena 7,9552 jauh lebih besar dari ε , maka pada tahap ini algoritma belum mencapai kondisi konvergen. Oleh karena itu, proses iterasi dilanjutkan kembali dengan memperbarui parameter-parameter pada setiap langkah. Proses pembaruan terus dilakukan hingga selisih perubahan derajat keanggotaan menjadi sangat kecil. Setelah iterasi berjalan hingga mencapai iterasi ke-54, diperoleh nilai sebagai berikut:

$$\|U_{t+1} - U_t\| = 0,000091523$$

Nilai ini sudah berada di bawah ambang batas $\varepsilon = 0,0001$, sehingga pada iterasi ke-54 algoritma dinyatakan telah konvergen dan proses *clustering* dapat dihentikan.

8. Hasil *Clustering*

Setelah algoritma *Fuzzy Gustafson-Kessel* dengan $m = 1,1$ mencapai konvergensi, penentuan hasil *clustering* data didasarkan pada matriks keanggotaan akhir. Setiap data dikelompokkan ke dalam *cluster* yang

memiliki nilai derajat keanggotaan terbesar, sehingga *cluster* tersebut menjadi *cluster* dominan yang paling merepresentasikan data. Hasil *clustering* data berdasarkan *cluster* dominan untuk $c = 2$ dan $m = 1,3$ disajikan pada Tabel 4.10.

Tabel 4.10 Hasil *Clustering* Dua *Cluster* ($m = 1,1$)

Data	Derajat Keanggotaan		Cluster
	1	2	
1	0,00000006	0,99999994	2
2	0,00000003	0,99999997	2
3	0,8820	0,1180	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$
502	0,9592	0,0408	1

4.3.2 Perhitungan untuk Variasi Nilai $m = 1,3$

Pada variasi ini, proses *clustering* dilakukan dengan menggunakan nilai $m = 1,3$. Seluruh tahapan mengikuti langkah-langkah yang telah dijelaskan pada bagian sebelumnya $m = 1,1$. Sebagaimana telah dijelaskan pada subbab sebelumnya, proses perhitungan manual tetap difokuskan pada kasus $c = 2$ sebagai contoh perhitungan. Walaupun demikian, seluruh tahapan perhitungan yang dilakukan pada masing-masing nilai c menetapkan prosedur yang sama.

1. Inisialisasi Matriks Keanggotaan Awal

Pada variasi nilai $m = 1,3$ proses diawali dengan inisialisasi U_0 . Matriks U_0 yang digunakan sama seperti pada perhitungan $m = 1,1$ karena pembangkitan U_0 tidak dipengaruhi oleh nilai m . Bentuk U_0 untuk masing-masing banyaknya *cluster* $c = 2$ sampai $c = 5$ tetap menggunakan hasil inisialisasi awal yang telah ditampilkan pada subbab sebelumnya.

2. Pusat *Cluster*

Perhitungan pusat *cluster* awal dilakukan dengan menggunakan matriks U_0 .

Sebagai contoh manual, perhitungan ditunjukkan untuk $c = 2$ menggunakan variabel pertama, yaitu Umur (X_1).

Perhitungan pusat *cluster* 1 dapat dituliskan sebagai berikut:

$$v_{1,1} = \frac{\sum_{i=1}^n (\mu_{i,1})^{1,3} x_{i,1}}{\sum_{i=1}^n (\mu_{i,1})^{1,3}}$$

$$v_{1,1} = \frac{(\mu_{1,1})^{1,3} x_{1,1} + (\mu_{2,1})^{1,3} x_{2,1} + \dots + (\mu_{502,1})^{1,3} x_{502,1}}{(\mu_{1,1})^{1,3} + (\mu_{2,1})^{1,3} + \dots + (\mu_{502,1})^{1,3}}$$

$$v_{1,1} = \frac{(0,6381)^{1,3} 0,625 + (0,5518)^{1,3} 0,625 + \dots + (0,4586)^{1,3} 0}{(0,6381)^{1,3} + (0,5518)^{1,3} + \dots + (0,4586)^{1,3}}$$

$$v_{1,1} = \frac{105,4563}{203,2803} = 0,5188$$

Perhitungan untuk pusat *cluster* 2:

$$v_{2,1} = \frac{\sum_{i=1}^n (\mu_{i,2})^{1,3} x_{i,1}}{\sum_{i=1}^n (\mu_{i,2})^{1,3}}$$

$$v_{2,1} = \frac{(\mu_{1,2})^{1,3} x_{1,1} + (\mu_{2,2})^{1,3} x_{2,1} + \dots + (\mu_{502,2})^{1,3} x_{502,1}}{(\mu_{1,2})^{1,3} + (\mu_{2,2})^{1,3} + \dots + (\mu_{502,2})^{1,3}}$$

$$v_{2,1} = \frac{(0,3619)^{1,3} 0,625 + (0,4482)^{1,3} 0,625 + \dots + (0,5414)^{1,3} 0}{(0,3619)^{1,3} + (0,4482)^{1,3} + \dots + (0,5414)^{1,3}}$$

$$v_{2,1} = \frac{114,6926}{223,2803} = 0,5127$$

Nilai tersebut merupakan pusat cluster untuk variabel Umur (X_1) pada *cluster* 1 dan *cluster* 2 dengan nilai $m = 1,3$. Perhitungan yang sama dilakukan pada variabel lainnya sehingga setiap pusat *cluster* berbentuk vektor berdimensi tujuh. Berdasarkan hasil perhitungan, pusat *cluster* untuk $c = 2$ dapat dinyatakan sebagai:

$$v_1 = (0,5188, 0,4902, 0,5036, 0,5402, 0,4832, 0,4976, 0,5053)$$

$$v_2 = (0,5127, 0,5104, 0,5302, 0,5279, 0,4813, 0,5137, 0,4764)$$

Setiap komponen pada vektor tersebut secara berurutan menyatakan nilai untuk variabel X_1 sampai dengan X_7 . Hasil lengkap pusat dua *cluster* yang terbentuk tetap ditampilkan pada Tabel 4.11 agar perbandingan nilai antarvariabel pada masing-masing *cluster* dapat terlihat dengan lebih jelas.

Tabel 4.11 Hasil Pusat Dua *Cluster* ($m = 1,3$)

Variabel	Pusat <i>Cluster</i>	
	<i>Cluster 1</i>	<i>Cluster 2</i>
X_1	0,5188	0,5127
X_2	0,4902	0,5104
X_3	0,5036	0,5302
X_4	0,5402	0,5279
X_5	0,4832	0,4813
X_6	0,4976	0,5137
X_7	0,5053	0,4764

3. Matriks Kovarian

Perhitungan matriks kovarian pada variasi $m = 1,3$ dilakukan dengan prosedur yang sama seperti pada variasi $m = 1,1$. Berdasarkan pusat *cluster* dan *weighting exponent*, diperoleh hasil akhir matriks kovarian untuk masing-masing *cluster* sebagai berikut:

F_1

$$= \begin{bmatrix} 0,0920 & 0,0033 & -0,0049 & 0,0064 & -0,0109 & -0,0059 & 0,0085 \\ 0,0033 & 0,1240 & -0,0139 & 0,0108 & 0,0028 & -0,0059 & 0,0074 \\ -0,0049 & -0,0139 & 0,1183 & -0,0045 & -0,0006 & 0,0010 & -0,0040 \\ 0,0064 & 0,0108 & -0,0045 & 0,1021 & 0,0042 & -0,0107 & 0,0052 \\ -0,0109 & 0,0028 & -0,0006 & 0,0042 & 0,1250 & -0,0043 & -0,0027 \\ -0,0059 & -0,0059 & 0,0010 & -0,0107 & -0,0043 & 0,1466 & -0,0235 \\ 0,0085 & 0,0074 & -0,0040 & 0,0052 & -0,0027 & -0,0235 & 0,1639 \end{bmatrix}$$

F_2

$$= \begin{bmatrix} 0,0948 & 0,0025 & 0,0027 & 0,0029 & -0,0091 & -0,0046 & 0,0069 \\ 0,0025 & 0,1172 & -0,0105 & 0,0018 & -0,0019 & 0,0073 & -0,0019 \\ 0,0027 & -0,0105 & 0,1167 & 0,0006 & -0,0086 & -0,0043 & 0,0096 \\ 0,0029 & 0,0018 & 0,0006 & 0,0918 & -0,0007 & -0,0054 & 0,0049 \\ -0,0091 & -0,0019 & -0,0086 & -0,0007 & 0,1277 & -0,0046 & -0,0074 \\ -0,0046 & 0,0073 & -0,0043 & -0,0054 & -0,0046 & 0,1317 & -0,0150 \\ 0,0069 & -0,0019 & 0,0096 & 0,0049 & -0,0074 & -0,0150 & 0,1649 \end{bmatrix}$$

4. Jarak Mahalanobis

Perhitungan jarak *Mahalanobis* pada variasi $m = 1,3$ dilakukan dengan prosedur yang sama seperti pada variasi $m = 1,1$, yaitu dengan membentuk A_k dari matriks kovarian masing-masing *cluster*, kemudian menghitung jarak setiap data terhadap pusat *cluster*. Untuk $c = 2$, hasil jarak *Mahalanobis* antara setiap data dengan pusat *cluster* 1 dan *cluster* 2 diperlihatkan pada Tabel 4.12.

Tabel 4.12 Hasil Jarak *Mahalanobis* Dua Cluster ($m = 1,3$)

Data	Cluster 1	Cluster 2
1	0,5143	0,5078
2	0,8761	0,8138
3	0,8107	0,8056
$\vdots$	$\vdots$	$\vdots$
502	1,0222	0,9613

5. Fungsi Objektif

Perhitungan nilai fungsi objektif pada variasi $m = 1,3$ dilakukan dengan prosedur yang sama seperti pada variasi $m = 1,1$ dengan menggunakan derajat keanggotaan dan jarak *Mahalanobis* pada masing-masing *cluster*.

Hasil perhitungan untuk $c = 2$ sebagai berikut:

Perhitungan untuk *cluster* 1

$$\begin{aligned} \sum (\mu_{i,k})^m D_{i,kA_k}^2 &= (\mu_{1,1})^{1,3} D_{1,1A_1}^2 + \cdots + (\mu_{502,1})^{1,3} D_{502,1A_1}^2 \\ &= (0,6831)^{1,3} (0,5143)^2 + \cdots + (0,4586)^{1,3} (1,0222)^2 \end{aligned}$$

$$= 0,1475 + \dots + 0,3792 = 172,1875$$

Perhitungan untuk *cluster* 2

$$\begin{aligned}\sum (\mu_{i,k})^m D_{i,kA_k}^2 &= (\mu_{1,2})^{1,3} D_{1,2A_2}^2 + \dots + (\mu_{502,2})^{1,3} D_{502,2A_2}^2 \\ &= (0,3619)^{1,3} (0,5078)^2 + \dots + (0,5414)^{1,3} (0,9613)^2 \\ &= 0,0688 + \dots + 0,4162 = 184,3163\end{aligned}$$

Maka untuk fungsi objektif pada iterasi ke-1 ini adalah:

$$\begin{aligned}J_{GK}^{(1)} &= \sum_{i=1}^{502} \sum_{k=1}^2 (\mu_{i,k})^{1,3} D_{i,kA_k}^2 \\ &= 172,1875 + 184,3163 = 356,5038\end{aligned}$$

6. Matriks Keanggotaan Baru

Pada variasi $m = 1,5$ pembaruan matriks keanggotaan dilakukan dengan prosedur yang sama seperti $m = 1,3$, yaitu berdasarkan jarak *Mahalanobis* setiap data terhadap pusat *cluster* dan nilai m . Sebagai contoh perhitungan, derajat keanggotaan data pertama terhadap *cluster* 1 dan *cluster* 2 dihitung sebagai berikut:

Dengan nilai $m = 1,3$, maka nilai pangkat:

$$\frac{2}{m-1} = \frac{2}{1,3-1} = 6,6667$$

Derajat keanggotaan data pertama terhadap *cluster* 1

$$\begin{aligned}\mu_{1,1}^{(t+1)} &= \left[\left(\frac{D_{1,1A_1}}{D_{1,1A_1}} \right)^{6,6667} + \left(\frac{D_{1,1A_1}}{D_{1,2A_2}} \right)^{6,6667} \right]^{-1} \\ &= \left[1 + \left(\frac{0,5143}{0,5078} \right)^{6,6667} \right]^{-1} = [1 + 1,0879]^{-1} \\ &= 0,4790\end{aligned}$$

Derajat keanggotaan data pertama terhadap *cluster* 2

$$\begin{aligned}\mu_{1,2}^{(t+1)} &= \left[\left(\frac{D_{1,2A_2}}{D_{1,1A_1}} \right)^{6,6667} + \left(\frac{D_{1,2A_2}}{D_{1,2A_2}} \right)^{6,6667} \right]^{-1} \\ &= \left[\left(\frac{0,5078}{0,5143} \right)^{6,6667} + 1 \right]^{-1} = [0,9192 + 1]^{-1} \\ &= 0,5210\end{aligned}$$

Beberapa hasil pembaruan derajat keanggotaan untuk dua *cluster* ditampilkan pada Tabel 4.13.

Tabel 4.13 Hasil Matriks Keanggotaan Baru dengan Dua *Cluster* ($m = 1,3$)

Data	<i>Cluster</i> 1	<i>Cluster</i> 2
1	0,4790	0,5210
2	0,3794	0,6206
3	0,4895	0,5105
⋮	⋮	⋮
502	0,3990	0,6010

7. Kriteria Konvergensi

Pada variasi $m = 1,3$ pemeriksaan konvergensi dilakukan dengan prosedur yang sama seperti $m = 1,1$, yaitu membandingkan perubahan matriks keanggotaan pada dua iterasi berturut-turut, $\|U_{t+1} - U_t\|$, terhadap nilai *threshold* $\varepsilon = 0,0001$. Sebagai hasil perhitungan pada iterasi awal, diperoleh:

Perhitungan untuk *cluster* 1:

$$\begin{aligned}\|U_{t+1} - U_t\| &= \sqrt{(0,4790 - 0,6381)^2 + \dots + (0,3990 - 0,4586)^2} \\ &= \sqrt{26,9166}\end{aligned}$$

Perhitungan untuk *cluster* 2:

$$\|U_{t+1} - U_t\| = \sqrt{(0,5210 - 0,3619)^2 + \dots + (0,6010 - 0,5414)^2}$$

$$= \sqrt{26,9166}$$

Karena perhitungan dilakukan dengan menjumlahkan hasil dari seluruh *cluster*, maka nilai akhirnya diperoleh sebagai berikut:

$$\|U_{t+1} - U_t\| = \sqrt{26,9166 + 26,9166} = \sqrt{53,8333}$$

$$\|U_{t+1} - U_t\| = 7,3371$$

Nilai tersebut masih jauh lebih besar dari $\varepsilon = 0,0001$, maka pada iterasi awal algoritma belum dinyatakan konvergen. Proses iterasi kemudian dilanjutkan hingga perubahan matriks keanggotaan semakin kecil. Setelah iterasi berjalan hingga mencapai iterasi ke-132, diperoleh nilai sebagai berikut:

$$\|U_{t+1} - U_t\| = 0,0000984$$

Dengan demikian, pada iterasi tersebut algoritma FGK dengan $m = 1,3$ telah mencapai konvergensi.

8. Hasil *Clustering*

Setelah algoritma *Fuzzy Gustafson-Kessel* dengan $m = 1,3$ mencapai konvergensi, penentuan hasil *clustering* data didasarkan pada matriks keanggotaan akhir. Setiap data dikelompokkan ke dalam *cluster* yang memiliki nilai derajat keanggotaan terbesar, sehingga *cluster* tersebut menjadi *cluster* dominan yang paling merepresentasikan data. Hasil *clustering* data berdasarkan *cluster* dominan untuk $c = 2$ dan $m = 1,3$ disajikan pada Tabel 4.14.

Tabel 4.14 Hasil *Clustering* Dua *Cluster* ($m = 1,3$)

Data	Derajat Keanggotaan		Cluster
	1	2	
1	0,0027	0,9973	2
2	0,0066	0,9934	2
3	0,4305	0,5695	2
⋮	⋮	⋮	⋮
502	0,7696	0,2304	1

4.3.3 Perhitungan untuk Variasi Nilai $m = 1,5$

Pada variasi ini, proses *clustering* dilakukan dengan menggunakan nilai $m = 1,5$. Seluruh tahapan mengikuti langkah-langkah yang telah dijelaskan pada bagian sebelumnya $m = 1,3$ dan $m = 1,1$. Sebagaimana telah dijelaskan pada subbab sebelumnya, proses perhitungan manual tetap difokuskan pada kasus $c = 2$ sebagai contoh perhitungan. Walaupun demikian, seluruh tahapan perhitungan yang dilakukan pada masing-masing nilai c menetapkan prosedur yang sama.

1. Inisialisasi Matriks Keanggotaan Awal

Pada variasi nilai $m = 1,5$ proses diawali dengan inisialisasi U_0 . Matriks U_0 yang digunakan sama seperti pada perhitungan $m = 1,3$ karena pembangkitan U_0 tidak dipengaruhi oleh nilai m . Bentuk U_0 untuk masing-masing banyaknya *cluster* $c = 2$ sampai $c = 5$ tetap menggunakan hasil inisialisasi awal yang telah ditampilkan pada subbab sebelumnya.

2. Pusat *Cluster*

Perhitungan pusat *cluster* awal dilakukan dengan menggunakan matriks U_0 . Sebagai contoh manual, perhitungan ditunjukkan untuk $c = 2$ menggunakan variabel pertama, yaitu Umur (X_1).

Perhitungan pusat *cluster* 1 dapat dituliskan sebagai berikut:

$$v_{1,1} = \frac{\sum_{i=1}^n (\mu_{i,1})^{1,5} x_{i,1}}{\sum_{i=1}^n (\mu_{i,1})^{1,5}}$$

$$v_{1,1} = \frac{(\mu_{1,1})^{1,5} x_{1,1} + (\mu_{2,1})^{1,5} x_{2,1} + \dots + (\mu_{502,1})^{1,5} x_{502,1}}{(\mu_{1,1})^{1,5} + (\mu_{2,1})^{1,5} + \dots + (\mu_{502,1})^{1,5}}$$

$$v_{1,1} = \frac{(0,6381)^{1,5} 0,625 + (0,5518)^{1,5} 0,625 + \dots + (0,4586)^{1,5} 0}{(0,6381)^{1,5} + (0,5518)^{1,5} + \dots + (0,4586)^{1,5}}$$

$$v_{1,1} = \frac{95,0421}{182,9538} = 0,5195$$

Perhitungan pusat *cluster* 2:

$$v_{2,1} = \frac{\sum_{i=1}^n (\mu_{i,2})^{1,5} x_{i,1}}{\sum_{i=1}^n (\mu_{i,2})^{1,5}}$$

$$v_{2,1} = \frac{(\mu_{1,2})^{1,5} x_{1,1} + (\mu_{2,2})^{1,5} x_{2,1} + \dots + (\mu_{502,2})^{1,5} x_{502,1}}{(\mu_{1,2})^{1,5} + (\mu_{2,2})^{1,5} + \dots + (\mu_{502,2})^{1,5}}$$

$$v_{2,1} = \frac{(0,3619)^{1,5} 0,625 + (0,4482)^{1,5} 0,625 + \dots + (0,5414)^{1,5} 0}{(0,3619)^{1,5} + (0,4482)^{1,5} + \dots + (0,5414)^{1,5}}$$

$$v_{2,1} = \frac{104,2803}{203,3678} = 0,5127$$

Nilai tersebut merupakan pusat cluster untuk variabel Umur (X_1) pada *cluster* 1 dan *cluster* 2 dengan nilai $m = 1,5$. Perhitungan yang sama dilakukan pada variabel lainnya sehingga setiap pusat *cluster* berbentuk vektor berdimensi tujuh. Berdasarkan hasil perhitungan, pusat *cluster* untuk $c = 2$ dapat dinyatakan sebagai:

$$v_1 = (0,5195, 0,4888, 0,5127, 0,5010, 0,4833, 0,4961, 0,5056)$$

$$v_2 = (0,5127, 0,5112, 0,5304, 0,5274, 0,4812, 0,5141, 0,4737)$$

Setiap komponen vektor menyatakan nilai untuk variabel X_1 sampai dengan X_7 , dan nilai tersebut juga disajikan dalam Tabel 4.15.

Tabel 4.15 Hasil Pusat Dua *Cluster* ($m = 1,5$)

Variabel	Pusat <i>Cluster</i>	
	<i>Cluster 1</i>	<i>Cluster 2</i>
X_1	0,5195	0,5127
X_2	0,4888	0,5112
X_3	0,5010	0,5304
X_4	0,5410	0,5274
X_5	0,4833	0,4812
X_6	0,4961	0,5141
X_7	0,5056	0,4737

3. Matriks Kovarian

Perhitungan matriks kovarian pada variasi $m = 1,5$ dilakukan dengan prosedur yang sama seperti pada variasi $m = 1,3$. Berdasarkan pusat *cluster* dan *weighting exponent*, diperoleh hasil akhir matriks kovarian untuk masing-masing *cluster* sebagai berikut:

F_1

$$= \begin{bmatrix} 0,0918 & 0,0034 & -0,0057 & 0,0065 & -0,0107 & -0,0060 & 0,0085 \\ 0,0034 & 0,1244 & -0,0140 & 0,0116 & 0,0030 & -0,0069 & 0,0082 \\ -0,0057 & -0,0140 & 0,1182 & -0,0051 & -0,0003 & 0,0021 & -0,0049 \\ 0,0065 & 0,0116 & -0,0051 & 0,1024 & 0,0043 & -0,0112 & 0,0053 \\ -0,0107 & 0,0030 & -0,0003 & 0,0043 & 0,1247 & -0,0044 & -0,0026 \\ -0,0060 & -0,0069 & 0,0021 & -0,0112 & -0,0044 & 0,1475 & -0,0240 \\ 0,0085 & 0,0082 & -0,0049 & 0,0053 & -0,0026 & -0,0240 & 0,1642 \end{bmatrix}$$

F_2

$$= \begin{bmatrix} 0,0950 & 0,0024 & 0,0028 & 0,0026 & -0,0087 & -0,0046 & 0,0067 \\ 0,0024 & 0,1169 & -0,0102 & 0,0016 & -0,0022 & 0,0078 & -0,0021 \\ 0,0028 & -0,0102 & 0,1165 & 0,0006 & -0,0091 & -0,0039 & 0,0102 \\ 0,0026 & 0,0016 & 0,0006 & 0,0911 & -0,0011 & -0,0053 & 0,0050 \\ -0,0087 & -0,0022 & -0,0091 & -0,0011 & 0,1276 & -0,0047 & -0,0078 \\ -0,0046 & 0,0078 & -0,0039 & -0,0053 & -0,0047 & 0,1310 & -0,0146 \\ 0,0067 & -0,0021 & 0,0102 & 0,0050 & -0,0078 & -0,0146 & 0,1652 \end{bmatrix}$$

4. Jarak Mahalanobis

Perhitungan jarak *Mahalanobis* pada variasi $m = 1,5$ dilakukan dengan prosedur yang sama seperti pada variasi $m = 1,3$, yaitu dengan membentuk A_k dari matriks kovarian masing-masing *cluster*, kemudian

menghitung jarak setiap data terhadap pusat *cluster*. Untuk $c = 2$, hasil jarak *Mahalanobis* antara setiap data dengan pusat *cluster* 1 dan *cluster* 2 diperlihatkan pada Tabel 4.16.

Tabel 4.16 Hasil Jarak *Mahalanobis* Dua *Cluster* ($m = 1,5$)

Data	Cluster 1	Cluster 2
1	0,5161	0,5086
2	0,8811	0,8118
3	0,8096	0,8043
$\vdots$	$\vdots$	$\vdots$
502	1,0255	0,9580

5. Fungsi Objektif

Perhitungan nilai fungsi objektif pada variasi $m = 1,5$ dilakukan dengan prosedur yang sama seperti pada variasi $m = 1,1$ dan $m = 1,3$ dengan menggunakan derajat keanggotaan dan jarak *Mahalanobis* pada masing-masing *cluster*. Hasil perhitungan untuk $c = 2$ sebagai berikut:

Perhitungan *cluster* 1

$$\begin{aligned}\sum (\mu_{i,k})^m D_{i,kA_k}^2 &= (\mu_{1,1})^{1,5} D_{1,1A_1}^2 + \dots + (\mu_{502,1})^{1,5} D_{502,1A_1}^2 \\ &= 155,0529\end{aligned}$$

Perhitungan *cluster* 2

$$\begin{aligned}\sum (\mu_{i,k})^m D_{i,kA_k}^2 &= (\mu_{1,2})^{1,5} D_{1,2A_2}^2 + \dots + (\mu_{502,2})^{1,5} D_{502,2A_2}^2 \\ &= 167,2045\end{aligned}$$

Maka untuk fungsi objektif pada iterasi ke-1 ini adalah:

$$\begin{aligned}J_{GK}^{(1)} &= \sum_{i=1}^{502} \sum_{k=1}^2 (\mu_{i,k})^{1,5} D_{i,kA_k}^2 \\ &= 155,0529 + 167,2045 = 322,2573\end{aligned}$$

6. Matriks Keanggotaan Baru

Pada variasi $m = 1,5$ pembaruan matriks keanggotaan dilakukan dengan prosedur yang sama seperti $m = 1,3$, yaitu berdasarkan jarak *Mahalanobis* setiap data terhadap pusat *cluster* dan nilai m . Sebagai contoh perhitungan, derajat keanggotaan data pertama terhadap *cluster* 1 dan *cluster* 2 dihitung sebagai berikut:

Dengan nilai $m = 1,5$, maka nilai pangkat:

$$\frac{2}{m-1} = \frac{2}{1,5-1} = 4$$

Derajat keanggotaan data pertama terhadap *cluster* 1

$$\begin{aligned}\mu_{1,1}^{(t+1)} &= \left[\left(\frac{D_{1,1A_1}}{D_{1,1A_1}} \right)^4 + \left(\frac{D_{1,1A_1}}{D_{1,2A_2}} \right)^4 \right]^{-1} \\ &= \left[1 + \left(\frac{0,5161}{0,5086} \right)^4 \right]^{-1} = 0,4853\end{aligned}$$

Derajat keanggotaan data pertama terhadap *cluster* 2

$$\begin{aligned}\mu_{1,2}^{(t+1)} &= \left[\left(\frac{D_{1,2A_2}}{D_{1,1A_1}} \right)^4 + \left(\frac{D_{1,2A_2}}{D_{1,2A_2}} \right)^4 \right]^{-1} \\ &= \left[\left(\frac{0,5086}{0,5161} \right)^4 + 1 \right]^{-1} = 0,5147\end{aligned}$$

Beberapa hasil pembaruan derajat keanggotaan untuk dua *cluster* ditampilkan pada Tabel 4.17.

Tabel 4.17 Hasil Matriks Keanggotaan Baru dengan Dua *Cluster* ($m = 1,5$)

Data	Cluster 1	Cluster 2
1	0,4853	0,5147
2	0,4188	0,5812
3	0,4935	0,5065
⋮	⋮	⋮
502	0,4323	0,5677

7. Kriteria Konvergensi

Pada variasi $m = 1,5$ pemeriksaan konvergensi dilakukan dengan prosedur yang sama seperti $m = 1,3$, yaitu membandingkan perubahan matriks keanggotaan pada dua iterasi berturut-turut, $\|U_{t+1} - U_t\|$, terhadap nilai *threshold* $\varepsilon = 0,0001$. Sebagai hasil perhitungan pada iterasi awal, diperoleh:

Perhitungan untuk *cluster* 1:

$$\begin{aligned}\|U_{t+1} - U_t\| &= \sqrt{(0,4853 - 0,6381)^2 + \dots + (0,4323 - 0,4586)^2} \\ &= \sqrt{26,9815}\end{aligned}$$

Perhitungan untuk *cluster* 2:

$$\begin{aligned}\|U_{t+1} - U_t\| &= \sqrt{(0,5147 - 0,3619)^2 + \dots + (0,5677 - 0,5414)^2} \\ &= \sqrt{26,9815}\end{aligned}$$

Karena perhitungan dilakukan dengan menjumlahkan hasil dari seluruh *cluster*, maka nilai akhirnya diperoleh sebagai berikut:

$$\begin{aligned}\|U_{t+1} - U_t\| &= \sqrt{26,9815 + 26,9815} = \sqrt{53,9630} \\ \|U_{t+1} - U_t\| &= 7,3459\end{aligned}$$

Nilai tersebut masih jauh lebih besar dari $\varepsilon = 0,0001$, maka pada iterasi awal algoritma belum dinyatakan konvergen. Proses iterasi kemudian dilanjutkan hingga perubahan matriks keanggotaan semakin kecil. Setelah iterasi ke-238, diperoleh nilai:

$$\|U_{t+1} - U_t\| = 0,0000998$$

Dengan demikian, pada iterasi tersebut algoritma FGK dengan $m = 1,5$ telah mencapai konvergensi.

8. Hasil *Clustering*

Setelah algoritma *Fuzzy Gustafson-Kessel* dengan $m = 1,5$ mencapai konvergensi, penentuan hasil *clustering* data didasarkan pada matriks keanggotaan akhir. Setiap data dikelompokkan ke dalam *cluster* yang memiliki nilai derajat keanggotaan terbesar, sehingga *cluster* tersebut menjadi *cluster* dominan yang paling merepresentasikan data. Hasil *clustering* data berdasarkan *cluster* dominan untuk $c = 2$ dan $m = 1,5$ disajikan pada Tabel 4.18.

Tabel 4.18 Hasil *Clustering* Dua *Cluster* ($m = 1,5$)

Data	Derajat Keanggotaan		Cluster
	1	2	
1	0,2658	0,7342	2
2	0,3290	0,6710	2
3	0,2255	0,7745	2
⋮	⋮	⋮	⋮
502	0,6597	0,3403	1

4.3.4 Perhitungan untuk Variasi Nilai $m = 1,7$

Pada variasi ini, proses *clustering* dilakukan dengan menggunakan $m = 1,7$. Seluruh tahapan mengikuti prosedur yang sama seperti pada subbab sebelumnya $m = 1,1$, $m = 1,3$, dan $m = 1,5$. Perhitungan manual tetap difokuskan pada kasus $c = 2$ dan menggunakan variabel pertama, yaitu Umur (X_1).

1. Inisialisasi Matriks Keanggotaan Awal

Pada variasi nilai $m = 1,7$ proses diawali dengan inisialisasi U_0 . Matriks U_0 yang digunakan sama seperti pada perhitungan $m = 1,1$, $m = 1,3$, dan $m = 1,5$ karena pembangkitan U_0 tidak dipengaruhi oleh nilai m . Bentuk U_0 untuk masing-masing $c = 2$ sampai $c = 5$ tetap menggunakan hasil

inisialisasi awal yang telah ditampilkan pada subbab sebelumnya.

2. Pusat *Cluster*

Perhitungan pusat *cluster* awal dilakukan dengan menggunakan matriks U_0 .

Sebagai contoh manual, perhitungan ditunjukkan untuk $c = 2$ menggunakan variabel pertama, yaitu Umur (X_1).

Perhitungan pusat untuk *cluster* 1 dapat dituliskan sebagai berikut:

$$v_{1,1} = \frac{\sum_{i=1}^n (\mu_{i,1})^{1,7} x_{i,1}}{\sum_{i=1}^n (\mu_{i,1})^{1,7}}$$

$$v_{1,1} = \frac{(\mu_{1,1})^{1,7} x_{1,1} + (\mu_{2,1})^{1,7} x_{2,1} + \dots + (\mu_{502,1})^{1,7} x_{502,1}}{(\mu_{1,1})^{1,7} + (\mu_{2,1})^{1,7} + \dots + (\mu_{502,1})^{1,7}}$$

$$v_{1,1} = \frac{(0,6381)^{1,7} 0,625 + (0,5518)^{1,7} 0,625 + \dots + (0,4586)^{1,7} 0}{(0,6381)^{1,7} + (0,5518)^{1,7} + \dots + (0,4586)^{1,7}}$$

$$v_{1,1} = \frac{86,1381}{165,5855} = 0,5202$$

Perhitungan pusat untuk *cluster* 2:

$$v_{2,1} = \frac{\sum_{i=1}^n (\mu_{i,2})^{1,7} x_{i,1}}{\sum_{i=1}^n (\mu_{i,2})^{1,7}}$$

$$v_{2,1} = \frac{(\mu_{1,2})^{1,7} x_{1,1} + (\mu_{2,2})^{1,7} x_{2,1} + \dots + (\mu_{502,2})^{1,7} x_{502,1}}{(\mu_{1,2})^{1,7} + (\mu_{2,2})^{1,7} + \dots + (\mu_{502,2})^{1,7}}$$

$$v_{2,1} = \frac{(0,3619)^{1,7} 0,625 + (0,4482)^{1,7} 0,625 + \dots + (0,5414)^{1,7} 0}{(0,3619)^{1,7} + (0,4482)^{1,7} + \dots + (0,5414)^{1,7}}$$

$$v_{2,1} = \frac{95,3182}{185,8781} = 0,5128$$

Nilai tersebut merupakan pusat *cluster* untuk variabel Umur (X_1) pada *cluster* 1 dan *cluster* 2 dengan nilai $m = 2$. Perhitungan yang sama dilakukan pada variabel lainnya sehingga setiap pusat *cluster* berbentuk

vektor berdimensi tujuh. Berdasarkan hasil perhitungan, pusat *cluster* untuk $c = 2$ dapat dinyatakan sebagai:

$$v_1 = (0,5202, 0,4874, 0,4983, 0,5417, 0,4834, 0,4947, 0,5056)$$

$$v_2 = (0,5128, 0,5118, 0,5306, 0,5269, 0,4812, 0,5144, 0,4710)$$

Setiap komponen vektor menyatakan nilai untuk variabel X_1 sampai dengan X_7 , dan nilai tersebut juga disajikan dalam Tabel 4.19.

Tabel 4.19 Hasil Pusat Dua *Cluster* ($m = 1,7$)

Variabel	Pusat <i>Cluster</i>	
	<i>Cluster 1</i>	<i>Cluster 2</i>
X_1	0,5202	0,5128
X_2	0,4874	0,5118
X_3	0,4983	0,5306
X_4	0,5417	0,5269
X_5	0,4834	0,4811
X_6	0,4947	0,5144
X_7	0,5056	0,4710

3. Matriks Kovarian

Perhitungan matriks kovarian pada variasi $m = 1,7$ dilakukan dengan prosedur yang sama seperti pada variasi $m = 1,1$, $m = 1,3$, dan $m = 1,5$. Berdasarkan pusat *cluster* dan *weighting exponent*, diperoleh hasil akhir matriks kovarian untuk masing-masing *cluster* sebagai berikut:

F_1

$$= \begin{bmatrix} 0,0917 & 0,0034 & -0,0064 & 0,0065 & -0,0104 & -0,0061 & 0,0086 \\ 0,0034 & 0,1248 & -0,0141 & 0,0124 & 0,0031 & -0,0078 & 0,0091 \\ -0,0064 & -0,0141 & 0,1181 & -0,0056 & -0,00002 & 0,0031 & -0,0058 \\ 0,0065 & 0,0124 & -0,0056 & 0,1026 & 0,0045 & -0,0117 & 0,0055 \\ -0,0104 & 0,0031 & -0,00002 & 0,0045 & 0,1243 & -0,0045 & -0,0026 \\ -0,0061 & -0,0078 & 0,0031 & -0,0117 & -0,0045 & 0,1483 & -0,0245 \\ 0,0086 & 0,0091 & -0,0058 & 0,0055 & -0,0026 & -0,0245 & 0,1645 \end{bmatrix}$$

F_2

$$= \begin{bmatrix} 0,0951 & 0,0023 & 0,0028 & 0,0023 & -0,0083 & -0,0046 & 0,0066 \\ 0,0023 & 0,1166 & -0,0100 & 0,0014 & -0,0025 & 0,0082 & -0,0022 \\ 0,0028 & -0,0100 & 0,1164 & 0,0005 & -0,0096 & -0,0035 & 0,0107 \\ 0,0023 & 0,0014 & 0,0005 & 0,0903 & -0,0014 & -0,0052 & 0,0051 \\ -0,0083 & -0,0025 & -0,0096 & -0,0014 & 0,1276 & -0,0048 & -0,0082 \\ -0,0046 & 0,0082 & -0,0035 & -0,0052 & -0,0048 & 0,1304 & -0,0142 \\ 0,0066 & -0,0022 & 0,0107 & 0,0051 & -0,0082 & -0,0142 & 0,1655 \end{bmatrix}$$

4. Jarak Mahalanobis

Perhitungan jarak *Mahalanobis* pada variasi $m = 1,7$ dilakukan dengan prosedur yang sama seperti pada variasi $m = 1,1$, $m = 1,3$, dan $m = 1,5$ yaitu dengan membentuk A_k dari matriks kovarian masing-masing *cluster*, kemudian menghitung jarak setiap data terhadap pusat *cluster*. Untuk $c = 2$, hasil jarak *Mahalanobis* antara setiap data dengan pusat *cluster* 1 dan *cluster* 2 diperlihatkan pada Tabel 4.20.

Tabel 4.20 Hasil Jarak *Mahalanobis* Dua *Cluster* ($m = 1,7$)

Data	Cluster 1	Cluster 2
1	0,5179	0,5094
2	0,8859	0,8102
3	0,8083	0,8031
⋮	⋮	⋮
502	1,0285	0,9550

5. Fungsi Objektif

Perhitungan nilai fungsi objektif pada variasi $m = 2$ dilakukan dengan prosedur yang sama seperti pada variasi $m = 1,1$, $m = 1,3$, dan $m = 1,5$ dengan menggunakan derajat keanggotaan dan jarak *Mahalanobis* pada masing-masing *cluster*. Hasil perhitungan untuk $c = 2$ sebagai berikut:

Perhitungan *cluster* 1

$$\begin{aligned} \sum (\mu_{i,k})^m D_{i,kA_k}^{1,7} &= (\mu_{1,1})^{1,7} D_{1,1A_1}^{1,7} + \dots + (\mu_{502,1})^{1,7} D_{502,1A_1}^{1,7} \\ &= (0,6381)^{1,7} (0,5179)^{1,7} + \dots + (0,4586)^{1,7} (1,0285)^{1,7} \end{aligned}$$

$$= 140,3808$$

Perhitungan *cluster* 2

$$\begin{aligned}\sum (\mu_{i,k})^m D_{i,kA_k}^{1,7} &= (\mu_{1,2})^{1,7} D_{1,2A_2}^{1,7} + \dots + (\mu_{502,2})^{1,7} D_{502,2A_2}^{1,7} \\ &= (0,3619)^{1,7} (0,5094)^{1,7} + \dots + (0,5414)^{1,7} (0,9550)^{1,7} \\ &= 152,5007\end{aligned}$$

Maka untuk fungsi objektif pada iterasi ke-1 ini adalah:

$$\begin{aligned}J_{GK}^{(1)} &= \sum_{i=1}^{502} \sum_{k=1}^2 (\mu_{i,k})^2 D_{i,kA_k}^2 \\ &= 140,3808 + 152,5007 = 292,8815\end{aligned}$$

6. Matriks Keanggotaan Baru

Pada variasi $m = 2$ pembaruan matriks keanggotaan dilakukan dengan prosedur yang sama seperti $m = 1,1$, $m = 1,3$, dan $m = 1,5$ yaitu berdasarkan jarak *Mahalanobis* setiap data terhadap pusat *cluster* dan nilai m . Sebagai contoh perhitungan, derajat keanggotaan data pertama terhadap *cluster* 1 dan *cluster* 2 dihitung sebagai berikut:

Dengan nilai $m = 2$, maka nilai pangkat:

$$\frac{2}{1,7-1} = \frac{2}{2-1} = 2,8571$$

Derajat keanggotaan data pertama terhadap *cluster* 1

$$\begin{aligned}\mu_{1,1}^{(t+1)} &= \left[\left(\frac{D_{1,1A_1}}{D_{1,1A_1}} \right)^{2,8571} + \left(\frac{D_{1,1A_1}}{D_{1,2A_2}} \right)^{2,8571} \right]^{-1} \\ &= \left[1 + \left(\frac{0,5179}{0,5094} \right)^{2,8571} \right]^{-1} = 0,4882\end{aligned}$$

Derajat keanggotaan data pertama terhadap *cluster* 2

$$\begin{aligned}\mu_{1,2}^{(t+1)} &= \left[\left(\frac{D_{1,2A_2}}{D_{1,1A_1}} \right)^2 + \left(\frac{D_{1,2A_2}}{D_{1,2A_2}} \right)^2 \right]^{-1} \\ &= \left[\left(\frac{0,5094}{0,5179} \right)^{2,8571} + 1 \right]^{-1} = 0,5118\end{aligned}$$

Beberapa hasil pembaruan derajat keanggotaan untuk dua *cluster* ditampilkan pada Tabel 4.21.

Tabel 4.21 Hasil Matriks Keanggotaan Baru dengan Dua *Cluster* ($m = 1,7$)

Data	Cluster 1	Cluster 2
1	0,4882	0,5118
2	0,4366	0,5634
3	0,4954	0,5046
$\vdots$	$\vdots$	$\vdots$
502	0,4472	0,5528

7. Kriteria Konvergensi

Pada variasi $m = 1,7$ pemeriksaan konvergensi dilakukan dengan prosedur yang sama seperti $m = 1,1$, $m = 1,3$, dan $m = 1,5$ yaitu membandingkan perubahan matriks keanggotaan pada dua iterasi berturut-turut, $\|U_{t+1} - U_t\|$, terhadap nilai *threshold* $\varepsilon = 0,0001$. Sebagai hasil perhitungan pada iterasi awal, diperoleh:

Perhitungan untuk *cluster* 1:

$$\begin{aligned}\|U_{t+1} - U_t\| &= \sqrt{(0,4882 - 0,6381)^2 + \dots + (0,4472 - 0,4586)^2} \\ &= \sqrt{27,1557}\end{aligned}$$

Perhitungan untuk *cluster* 2:

$$\begin{aligned}\|U_{t+1} - U_t\| &= \sqrt{(0,5118 - 0,3619)^2 + \dots + (0,5528 - 0,5414)^2} \\ &= \sqrt{27,1557}\end{aligned}$$

Karena perhitungan dilakukan dengan menjumlahkan hasil dari seluruh

cluster, maka nilai akhirnya diperoleh sebagai berikut:

$$\|U_{t+1} - U_t\| = \sqrt{27,1557 + 27,1557} = \sqrt{54,2913}$$

$$\|U_{t+1} - U_t\| = 7,3683$$

Nilai tersebut masih jauh lebih besar dari $\varepsilon = 0,0001$, maka pada iterasi awal algoritma belum dinyatakan konvergen. Proses iterasi kemudian dilanjutkan hingga perubahan matriks keanggotaan semakin kecil. Setelah iterasi ke-41, diperoleh nilai:

$$\|U_{t+1} - U_t\| = 0,0000963$$

Dengan demikian, pada iterasi tersebut algoritma FGK dengan $m = 1,7$ telah mencapai konvergensi.

8. Hasil *Clustering*

Setelah algoritma *Fuzzy Gustafson-Kessel* dengan $m = 1,7$ mencapai konvergensi, penentuan hasil *clustering* data didasarkan pada matriks keanggotaan akhir. Setiap data dikelompokkan ke dalam *cluster* yang memiliki nilai derajat keanggotaan terbesar, sehingga *cluster* tersebut menjadi *cluster* dominan yang paling merepresentasikan data. Hasil *clustering* data berdasarkan *cluster* dominan untuk $c = 2$ dan $m = 2$ disajikan pada Tabel 4.22.

Tabel 4.22 Hasil *Clustering* Dua *Cluster* ($m = 1,7$)

Data	Derajat Keanggotaan		Cluster
	1	2	
1	0,499991	0,500009	2
2	0,499965	0,500035	2
3	0,499968	0,500032	2
⋮	⋮	⋮	⋮
502	0,500019	0,499981	1

4.3.5 Perhitungan untuk Variasi Nilai $m = 1,9$

Pada variasi ini, proses *clustering* dilakukan dengan menggunakan nilai $m = 1,9$. Seluruh tahapan mengikuti prosedur yang sama seperti pada subbab sebelumnya $m = 1,1$, $m = 1,3$, $m = 1,5$, dan $m = 1,7$. Perhitungan manual tetap difokuskan pada kasus $c = 2$ dan menggunakan variabel pertama, yaitu Umur (X_1).

1. Inisialisasi Matriks Keanggotaan Awal

Pada variasi nilai $m = 1,7$ proses diawali dengan inisialisasi U_0 . Matriks U_0 yang digunakan sama seperti pada perhitungan $m = 1,1$, $m = 1,3$, $m = 1,5$, dan $m = 1,7$ karena pembangkitan U_0 tidak dipengaruhi oleh nilai m . Bentuk U_0 untuk masing-masing nilai $c = 2$ sampai $c = 5$ tetap menggunakan hasil inisialisasi awal yang telah ditampilkan pada subbab sebelumnya.

2. Pusat Cluster

Perhitungan pusat *cluster* awal dilakukan dengan menggunakan matriks U_0 . Sebagai contoh manual, perhitungan ditunjukkan untuk $c = 2$ menggunakan variabel pertama, yaitu Umur (X_1).

Perhitungan pusat untuk *cluster* 1 dapat dituliskan sebagai berikut:

$$v_{1,1} = \frac{\sum_{i=1}^n (\mu_{i,1})^{1,9} x_{i,1}}{\sum_{i=1}^n (\mu_{i,1})^{1,9}}$$

$$v_{1,1} = \frac{(\mu_{1,1})^{1,9} x_{1,1} + (\mu_{2,1})^{1,9} x_{2,1} + \dots + (\mu_{502,1})^{1,9} x_{502,1}}{(\mu_{1,1})^{1,9} + (\mu_{2,1})^{1,9} + \dots + (\mu_{502,1})^{1,9}}$$

$$v_{1,1} = \frac{(0,6381)^{1,9} 0,625 + (0,5518)^{1,9} 0,625 + \dots + (0,4586)^{1,9} 0}{(0,6381)^{1,9} + (0,5518)^{1,9} + \dots + (0,4586)^{1,9}}$$

$$v_{1,1} = \frac{78,4616}{150,6163} = 0,5209$$

Perhitungan pusat untuk *cluster* 2:

$$v_{2,1} = \frac{\sum_{i=1}^n (\mu_{i,2})^{1,9} x_{i,1}}{\sum_{i=1}^n (\mu_{i,2})^{1,9}}$$

$$v_{2,1} = \frac{(\mu_{1,2})^{1,9} x_{1,1} + (\mu_{2,2})^{1,9} x_{2,1} + \dots + (\mu_{502,2})^{1,9} x_{502,1}}{(\mu_{1,2})^{1,9} + (\mu_{2,2})^{1,9} + \dots + (\mu_{502,2})^{1,9}}$$

$$v_{2,1} = \frac{(0,3619)^{1,9} 0,625 + (0,4482)^{1,9} 0,625 + \dots + (0,5414)^{1,9} 0}{(0,3619)^{1,9} + (0,4482)^{1,9} + \dots + (0,5414)^{1,9}}$$

$$v_{2,1} = \frac{87,5473}{170,7058} = 0,5129$$

Nilai tersebut merupakan pusat *cluster* untuk variabel Umur (X_1) pada *cluster* 1 dan *cluster* 2 dengan nilai $m = 1,9$. Perhitungan yang sama dilakukan pada variabel lainnya sehingga setiap pusat *cluster* berbentuk vektor berdimensi tujuh. Berdasarkan hasil perhitungan, pusat *cluster* untuk $c = 2$ dapat dinyatakan sebagai:

$$v_1 = (0,5209, 0,4859, 0,4956, 0,5423, 0,4834, 0,4933, 0,5055)$$

$$v_2 = (0,5129, 0,5125, 0,5306, 0,5265, 0,4811, 0,5146, 0,4683)$$

Setiap komponen vektor menyatakan nilai untuk variabel X_1 sampai dengan X_7 , dan nilai tersebut juga disajikan dalam Tabel 4.23.

Tabel 4.23 Hasil Pusat Dua *Cluster* ($m = 1,9$)

Variabel	Pusat <i>Cluster</i>	
	<i>Cluster</i> 1	<i>Cluster</i> 2
X_1	0,5209	0,5129
X_2	0,4859	0,5125
X_3	0,4956	0,5306
X_4	0,5423	0,5265
X_5	0,4834	0,4811
X_6	0,4933	0,5146
X_7	0,5055	0,4683

3. Matriks Kovarian

Perhitungan matriks kovarian pada variasi $m = 2,9$ dilakukan dengan prosedur yang sama seperti pada variasi $m = 1,1$, $m = 1,3$, $m = 1,5$, dan $m = 1,7$. Berdasarkan pusat *cluster* dan *weighting exponent*, diperoleh hasil akhir matriks kovarian untuk masing-masing *cluster* sebagai berikut:

F_1

$$= \begin{bmatrix} 0,0916 & 0,0034 & -0,0072 & 0,0066 & -0,0101 & -0,0061 & 0,0087 \\ 0,0034 & 0,1251 & -0,0142 & 0,0131 & 0,0032 & -0,0087 & 0,0099 \\ -0,0072 & -0,0142 & 0,01180 & -0,0062 & -0,0002 & 0,0042 & -0,0067 \\ 0,0066 & 0,0131 & -0,0062 & 0,1029 & 0,0046 & -0,0121 & 0,0057 \\ -0,0101 & 0,0032 & -0,0002 & 0,0046 & 0,1239 & -0,0047 & -0,0026 \\ -0,0061 & -0,0087 & 0,0042 & -0,0121 & -0,0047 & 0,1491 & -0,0249 \\ 0,0087 & 0,0099 & -0,0067 & 0,0057 & -0,0026 & -0,0249 & 0,1649 \end{bmatrix}$$

F_2

$$= \begin{bmatrix} 0,0952 & 0,0023 & 0,0028 & 0,0020 & -0,0079 & -0,0046 & 0,0065 \\ 0,0023 & 0,1163 & -0,0097 & 0,0014 & -0,0027 & 0,0086 & -0,0023 \\ 0,0028 & -0,0097 & 0,1162 & 0,0004 & -0,0100 & -0,0029 & 0,0111 \\ 0,0020 & 0,0014 & 0,0004 & 0,0896 & -0,0016 & -0,0051 & 0,0052 \\ -0,0079 & -0,0027 & -0,0100 & -0,0016 & 0,1275 & -0,0048 & -0,0086 \\ -0,0046 & 0,0086 & -0,0029 & -0,0051 & -0,0048 & 0,1298 & -0,0139 \\ 0,0065 & -0,0023 & 0,0111 & 0,0052 & -0,0086 & -0,0139 & 0,1658 \end{bmatrix}$$

4. Jarak Mahalanobis

Perhitungan jarak *Mahalanobis* pada variasi $m = 1,9$ dilakukan dengan prosedur yang sama seperti pada variasi $m = 1,1$, $m = 1,3$, $m = 1,5$, dan $m = 1,7$, yaitu dengan membentuk A_k dari matriks kovarian masing-masing *cluster*, kemudian menghitung jarak setiap data terhadap pusat *cluster*. Untuk $c = 2$, hasil jarak *Mahalanobis* antara setiap data dengan pusat *cluster* 1 dan *cluster* 2 diperlihatkan pada Tabel 4.24.

Tabel 4.24 Hasil Jarak *Mahalanobis* Dua *Cluster* ($m = 1,9$)

Data	Cluster 1	Cluster 2
1	0,5198	0,5105
2	0,8905	0,8089
3	0,8068	0,8019
$\vdots$	$\vdots$	$\vdots$
502	1,0313	0,9522

5. Fungsi Objektif

Perhitungan nilai fungsi objektif pada variasi $m = 1,9$ dilakukan dengan prosedur yang sama seperti pada variasi $m = 1,1$, $m = 1,3$, $m = 1,5$, dan $m = 1,7$ dengan menggunakan derajat keanggotaan dan jarak *Mahalanobis* pada masing-masing *cluster*. Hasil perhitungan untuk $c = 2$ sebagai berikut:

Perhitungan *cluster* 1

$$\begin{aligned}
 \sum (\mu_{i,k})^m D_{i,kA_k}^2 &= (\mu_{1,1})^{1,9} D_{1,1A_1}^2 + \dots + (\mu_{502,1})^{1,9} D_{502,1A_1}^2 \\
 &= (0,6381)^{1,9} (0,5198)^2 + \dots + (0,4586)^{1,9} (1,0313)^2 \\
 &= 98,5446
 \end{aligned}$$

Perhitungan *cluster* 2

$$\begin{aligned}
 \sum (\mu_{i,k})^m D_{i,kA_k}^2 &= (\mu_{1,2})^{1,9} D_{1,2A_2}^2 + \dots + (\mu_{502,2})^{1,9} D_{502,2A_2}^2 \\
 &= (0,3619)^{1,9} (0,5105)^2 + \dots + (0,5414)^{1,9} (0,9522)^2 \\
 &= 110,3936
 \end{aligned}$$

Maka untuk fungsi objektif pada iterasi ke-1 ini adalah:

$$\begin{aligned}
 J_{GK}^{(1)} &= \sum_{i=1}^{502} \sum_{k=1}^2 (\mu_{i,k})^{1,9} D_{i,kA_k}^2 \\
 &= 127,7111 + 139,7733 = 267,4843
 \end{aligned}$$

6. Matriks Keanggotaan Baru

Pada variasi $m = 1,9$ pembaruan matriks keanggotaan dilakukan dengan prosedur yang sama seperti $m = 1,1$, $m = 1,3$, $m = 1,5$, dan $m = 1,7$, yaitu berdasarkan jarak *Mahalanobis* setiap data terhadap pusat *cluster* dan nilai *weighting exponent*. Sebagai contoh perhitungan, derajat keanggotaan data pertama terhadap *cluster* 1 dan *cluster* 2 dihitung sebagai berikut:

Dengan nilai $m = 1,9$, maka nilai pangkat:

$$\frac{2}{m-1} = \frac{2}{1,9-1} = 2,2222$$

Derajat keanggotaan data pertama terhadap *cluster* 1

$$\begin{aligned}\mu_{1,1}^{(t+1)} &= \left[\left(\frac{D_{1,1A_1}}{D_{1,1A_1}} \right)^{2,2222} + \left(\frac{D_{1,1A_1}}{D_{1,2A_2}} \right)^{2,2222} \right]^{-1} \\ &= \left[1 + \left(\frac{0,5198}{0,5105} \right)^{2,2222} \right]^{-1} = 0,4900\end{aligned}$$

Derajat keanggotaan data pertama terhadap *cluster* 2

$$\begin{aligned}\mu_{1,2}^{(t+1)} &= \left[\left(\frac{D_{1,2A_2}}{D_{1,1A_1}} \right)^{2,2222} + \left(\frac{D_{1,2A_2}}{D_{1,2A_2}} \right)^{2,2222} \right]^{-1} \\ &= \left[\left(\frac{0,5105}{0,5198} \right)^{2,2222} + 1 \right]^{-1} = 0,5100\end{aligned}$$

Beberapa hasil pembaruan derajat keanggotaan untuk dua *cluster* ditampilkan pada Tabel 4.25.

Tabel 4.25 Hasil Matriks Keanggotaan Baru dengan Dua *Cluster* ($m = 1,9$)

Data	<i>Cluster 1</i>	<i>Cluster 2</i>
1	0,4900	0,5100
2	0,4469	0,5531
3	0,4966	0,5034
$\vdots$	$\vdots$	$\vdots$
502	0,4558	0,5442

7. Kriteria Konvergensi

Pada variasi $m = 1,9$ pemeriksaan konvergensi dilakukan dengan prosedur yang sama seperti $m = 1,1$, $m = 1,3$, $m = 1,5$, dan $m = 1,7$, yaitu membandingkan perubahan matriks keanggotaan pada dua iterasi berturut-turut, $\|U_{t+1} - U_t\|$, terhadap nilai *threshold* $\varepsilon = 0,0001$. Sebagai hasil perhitungan pada iterasi awal, diperoleh:

Perhitungan untuk *cluster 1*:

$$\begin{aligned}\|U_{t+1} - U_t\| &= \sqrt{(0,4900 - 0,6381)^2 + \dots + (0,4558 - 0,4586)^2} \\ &= \sqrt{27,2784}\end{aligned}$$

Perhitungan untuk *cluster 2*:

$$\begin{aligned}\|U_{t+1} - U_t\| &= \sqrt{(0,5100 - 0,3619)^2 + \dots + (0,5442 - 0,5414)^2} \\ &= \sqrt{27,2784}\end{aligned}$$

Karena perhitungan dilakukan dengan menjumlahkan hasil dari seluruh *cluster*, maka nilai akhirnya diperoleh sebagai berikut:

$$\begin{aligned}\|U_{t+1} - U_t\| &= \sqrt{27,2784 + 27,2784} = \sqrt{54,5568} \\ \|U_{t+1} - U_t\| &= 7,3863\end{aligned}$$

Nilai tersebut masih jauh lebih besar dari $\varepsilon = 0,0001$, maka pada iterasi awal algoritma belum dinyatakan konvergen. Proses iterasi kemudian dilanjutkan hingga perubahan matriks keanggotaan semakin kecil. Setelah

iterasi ke-24, diperoleh nilai:

$$\|U_{t+1} - U_t\| = 0,00009035$$

Dengan demikian, pada iterasi tersebut algoritma FGK dengan $m = 1,9$ telah mencapai konvergensi.

8. Hasil *Clustering*

Setelah algoritma *Fuzzy Gustafson-Kessel* dengan $m = 1,9$ mencapai konvergensi, penentuan hasil *clustering* data didasarkan pada matriks keanggotaan akhir. Setiap data dikelompokkan ke dalam *cluster* yang memiliki nilai derajat keanggotaan terbesar, sehingga *cluster* tersebut menjadi *cluster* dominan yang paling merepresentasikan data. Hasil *clustering* data berdasarkan *cluster* dominan untuk $c = 2$ dan $m = 1,9$ disajikan pada Tabel 4.26.

Tabel 4.26 Hasil *Clustering* Dua *Cluster* ($m = 1,9$)

Data	Derajat Keanggotaan		Cluster
	1	2	
1	0,50000003	0,49999997	1
2	0,49997940	0,50002060	2
3	0,49999739	0,50002060	2
⋮	⋮	⋮	⋮
502	0,50000547	0,49999453	1

4.3.6 Perhitungan untuk Variasi Nilai $m = 2,1$

Pada variasi ini, proses *clustering* dilakukan dengan *weighting exponent* $m = 2,1$ dan mengikuti tahapan yang sama seperti pada variasi m sebelumnya (1,1 , 1,3 , 1,5 , 1,7 , dan 1,9). Perhitungan manual difokuskan pada kasus $c = 2$ dengan menggunakan variabel Umur (X_1).

1. Inisialisasi Matriks Keanggotaan Awal

Pada variasi nilai $m = 2,1$, proses diawali dengan inisialisasi U_0 . Matriks

U_0 yang digunakan sama dengan variasi m lainnya (1,1 , 1,3 , 1,5 , 1,7 , dan 1,9) karena pembentukannya tidak dipengaruhi oleh nilai m . Untuk $c = 2$ hingga $c = 5$, digunakan matriks inisialisasi awal yang telah disajikan pada subbab sebelumnya.

2. Pusat *Cluster*

Perhitungan pusat *cluster* awal dilakukan dengan menggunakan matriks U_0 .

Sebagai contoh manual, perhitungan ditunjukkan untuk $c = 2$ menggunakan variabel pertama, yaitu Umur (X_1).

Perhitungan pusat untuk *cluster* 1 dapat dituliskan sebagai berikut:

$$v_{1,1} = \frac{\sum_{i=1}^n (\mu_{i,1})^{2,1} x_{i,1}}{\sum_{i=1}^n (\mu_{i,1})^{2,1}}$$

$$v_{1,1} = \frac{(\mu_{1,1})^{2,1} x_{1,1} + (\mu_{2,1})^{2,1} x_{2,1} + \dots + (\mu_{502,1})^{2,1} x_{502,1}}{(\mu_{1,1})^{2,1} + (\mu_{2,1})^{2,1} + \dots + (\mu_{502,1})^{2,1}}$$

$$v_{1,1} = \frac{(0,6381)^{2,1} 0,625 + (0,5518)^{2,1} 0,625 + \dots + (0,4586)^{2,1} 0}{(0,6381)^{2,1} + (0,5518)^{2,1} + \dots + (0,4586)^{2,1}}$$

$$v_{1,1} = \frac{71,7955}{137,6179} = 0,5217$$

Perhitungan pusat untuk *cluster* 2:

$$v_{2,1} = \frac{\sum_{i=1}^n (\mu_{i,2})^{2,1} x_{i,1}}{\sum_{i=1}^n (\mu_{i,2})^{2,1}}$$

$$v_{2,1} = \frac{(\mu_{1,2})^{2,1} x_{1,1} + (\mu_{2,2})^{2,1} x_{2,1} + \dots + (\mu_{502,2})^{2,1} x_{502,1}}{(\mu_{1,2})^{2,1} + (\mu_{2,2})^{2,1} + \dots + (\mu_{502,2})^{2,1}}$$

$$v_{2,1} = \frac{(0,3619)^{2,1} 0,625 + (0,4482)^{2,1} 0,625 + \dots + (0,5414)^{2,1} 0}{(0,3619)^{2,1} + (0,4482)^{2,1} + \dots + (0,5414)^{2,1}}$$

$$v_{2,1} = \frac{80,7651}{157,4644} = 0,5129$$

Nilai tersebut merupakan pusat cluster untuk variabel Umur (X_1) pada *cluster* 1 dan *cluster* 2 dengan nilai $m = 2,1$. Perhitungan yang sama dilakukan pada variabel lainnya sehingga setiap pusat *cluster* berbentuk vektor berdimensi tujuh. Berdasarkan hasil perhitungan, pusat *cluster* untuk $c = 2$ dapat dinyatakan sebagai:

$$v_1 = (0,5217, 0,4845, 0,4929, 0,5428, 0,4834, 0,4919, 0,5052)$$

$$v_2 = (0,5129, 0,5131, 0,5305, 0,5263, 0,4811, 0,5147, 0,4657)$$

Setiap komponen vektor menyatakan nilai untuk variabel X_1 sampai dengan X_7 , dan nilai tersebut juga disajikan dalam Tabel 4.27.

Tabel 4.27 Hasil Pusat Dua *Cluster* ($m = 2,1$)

Variabel	Pusat <i>Cluster</i>	
	<i>Cluster</i> 1	<i>Cluster</i> 2
X_1	0,5217	0,5129
X_2	0,4845	0,5131
X_3	0,4929	0,5305
X_4	0,5428	0,5263
X_5	0,4834	0,4811
X_6	0,4919	0,5147
X_7	0,5052	0,4657

3. Matriks Kovarian

Perhitungan matriks kovarian pada variasi $m = 2,1$ dilakukan dengan prosedur yang sama seperti pada variasi $m = 1,1$, $m = 1,3$, $m = 1,5$, $m = 1,7$, dan $m = 1,9$. Berdasarkan pusat *cluster* dan *weighting exponent*, diperoleh hasil akhir matriks kovarian untuk masing-masing *cluster* sebagai berikut:

F_1

$$= \begin{bmatrix} 0,0915 & 0,0034 & -0,0080 & 0,0066 & -0,0098 & -0,0061 & 0,0088 \\ 0,0034 & 0,1254 & -0,0143 & 0,0138 & 0,0033 & -0,0096 & 0,0107 \\ -0,0080 & -0,0143 & 0,0179 & -0,0068 & -0,0003 & 0,0053 & -0,0075 \\ 0,0066 & 0,0138 & -0,0068 & 0,1031 & 0,0047 & -0,0126 & 0,0059 \\ -0,0098 & 0,0033 & -0,0003 & 0,0047 & 0,1234 & -0,0049 & -0,0026 \\ -0,0062 & -0,0096 & 0,0053 & -0,0126 & -0,0049 & 0,1498 & -0,0253 \\ 0,0088 & 0,0107 & -0,0075 & 0,0059 & -0,0026 & -0,0253 & 0,1652 \end{bmatrix}$$

 F_2

$$= \begin{bmatrix} 0,0953 & 0,0022 & 0,0028 & 0,0018 & -0,0074 & -0,0047 & 0,0063 \\ 0,0022 & 0,1160 & -0,0095 & 0,0013 & -0,0029 & 0,0089 & -0,0023 \\ 0,0028 & -0,0095 & 0,1161 & 0,0002 & -0,0103 & -0,0024 & 0,0115 \\ 0,0018 & 0,0013 & 0,0002 & 0,0889 & -0,0019 & -0,0051 & 0,0053 \\ -0,0074 & -0,0029 & -0,0103 & -0,0019 & 0,1274 & -0,0048 & -0,0090 \\ -0,0047 & 0,0089 & -0,0024 & -0,0051 & -0,0048 & 0,1293 & -0,0136 \\ 0,0063 & -0,0023 & 0,0115 & 0,0053 & -0,0090 & -0,0136 & 0,1661 \end{bmatrix}$$

4. Jarak Mahalanobis

Perhitungan jarak *Mahalanobis* pada variasi $m = 2,1$ dilakukan dengan prosedur yang sama seperti pada variasi $mm = 1,1$, $m = 1,3$, $m = 1,5$, $m = 1,7$, dan $m = 1,9$, yaitu dengan membentuk A_k dari matriks kovarian masing-masing *cluster*, kemudian menghitung jarak setiap data terhadap pusat *cluster*. Untuk $c = 2$, hasil jarak *Mahalanobis* antara setiap data dengan pusat *cluster* 1 dan *cluster* 2 diperlihatkan pada Tabel 4.28.

Tabel 4.28 Hasil Jarak *Mahalanobis* Dua Cluster ($m = 2,1$)

Data	Cluster 1	Cluster 2
1	0,5216	0,5116
2	0,8949	0,8079
3	0,8052	0,8009
⋮	⋮	⋮
502	1,0338	0,9496

5. Fungsi Objektif

Perhitungan nilai fungsi objektif pada variasi $m = 2,1$ dilakukan dengan prosedur yang sama seperti pada variasi $m = 1,1$, $m = 1,3$, $m = 1,5$,

$m = 1,7$, dan $m = 1,9$ dengan menggunakan derajat keanggotaan dan jarak *Mahalanobis* pada masing-masing *cluster*. Hasil perhitungan untuk $c = 2$ sebagai berikut:

Perhitungan *cluster* 1

$$\begin{aligned}\sum (\mu_{i,k})^m D_{i,kA_k}^2 &= (\mu_{1,1})^{2,1} D_{1,1A_1}^2 + \dots + (\mu_{502,1})^{2,1} D_{502,1A_1}^2 \\ &= (0,6381)^{2,1} (0,5216)^2 + \dots + (0,4586)^{2,1} (1,0338)^2 \\ &= 116,6899\end{aligned}$$

Perhitungan *cluster* 2

$$\begin{aligned}\sum (\mu_{i,k})^m D_{i,kA_k}^2 &= (\mu_{1,2})^{2,1} D_{1,2A_2}^2 + \dots + (\mu_{502,2})^{2,1} D_{502,2A_2}^2 \\ &= (0,3619)^{2,1} (0,5116)^2 + \dots + (0,5414)^{2,1} (0,9496)^2 \\ &= 128,6835\end{aligned}$$

Maka untuk fungsi objektif pada iterasi ke-1 ini adalah:

$$\begin{aligned}J_{GK}^{(1)} &= \sum_{i=1}^{502} \sum_{k=1}^2 (\mu_{i,k})^{2,1} D_{i,kA_k}^2 \\ &= 116,6899 + 128,6835 = 245,3734\end{aligned}$$

6. Matriks Keanggotaan Baru

Pada variasi $m = 2,1$ pembaruan matriks keanggotaan dilakukan dengan prosedur yang sama seperti $m = 1,1$, $m = 1,3$, $m = 1,5$, $m = 1,7$, dan $m = 1,9$, yaitu berdasarkan jarak *Mahalanobis* setiap data terhadap pusat *cluster* dan nilai m . Sebagai contoh perhitungan, derajat keanggotaan data pertama terhadap *cluster* 1 dan *cluster* 2 dihitung sebagai berikut:

Dengan nilai $m = 2,1$, maka nilai pangkat:

$$\frac{2}{2,1 - 1} = \frac{2}{2,1 - 1} = 1,8182$$

Derajat keanggotaan data pertama terhadap *cluster* 1

$$\begin{aligned}\mu_{1,1}^{(t+1)} &= \left[\left(\frac{D_{1,1A_1}}{D_{1,1A_1}} \right)^{1,8182} + \left(\frac{D_{1,1A_1}}{D_{1,2A_2}} \right)^{1,8182} \right]^{-1} \\ &= \left[1 + \left(\frac{0,5216}{0,5116} \right)^{1,8182} \right]^{-1} = 0,4913\end{aligned}$$

Derajat keanggotaan data pertama terhadap *cluster* 2

$$\begin{aligned}\mu_{1,2}^{(t+1)} &= \left[\left(\frac{D_{1,2A_2}}{D_{1,1A_1}} \right)^{1,8182} + \left(\frac{D_{1,2A_2}}{D_{1,2A_2}} \right)^{1,8182} \right]^{-1} \\ &= \left[\left(\frac{0,5116}{0,5216} \right)^1 + 1 \right]^{-1} = 0,5087\end{aligned}$$

Beberapa hasil pembaruan derajat keanggotaan untuk dua *cluster* ditampilkan pada Tabel 4.29.

Tabel 4.29 Hasil Matriks Keanggotaan Baru dengan Dua *Cluster* ($m = 2,1$)

Data	<i>Cluster</i> 1	<i>Cluster</i> 2
1	0,4913	0,5087
2	0,4537	0,5463
3	0,4975	0,5025
⋮	⋮	⋮
502	0,4615	0,5385

7. Kriteria Konvergensi

Pada variasi $m = 2,1$ pemeriksaan konvergensi dilakukan dengan prosedur yang sama seperti $m = 1,1$, $m = 1,3$, $m = 1,5$, $m = 1,7$, dan $m = 1,9$, yaitu membandingkan perubahan matriks keanggotaan pada dua iterasi berturut-turut, $\|U_{t+1} - U_t\|$, terhadap nilai *threshold* $\varepsilon = 0,0001$. Sebagai hasil perhitungan pada iterasi awal, diperoleh:

Perhitungan untuk *cluster* 1:

$$\|U_{t+1} - U_t\| = \sqrt{(0,4913 - 0,6381)^2 + \dots + (0,4615 - 0,4586)^2}$$

$$= \sqrt{27,3815}$$

Perhitungan untuk *cluster* 2:

$$\begin{aligned} \|U_{t+1} - U_t\| &= \sqrt{(0,5087 - 0,3619)^2 + \dots + (0,5385 - 0,5414)^2} \\ &= \sqrt{27,3815} \end{aligned}$$

Karena perhitungan dilakukan dengan menjumlahkan hasil dari seluruh *cluster*, maka nilai akhirnya diperoleh sebagai berikut:

$$\begin{aligned} \|U_{t+1} - U_t\| &= \sqrt{27,3815 + 27,3815} = \sqrt{54,7630} \\ \|U_{t+1} - U_t\| &= 7,4002 \end{aligned}$$

Nilai tersebut masih jauh lebih besar dari $\varepsilon = 0,0001$, maka pada iterasi awal algoritma belum dinyatakan konvergen. Proses iterasi kemudian dilanjutkan hingga perubahan matriks keanggotaan semakin kecil. Setelah iterasi ke-19, diperoleh nilai:

$$\|U_{t+1} - U_t\| = 0,000085$$

Dengan demikian, pada iterasi tersebut algoritma FGK dengan $m = 2,1$ telah mencapai konvergensi.

8. Hasil *Clustering*

Setelah algoritma *Fuzzy Gustafson-Kessel* dengan $m = 2,1$ mencapai konvergensi, penentuan hasil *clustering* data didasarkan pada matriks keanggotaan akhir. Setiap data dikelompokkan ke dalam *cluster* yang memiliki nilai derajat keanggotaan terbesar, sehingga *cluster* tersebut menjadi *cluster* dominan yang paling merepresentasikan data. Hasil *clustering* data berdasarkan *cluster* dominan untuk $c = 2$ dan $m = 3$ disajikan pada Tabel 4.30.

Tabel 4.30 Hasil *Clustering* Dua *Cluster* ($m = 2,1$)

Data	Derajat Keanggotaan		Cluster
	1	2	
1	0,500001	0,499999	1
2	0,499986	0,500014	2
3	0,500002	0,499999	1
⋮	⋮	⋮	⋮
502	0,5000003	0,499997	1

4.4 Evaluasi Hasil *Clustering* dengan *Xie-Beni Index*

Evaluasi hasil *clustering* dilakukan untuk menganalisis keterkaitan variasi nilai *weighting exponent* (m) terhadap kualitas *cluster* yang dihasilkan oleh algoritma *Fuzzy Gustafson-Kessel* (FGK). Parameter m berperan dalam mengatur tingkat kekaburan derajat keanggotaan data terhadap masing-masing *cluster*, sehingga perubahan nilai m bisa memengaruhi tingkat kekompakan dan keterpisahan *cluster*. Selain itu, evaluasi ini juga mempertimbangkan variasi banyaknya *cluster* (c) sebagai parameter pendukung, yaitu $c = 2, 3, 4$, dan 5 .

Variasi banyaknya *cluster* digunakan sebagai parameter pembanding untuk melihat pola perubahan nilai *Xie-Beni Index* pada setiap nilai *weighting exponent* pada struktur *cluster* yang berbeda. Kualitas *cluster* dievaluasi menggunakan *Xie-Beni Index* (XB), di mana nilai XB yang lebih kecil menunjukkan *cluster* yang lebih kompak dan memiliki pemisahan antarklaster yang lebih baik.

Evaluasi dilakukan dengan menggunakan pusat *cluster* dan matriks keanggotaan yang diperoleh pada iterasi terakhir setelah algoritma FGK mencapai kondisi konvergen. Untuk setiap kombinasi nilai m dan c , nilai *Xie-Beni Index* dihitung berdasarkan hasil *clustering* tersebut. Untuk memperjelas proses evaluasi secara matematis, ditunjukkan satu contoh perhitungan *Xie-Beni Index* pada salah

satu kombinasi parameter, yaitu nilai $m = 1,3$ dan $c = 2$ dengan menggunakan persamaan (2.9).

Komponen yang diperlukan dalam perhitungan *Xie-Beni Index* meliputi *dataset* hasil *preprocessing*, matriks keanggotaan pada iterasi konvergen, dan pusat *cluster* dari matriks keanggotaan pada iterasi konvergen. Data yang digunakan dalam proses evaluasi yang telah ditunjukkan pada Tabel 4.6. Kemudian, untuk matriks keanggotaan ketika $c = 2$ dan $m = 1,3$ ditunjukkan pada Tabel 4.31.

Tabel 4.31 Matriks Keanggotaan untuk Evaluasi *Xie-Beni Index*

Cluster 1	Cluster 2
0,002722	0,997278
0,006571	0,993429
0,430502	0,569498
⋮	⋮
0,769612	0,230388

Berdasarkan matriks keanggotaan pada iterasi konvergen yang ditunjukkan pada Tabel 4.27, selanjutnya dilakukan perhitungan pusat *cluster* untuk masing-masing *cluster*. Pusat *cluster* ini merupakan salah satu komponen utama yang diperlukan dalam perhitungan *Xie-Beni Index*. Adapun untuk perhitungan pusat *cluster* ketika $c = 2$, $m = 1,3$ disajikan berikut ini.

Perhitungan untuk pusat *cluster* 1 dapat dituliskan sebagai berikut:

$$v_{1,1} = \frac{\sum_{i=1}^n (\mu_{i,1})^{1,3} x_{i,1}}{\sum_{i=1}^n (\mu_{i,1})^{1,3}}$$

$$v_{1,1} = \frac{(\mu_{1,1})^{1,3} x_{1,1} + (\mu_{2,1})^{1,3} x_{2,1} + \dots + (\mu_{502,1})^{1,3} x_{502,1}}{(\mu_{1,1})^{1,3} + (\mu_{2,1})^{1,3} + \dots + (\mu_{502,1})^{1,3}}$$

$$v_{1,1} = \frac{(0,002722)^{1,3} 0,625 + (0,006571)^{1,3} 0,625 + \dots + (0,769612)^{1,3} 0}{(0,002722)^{1,3} + (0,006571)^{1,3} + \dots + (0,769612)^{1,3}}$$

$$v_{1,1} = \frac{120,505}{240,515} = 0,50103$$

Untuk perhitungan pusat *cluster* 2 dilakukan dengan cara yang sama, sehingga hasil untuk pusat dua *cluster* disajikan pada Tabel 4.32 di bawah ini.

Tabel 4.32 Pusat *Cluster* untuk Evaluasi *Xie-Beni Index*

Variabel	Pusat <i>Cluster</i>	
	<i>Cluster 1</i>	<i>Cluster 2</i>
X ₁	0,50103	0,52660
X ₂	0,54225	0,46112
X ₃	0,23558	0,81129
X ₄	0,50980	0,56043
X ₅	0,50858	0,45665
X ₆	0,52117	0,48959
X ₇	0,46861	0,52226

Perhitungan nilai *Xie-Beni Index* diawali dengan menghitung nilai S berdasarkan persamaan (2.9). Nilai S digunakan untuk mengukur tingkat kekompakan internal *cluster* (*intra-class compactness*). Nilai S dihitung dengan menggunakan persamaan dibawah ini.

$$S = \frac{\sum_{k=1}^c \sum_{i=1}^n (\mu_{i,k})^m \|x_i - v_k\|^2}{n}$$

Pada contoh perhitungan dengan $c = 2$ dan nilai $m = 1,3$ tahapan perhitungan ditunjukkan sebagai berikut:

$$S = \frac{\sum_{k=1}^2 \sum_{i=1}^{502} (\mu_{i,k})^m \|x_i - v_k\|^2}{n}$$

$$S = \frac{\sum_{k=1}^2 \sum_{i=1}^{502} (\mu_{i,k})^m \sum_{j=1}^7 (x_{i,j} - v_{k,j})^2}{n}$$

$$S = \frac{1}{502} \left[\begin{aligned} & \left[(\mu_{1,1})^3 \left((x_{1,1} - v_{1,1})^2 + \dots + (x_{1,7} - v_{1,7})^2 \right) + \dots + (\mu_{502,1})^3 \right. \\ & \quad \left. \left((x_{502,1} - v_{1,1})^2 + \dots + (x_{502,7} - v_{1,7})^2 \right) \right] + \\ & \left[(\mu_{1,2})^3 \left((x_{1,1} - v_{2,1})^2 + \dots + (x_{1,7} - v_{2,7})^2 \right) + \dots + (\mu_{502,2})^3 \right. \\ & \quad \left. \left((x_{502,1} - v_{2,1})^2 + \dots + (x_{502,7} - v_{2,7})^2 \right) \right] \end{aligned} \right]$$

$$\begin{aligned}
&= \frac{1}{502} [(0,002722)^{1.3} ((0,625 - 0,50103)^2 + \dots + (0,5 - 0,46861)^2) \\
&\quad + \dots + (0,769612)^3 ((0 - 0,50103)^2) + \dots + (0 - 0,46861)^2)] + \\
&\quad [(0,997278)^3 ((0,625 - 0,52660)^2) + \dots + (0,5 - 0,52226)^2) + \\
&\quad \dots + (0,230388)^3 ((0 - 0,52660)^2 + \dots + (0 - 0,52226)^2)] \\
&= \frac{1}{502} [(0,00046)((0,01537) + \dots + (0,00099)) + \dots + (0,71146) \\
&\quad ((0,25102) + \dots + (0,21960))] + [(0,99646)((0,00968) + \dots + \\
&\quad (0,00049)) + \dots + (0,14832)((0,27730) + \dots + (0,27275))] \\
&= \frac{1}{502} [(0,00024) + \dots + (0,74442)] + [(0,16791) + \dots + (0,18454)] \\
&= \frac{1}{502} [186,0185 + 177,0715] = \frac{1}{502} [363,0901]
\end{aligned}$$

$$S = 0,723287$$

Setelah diperoleh nilai S , langkah selanjutnya berdasarkan persamaan (2.9) adalah menghitung nilai d_{min} . Nilai ini dinyatakan sebagai jarak minimum antar pusat *cluster* yang terbentuk. Nilai d_{min} dihitung dengan menggunakan persamaan berikut ini.

$$d_{min} = \min_{i \neq k} \|v_i - v_k\|^2$$

Pada contoh perhitungan dengan $c = 2$, nilai d_{min} dihitung menggunakan pusat *cluster* pada iterasi konvergen yang disajikan pada Tabel 4.28. Karena hanya terdapat dua *cluster*, maka nilai d_{min} diperoleh langsung dari jarak kuadrat antara pusat *cluster* 1 dan pusat *cluster* 2. Karena pusat *cluster* terdiri dari 7 variabel, maka perhitungan jarak kuadrat antarpusat *cluster* ditunjukkan sebagai berikut:

$$\|v_i - v_k\|^2 = \sum_{j=1}^7 (v_{1,j} - v_{2,j})^2$$

$$\begin{aligned}
d_{min} &= (v_{1,1} - v_{2,1})^2 + (v_{1,2} - v_{2,2})^2 + (v_{1,3} - v_{2,3})^2 + (v_{1,4} - v_{2,4})^2 + \\
&\quad (v_{1,5} - v_{2,5})^2 + (v_{1,6} - v_{2,6})^2 + (v_{1,7} - v_{2,7})^2 \\
&= (0,50103 - 0,52660)^2 + (0,54225 - 0,46111)^2 + \\
&\quad (0,23558 - 0,811286)^2 + \dots + (0,46861 - 0,52226)^2 \\
&= 0,00065 + 0,00658 + 0,331432 + \dots + 0,002878 \\
&= 0,347804
\end{aligned}$$

Ketika nilai S dan d_{min} sudah didapatkan, langkah selanjutnya adalah menghitung nilai *Xie-Beni Index* sesuai persamaan (2.9).

$$\begin{aligned}
v_{XB} &= \frac{S}{d_{min}} \\
v_{XB} &= \frac{0,723287}{0,347804} \\
v_{XB} &= 2,079584
\end{aligned}$$

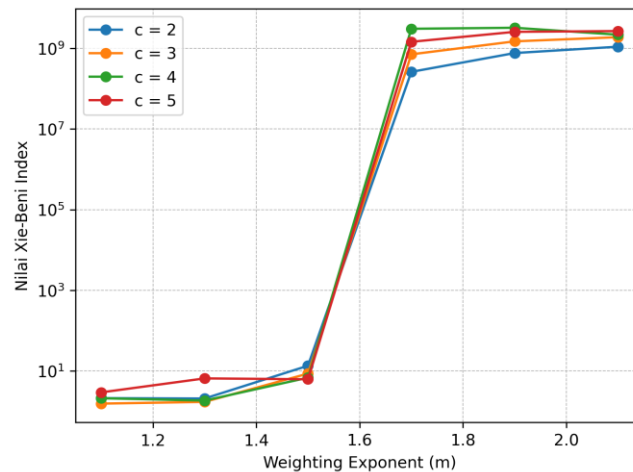
Perhitungan nilai *Xie-Beni Index* dilakukan untuk seluruh kombinasi parameter m dan c yang digunakan dalam penelitian ini dengan prosedur yang sama seperti perhitungan sebelumnya. Hasil perhitungan tersebut dirangkum dan disajikan pada Tabel 4.33.

Tabel 4.33 Hasil Evaluasi *Clustering* dengan *Xie-Beni Index*

Banyaknya <i>Cluster</i> (<i>c</i>)	<i>Weighting Exponent</i> (<i>m</i>)	Hasil Nilai <i>Xie-Beni Index</i>
2	1,1	2,1243402611
	1,3	2,0795837205
	1,5	13,5108918119
	1,7	262407710,05642
	1,9	767503590,26202
	2,1	1103379379,3075
3	1,1	1,5542301750
	1,3	1,7156973536
	1,5	8,6214837453
	1,7	713761178,63209
	1,9	1496369014,1610
	2,1	1899745079,1890
4	1,1	2,1092060904
	1,3	1,8476430912
	1,5	6,8329067408
	1,7	3071516307,3882
	1,9	3256546739,4304
	2,1	2190180230,1422
5	1,1	2,9550053889
	1,3	6,5562159904
	1,5	6,2393741202
	1,7	1457974274,5305
	1,9	2576713854,7351
	2,1	2693981637,9315

Berdasarkan Tabel 4.29, diperoleh nilai *Xie-Beni Index* (XB) untuk setiap kombinasi *weighting exponent* (*m*) dan banyaknya *cluster* (*c*). Nilai XB terkecil secara keseluruhan diperoleh pada kombinasi $c = 3$ dan $m = 1,1$ yaitu sebesar 1,55423. Selain itu, nilai XB yang relatif kecil juga diperoleh pada $m = 1,3$, khususnya pada $c = 3$ dan $c = 4$. Sementara itu, pada nilai *m* yang lebih besar, nilai XB cenderung lebih tinggi dibandingkan pada *m* kecil untuk seluruh variasi banyaknya *cluster*. Hal ini menunjukkan bahwa pemilihan nilai *m* memengaruhi kualitas hasil *clustering*.

Untuk memperjelas pola perubahan nilai *Xie-Beni Index* terhadap variasi *weighting exponent* (m), hasil evaluasi pada Tabel 4.29 selanjutnya divisualisasikan dalam bentuk grafik.



Gambar 4.1 Grafik *Xie-Beni Index* terhadap Variasi *Weighting Exponent*

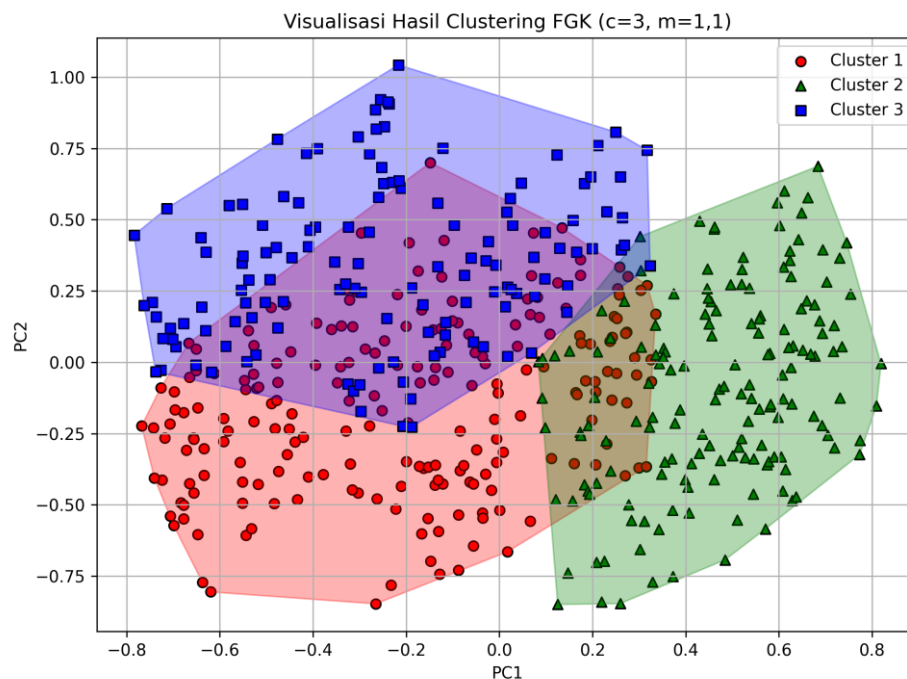
Gambar 4.1 menunjukkan pola perubahan nilai *Xie-Beni Index* (XB) terhadap variasi *weighting exponent* (m). Berdasarkan grafik tersebut, terlihat bahwa nilai XB cenderung meningkat seiring bertambahnya nilai m , dengan lonjakan nilai XB pada $m \geq 1,7$.

Lonjakan tersebut menunjukkan bahwa semakin besar nilai m , derajat keanggotaan menjadi semakin kabur, sehingga menurunkan kekompakan dan kejelasan pemisahan antar *cluster*. Selain itu, pola yang relatif konsisten pada seluruh kurva menunjukkan bahwa parameter m memiliki pengaruh yang lebih dominan dibandingkan variasi banyaknya *cluster* dalam menentukan kualitas hasil *clustering*.

Berdasarkan hasil evaluasi tersebut, dapat disimpulkan bahwa kombinasi parameter optimal dalam penelitian ini adalah $c = 3$ dan $m = 1,1$, karena menghasilkan nilai *Xie-Beni Index* paling minimum dibandingkan kombinasi parameter lainnya.

4.5 Visualisasi dan Interpretasi Hasil *Clustering* dengan Parameter Optimal

Hasil *clustering* dengan parameter optimal selanjutnya dianalisis melalui visualisasi menggunakan metode reduksi dimensi *Principal Component Analysis* (PCA). Visualisasi ini bertujuan untuk menampilkan sebaran data dan pemisahan antarklaster pada ruang dua dimensi untuk kombinasi parameter $c = 3$ dan $m = 1,1$, sebagaimana ditunjukkan pada gambar berikut.



Gambar 4.2 Visualisasi Hasil *Clustering*

Penggunaan PCA dalam penelitian ini didasarkan pada karakteristik data yang bersifat multivariat dengan tujuh variabel. Secara matematis, data tersebut berada pada ruang berdimensi tujuh sehingga tidak dapat diamati secara langsung. Visualisasi menggunakan sebagian variabel saja berpotensi menghilangkan informasi penting dan tidak merepresentasikan struktur data secara menyeluruh. Oleh karena itu, diperlukan teknik reduksi dimensi yang mampu merangkum informasi utama dari seluruh variabel ke dalam representasi yang lebih sederhana.

PCA dipilih karena mentransformasikan variabel asli menjadi komponen utama yang bersifat ortogonal dan disusun berdasarkan proporsi variasi terbesar yang mampu dijelaskan. Dua komponen utama pertama umumnya telah merepresentasikan sebagian besar struktur variansi dalam data, sehingga proyeksi ke ruang dua dimensi tetap mempertahankan pola global dan hubungan antardata. Dengan demikian, visualisasi hasil clustering pada ruang PCA dua dimensi bukan sekadar penyederhanaan tampilan, tetapi merupakan proyeksi matematis yang mempertahankan informasi dominan dari keseluruhan variabel. Hal ini memungkinkan evaluasi visual terhadap kekompakan dan keterpisahan antarkluster secara lebih representatif dan selaras dengan hasil evaluasi numerik yang diperoleh sebelumnya.

Berdasarkan Gambar 4.2, hasil *clustering* FGK dengan parameter $c = 3$ dan $m = 1,1$ membagi data pelajar ke dalam tiga *cluster*, yaitu *cluster* 1 sebanyak 209 pelajar, *cluster* 2 sebanyak 161 pelajar, dan *cluster* 3 sebanyak 132 pelajar. Untuk mengetahui karakteristik masing-masing *cluster*, dilakukan analisis deskriptif dengan menghitung nilai rata-rata setiap variabel pada tiap *cluster*. Nilai rata-rata ini diperoleh dari hasil pengelompokan FGK dan disajikan dalam bentuk Tabel 4.34.

Tabel 4.34 Nilai Rata-rata Variabel Tiap *Cluster*

<i>Cluster</i>	Variabel						
	X₁	X₂	X₃	X₄	X₅	X₆	X₇
1	0,5374	0,4904	0,5036	0,5179	0,2380	0,5502	0,2608
2	0,5198	0,5078	0,5264	0,5487	0,4674	0,4534	1
3	0,4740	0,5095	0,5341	0,5404	0,8864	0,5025	0,2386

Tabel tersebut menunjukkan perbedaan nilai rata-rata setiap variabel pada masing-masing *cluster*, yang selanjutnya digunakan sebagai dasar dalam menginterpretasikan karakteristik dan tingkat risiko depresi tiap *cluster*.

Cluster 1:

Cluster 1 memiliki nilai rata-rata variabel yang umumnya berada pada tingkat menengah. Nilai stres finansial (X_5) relatif rendah dibandingkan *cluster* lainnya, yang menunjukkan tekanan finansial yang tidak terlalu tinggi. Namun, nilai pola makan (X_7) juga relatif rendah, yang mengindikasikan pola makan yang kurang baik. Sementara itu, tekanan akademik (X_2), kepuasan belajar (X_3), jam belajar (X_4), dan durasi tidur (X_6) berada pada tingkat sedang. Variabel umur (X_1) juga tidak menunjukkan perbedaan yang mencolok dibandingkan *cluster* lain. Secara keseluruhan, karakteristik ini menunjukkan kondisi pelajar yang relatif stabil dengan risiko yang tidak terlalu tinggi, sehingga dapat dikategorikan sebagai pelajar dengan risiko depresi sedang.

Cluster 2:

Cluster 2 menunjukkan karakteristik yang cenderung lebih baik dibandingkan *cluster* lainnya. Hal ini terlihat dari nilai kepuasan belajar (X_3) yang cukup tinggi serta nilai jam belajar (X_4) yang paling tinggi dibandingkan *cluster* lain. Selain itu, pola makan (X_7) memiliki nilai paling tinggi, yang mengindikasikan pola hidup yang lebih baik. Meskipun demikian, nilai stres finansial (X_5) dan tekanan akademik (X_2) berada pada tingkat sedang. Durasi tidur (X_6) relatif lebih rendah dibandingkan *cluster* lain, namun masih dalam batas wajar. Secara umum, keseimbangan antara faktor protektif dan faktor risiko menunjukkan bahwa *cluster* ini merepresentasikan pelajar dengan kondisi yang relatif baik, sehingga dikategorikan sebagai pelajar dengan risiko depresi rendah.

Cluster 3:

Cluster 3 menunjukkan karakteristik yang kurang menguntungkan

dibandingkan *cluster* lainnya. Hal ini ditunjukkan oleh nilai stres finansial (X_5) yang sangat tinggi, yang menjadi faktor risiko utama pada *cluster* ini. Selain itu, tekanan akademik (X_2) juga berada pada tingkat yang relatif tinggi. Meskipun kepuasan belajar (X_3) memiliki nilai yang cukup tinggi, kondisi tersebut belum mampu mengimbangi tingginya tekanan yang dialami. Nilai pola makan (X_7) relatif rendah, yang menunjukkan pola hidup yang kurang baik. Variabel lain seperti jam belajar (X_4), durasi tidur (X_6), dan umur (X_1) berada pada tingkat sedang dan tidak menjadi pembeda utama. Secara keseluruhan, dominasi faktor risiko menunjukkan bahwa *cluster* ini dikategorikan sebagai pelajar dengan risiko depresi tinggi.

4.6 Kajian Integrasi Al-Qur'an terhadap Risiko Depresi Pelajar

Depresi pada pelajar tidak muncul secara tiba-tiba, melainkan berkaitan dengan berbagai tekanan yang menyertai kehidupan akademik dan sosial. Tuntutan belajar yang tinggi, keterbatasan kondisi ekonomi, rendahnya kepuasan dalam proses pembelajaran, serta pola hidup yang kurang seimbang dapat menjadi beban psikologis yang dirasakan berat oleh individu. Ketika tekanan-tekanan tersebut terakumulasi dan berlangsung dalam jangka waktu tertentu, kondisi ini berpotensi mengganggu keseimbangan mental pelajar.

Dalam perspektif Islam, kondisi terbebani dan menghadapi kesulitan merupakan bagian dari realitas kehidupan manusia. Al-Qur'an menegaskan hal tersebut dalam firman Allah SWT pada Q.S Al-Baqarah ayat 286:

لَا يُكَلِّفُ اللَّهُ نَفْسًا إِلَّا وُسْعَهَا ۚ لَهَا مَا كَسَبَتْ وَعَلَيْهَا مَا اكْتَسَبَتْ ۗ رَبَّنَا لَا تُؤَاخِذْنَا إِن نَّسِينَا أَوْ أَخْطَأْنَا ۗ رَبَّنَا وَلَا تَحْمِلْ عَلَيْنَا إَصْرًا كَمَا حَمَلْتَهُ عَلَى الَّذِينَ مِن قَبْلِنَا ۗ رَبَّنَا وَلَا تُحَمِّلْنَا مَا لَا طَاقَةَ لَنَا بِهِ ۖ وَاعْفُ عَنَّا وَارْحَمْنَا ۚ أَنْتَ مَوْلَانَا فَانصُرْنَا عَلَى الْقَوْمِ الْكَافِرِينَ ﴿٢٨٦﴾

Artinya: “Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya. Dia mendapat (pahala) dari kebajikan yang dikerjakannya dan

dia mendapat (siksa) dari (kejahatan) yang diperbuatnya. (Mereka berdoa), 'Ya Tuhan kami, janganlah Engkau hukum kami jika kami lupa atau kami melakukan kesalahan. Ya Tuhan kami, janganlah Engkau bebani kami dengan beban yang berat sebagaimana Engkau bebankan kepada orang-orang sebelum kami. Ya Tuhan kami, janganlah Engkau pikulkan kepada kami apa yang tidak sanggup kami memikulnya. Maafkanlah kami, ampunilah kami, dan rahmatilah kami. Engkaulah pelindung kami, maka tolonglah kami menghadapi orang-orang kafir'."

Menurut Tafsir Al-Misbah, ayat ini merupakan penegasan bahwa seluruh beban dan tanggung jawab yang diberikan kepada manusia selalu berada dalam batas kemampuan dan kelapangan manusia itu sendiri. Beban yang dihadapi manusia dapat berada pada tiga kondisi, yaitu mudah dilaksanakan, sulit dilaksanakan, atau mampu dilaksanakan tetapi dirasakan berat. Namun demikian, Allah tidak menghendaki kesulitan bagi hamba-Nya. Ketika suatu beban dirasakan berat karena kondisi tertentu, maka selalu tersedia bentuk kelonggaran dan kemudahan yang disesuaikan dengan keadaan manusia tersebut (Shihab, 2002).

Makna kesanggupan dalam ayat ini juga dipahami sebagai kelapangan, yaitu kemudahan dalam menjalani tugas dan tanggung jawab. Oleh karena itu, kesulitan yang dirasakan manusia bukanlah bentuk pembebanan di luar kemampuan, melainkan bagian dari dinamika kehidupan yang tetap berada dalam koridor kemampuan individu. Perbedaan kondisi, situasi, dan kapasitas pribadi menyebabkan beban yang sama dapat dirasakan berbeda oleh setiap orang.

Keterkaitan makna ayat ini dengan kajian yang dilakukan memberikan pemahaman bahwa kondisi depresi pada pelajar dapat dipandang sebagai refleksi dari akumulasi beban yang dirasakan berat oleh individu, meskipun beban tersebut pada dasarnya masih berada dalam batas kemampuan manusia. Oleh karena itu, perbedaan kondisi psikologis pelajar perlu disikapi secara bijaksana, tanpa memberikan penilaian yang bersifat menghakimi, serta mendorong adanya

perhatian, dukungan, dan pendampingan yang sesuai dengan kemampuan dan kebutuhan masing-masing individu.

Beban yang dirasakan pelajar tidak hanya berasal dari tuntutan akademik, tetapi juga dari kondisi sosial dan ekonomi yang melingkupi kehidupan mereka. Keterbatasan ekonomi dan biaya pendidikan dapat menimbulkan kecemasan serta rasa tidak aman terhadap masa depan. Tekanan ini diperkuat oleh orientasi berlebihan pada prestasi, perbandingan sosial, dan standar keberhasilan yang bersifat material. Ketika pencapaian akademik dijadikan satu-satunya tolok ukur nilai diri, kegagalan dapat memicu perasaan tidak berharga, putus asa, dan stres berkepanjangan yang meningkatkan risiko depresi. Dengan demikian, depresi pada pelajar merupakan fenomena multifaktor yang dipengaruhi oleh tekanan akademik, ekonomi, dan konstruksi sosial mengenai keberhasilan.

Dalam perspektif Islam, penanggulangan tekanan psikologis tersebut dilakukan melalui penguatan nilai-nilai spiritual sebagai strategi koping (Irawan dkk, 2025). Nilai tawakal membantu pelajar mereduksi kecemasan dengan menyadari bahwa hasil usaha, baik dalam aspek akademik maupun ekonomi, berada dalam ketetapan Allah SWT. Sabar menumbuhkan keteguhan dalam menghadapi kesulitan tanpa kehilangan harapan, sedangkan syukur membantu individu memaknai kondisi hidup secara lebih seimbang dan tidak semata-mata bergantung pada capaian material. Praktik ibadah seperti shalat dan dzikir berperan dalam menenangkan emosi dan menjaga stabilitas psikologis, sementara dukungan sosial berbasis ukhuwah membantu mengurangi rasa terisolasi. Nilai-nilai tersebut berkontribusi dalam menjaga keseimbangan mental pelajar serta mendukung upaya pengelolaan risiko depresi.

Selain nilai tawakal, sabar, dan syukur, konsep khauf dan raja' sebagaimana dijelaskan oleh Al-Ghazali juga memiliki relevansi dalam pengelolaan kondisi psikologis pelajar. Al-khauf merupakan kondisi batin yang lahir dari pengetahuan ('ilm) tentang keagungan Allah SWT serta kesadaran akan konsekuensi dari setiap perbuatan manusia. Rasa takut ini berfungsi mendorong pengendalian diri dan perbaikan perilaku melalui penahanan hawa nafsu, menjauhi perilaku negatif, serta mengarahkan individu pada introspeksi, taubat, dan ketaatan. Dalam konteks depresi, khauf membantu individu mengurangi keterikatan berlebihan pada aspek duniawi, memperkuat kedekatan dengan Allah SWT, dan menumbuhkan ketenangan batin, sehingga mendukung keseimbangan psikologis dan mencegah munculnya keputusasaan (Dacholfany, 2014).

Sebagai pasangan dari khauf, Al-Ghazali menjelaskan konsep al-raja', yaitu sikap berharap kepada Allah SWT yang lahir dari pengetahuan dan diwujudkan melalui amal nyata. Raja' tidak dimaknai sebagai angan-angan semata, melainkan harapan yang disertai dengan usaha untuk memperkuat iman, meningkatkan ketaatan, serta menjauhi perilaku tercela. Dalam kaitannya dengan risiko depresi, raja' berperan menumbuhkan sikap optimis, mencegah rasa putus asa, dan mendorong individu untuk terus memperbaiki diri serta berharap pada rahmat dan ampunan Allah SWT. Keseimbangan antara khauf dan raja' tersebut menjadi landasan spiritual yang mendukung ketahanan psikologis pelajar dalam menghadapi tekanan akademik, sosial, dan ekonomi (Dacholfany, 2014).

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil optimalisasi variasi nilai *weighting exponent* pada algoritma *Fuzzy Gustafson-Kessel* (FGK), diperoleh bahwa parameter tersebut berpengaruh terhadap kualitas *clustering*, di mana nilai m yang lebih kecil cenderung menghasilkan kualitas *cluster* yang lebih baik, sedangkan peningkatan nilai m menyebabkan kualitas *clustering* menurun akibat meningkatnya tingkat kekaburan derajat keanggotaan. Berdasarkan hasil tersebut, parameter optimal diperoleh pada kombinasi $c = 3$ dan $m = 1,1$ yang menghasilkan kualitas *clustering* optimal. Hasil *clustering* dengan parameter optimal tersebut mampu mengelompokkan data pelajar ke dalam kategori risiko depresi rendah, sedang, dan tinggi, sehingga menunjukkan bahwa optimalisasi parameter pada algoritma FGK dapat menghasilkan pengelompokan data yang lebih representatif dalam analisis risiko depresi pelajar.

5.2 Saran untuk Penelitian Lanjutan

Berdasarkan hasil dan keterbatasan penelitian ini, penelitian selanjutnya disarankan untuk mengkaji sensitivitas algoritma *Fuzzy Gustafson-Kessel* (FGK) terhadap variasi nilai *threshold* (ϵ) yang digunakan sebagai kriteria konvergensi. Variasi nilai *threshold* diharapkan dapat memberikan gambaran mengenai pengaruh kriteria penghentian iterasi terhadap stabilitas hasil *clustering*, jumlah iterasi yang diperlukan, serta kualitas *cluster* yang dihasilkan.

Selain itu, penelitian lanjutan dapat mempertimbangkan penggunaan metode evaluasi *cluster* tambahan, seperti *Partition Coefficient* (PC), *Partition Entropy*

(PE), maupun indeks validitas lain yang relatif jarang digunakan pada implementasi *Fuzzy Gustafson-Kessel*, seperti indeks *Fukuyama-Sugeno*. Penggunaan beberapa indeks evaluasi diharapkan dapat memberikan sudut pandang yang lebih beragam dalam menilai kualitas struktur *cluster* serta memperkuat interpretasi hasil *clustering* yang diperoleh.

DAFTAR PUSTAKA

- Amrullah, A. M. K. (n.d.). *Tafsir Al-Azhar Jilid 8*.
- Ayuningtias, I., Debataraja, N. N., & Imro'ah, N. (2019). Analisis Cluster Non-Hirarki dengan Metode K-Modes. *Buletin Ilmiah Mat. Stat. Dan Terapannya (Bimaster)*, 08(4), 909–916.
- Bezdek, J. C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press.
- Cebeci, Z., & Yildiz, F. (2015). Comparison of K-Means and Fuzzy C-Means Algorithms on Different Cluster Structures. *Journal of Agricultural Informatics*, 6(3).
- Dacholfany, M. I. (2014). Al-Khaud dan Al-Raja' Menurut Al-Ghazali. *As-Salam*, 5(1).
- Dillasamola, D. (2021). *Buku Ajar Teknik Evaluasi Bioaktivitas*. LPPM - Universitas Andalas.
- Duan, Y., Liu, Q., & Xia, S. (2018). An improved initialization center k-means clustering algorithm based on distance and density. *AIP Conference Proceedings*, 1955.
- Fajri, M., Rais, & Handayani, L. (2023). Regions Grouping in Central Sulawesi Province by Transmitted Disease using Fuzzy Gustafson Kessel. *Barekeng*, 17(1), 275–284.
- Feng, Z., & Fan, J. C. (2016). A novel validity index in fuzzy clustering algorithm. *International Journal of Wireless and Mobile Computing*, 10(2), 183–190.
- Firdawanti, A. R., Ahmad, H. F., & Agustiani, N. (2025). Analisis Perbandingan Kinerja Algoritma K-Means dan K-Medoids dengan Reduksi Dimensi PCA pada Indikator Kesehatan dan Sosial. *Bulletin of Computer Science Research*, 5(5), 1235–1247.
- Gustafson, D. E., & Kessel, W. C. (1978). Fuzzy Clustering with a Fuzzy Covariance Matrix. *Proceedings of the IEEE Conference on Decision and Control*, 761–766.
- Han, H. (2023). Fuzzy clustering algorithm for university Students' Psychological Fitness and Performance Detection. *Heliyon*, 9(8).
- Hapsari, R. K., Sumarni, T., Hidayatulloh, T., Farida, Herianto, Elfaladonna, F., Aini, N., Nurfitria, Kurniawan, H., Kraugusteeliana, Yusuf, L., & Widiatama, Y. (2024). *Data Mining*. Indie Press. www.indiepress.id
- Hasanah, M. (2024). Konsep Depresi Perspektif Al-Qur'an: Analisis Ayat-ayat Psikologi. *Al Bustan: Jurnal Studi Islam Dan Sosial Keagamaan*, 1(1), 71–85.
- Herdian, C., Kamila, A., & Agung Musa Budidarma, I. G. (2024). Studi Kasus Feature Engineering Untuk Data Teks: Perbandingan Label Encoding dan

- One-Hot Encoding Pada Metode Linear Regresi. *Technologia : Jurnal Ilmiah*, 15(1), 93.
- Hermansyah, ud, Aditia Hamdan, R., Sidik, F., & Wibowo, A. (2020). Klasterisasi Data Travel Umroh di Marketplace Umroh.com Menggunakan Metode K-Means. *Jurnal Ilmu Komputer*, XIII(2).
- Herzog, R., Köhne, F., Kreis, L., & Schiela, A. (2023). *Frobenius-Type Norms and Inner Products of Matrices and Linear Maps with Applications to Neural Network Training*.
- Hidayat, T. (2022). *Analisis Cluster dengan Metode Ensemble ROCK dalam Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan berdasarkan Indikator Kesejahteraan Rakyat*. Universitas Negeri Makassar.
- Irawan, J., Rusmandi, H., Shabet, A., Warsah, I., & Rizal, S. (2025). Analisis Psikologi Islam dalam Mengatasi Stres da Kecemasan pada Remaja. *Jurnal Literasiologi*, 14(3).
- Jatmiko, N. A. (2023). *Depression Student Dataset*. <https://www.kaggle.com/datasets/ikynahidwin/depression-student-dataset>
- Kementerian Agama. (2022). *Al-Qur'an dan Terjemahannya. Lajnah Pentashihan Mushaf Al-Qur'an Badan Litbang Kementerian Agama RI*. 2022.
- Kiwasi Wororomi, J., Reba, F., Aleksander Mandowen, S., Sroyer, A. M., Manurung, halomoan E., Koibur, M. E., Pujiarini, E. H., Edy, M. R., Yuliana, H., Yusuf, M., & Dinata, R. M. (2024). *Data Mining (Memahami Pola di Balik Angka)*. Eureka Media Aksara.
- Muhima, R. R., Kurniawan, M., Wardhana, S. R., Yudhana, A., Sunardi, S., Rahmawati, W. M., & Yulastuti, G. E. (2021). *Kupas Tuntas Algoritma Clustering: Konsep, Perhitungan Manual, dan Program* (E. Risanto, Ed.). ANDI.
- Nurzida, A., Utami, I. T., & Rochayani, M. Y. (2024). Perbandingan Metode Fuzzy C-Means dan Gustafson-Kessel dalam Penentuan Cluster Tingkat Risiko Penularan Tuberculosis terhadap Penyakit di Jawa Timur. *Jurnal Gaussian*, 13(2), 373–382.
- Pal, N. R., & Bezdek, J. C. (1995). On Cluster Validity for the Fuzzy c-Means Model. *IEEE Transactions on Fuzzy Systems*, 3(3), 370–379.
- Paradilla, Y. S., Hayati, M. N., & Sifriyani. (2023). Implementation of The Fuzzy Gustafson-Kessel Method on Grouping Districts/Cities in Kalimantan Island Based on Poverty Issues Factors. *Barekeng*, 17(1), 125–134.
- Prastanika, W. W., & Wijayanto, A. W. (2023). Analisis Hard dan Soft Clustering Untuk Pengelompokan Indikator Ketahanan Pangan Indonesia 2021. *Jurnal Sistem Dan Teknologi Informasi (JustIN)*, 11(4), 596.
- Putri, F. A. A., & Rochmawati, N. (2021). Penerapan Algoritma Gustafson-Kessel untuk Clustering Tweets Mention Akun Go-jek dan Grab Indonesia. *Journal*

of Informatics and Computer Science (JINACS), 2, 204–212.

- Rahakbauw, D. L., Ilwaru, V. Y. I., & Hahury, M. H. (2017). Implementasi Fuzzy C-Means Clustering dalam Penentuan Beasiswa. *Barekeng Jurnal Ilmu Matematika Dan Terapan*, 11(1), 1–11.
- Rahayu, P. W., Sudipa, I. G. I., Suryani, S., Surachman, A., Ridwan, A., Darmawiguna, I. G. M., Sutoyo, Muh. N., Slamet, I., Harlina, S., & Maysanjaya, I. M. D. (2024). *Buku Ajar Data Mining* (E. Efitra, Ed.). PT. Sonpedia Publishing Indonesia.
- Ramadani, I. R., Rozzaq, B. K., & Fauziyah, T. (2024). Depresi, Penyebab Dan Gejala Depresi. *BERSATU: Jurnal Pendidikan Bhinneka Tunggal Ika*, 2(2), 88–99.
- Ratu, B., Elfira, N., Lestari Sirumpa, Y., Andini, F., & Sateki, K. (2025). Depresi di Kalangan Mahasiswa Depression Among College Students. *Jurnal Ilmu Pendidikan Ahlusunnah*, VIII.
- Rindengan, A. J., & Langi, Y. A. R. (2019). *Sistem Fuzzy*. Patra Media Grafindo.
- Rosita, D., Elmunsyah, H., & H, A. N. (2021). *Modul Ajar Fuzzy*. Ahlimedia Press.
- Saputri, I. A., & Nurrahima, A. (2020). Depresi Anak Usia Sekolah. *Journal of Holistic Nursing and Health Science*, 3(2), 50–58.
- Shihab, M. Quraish. (2002). *Tafsir al-Mishbah : Pesan, Kesan dan Keserasian Al-Qur'an*. Lentera Hati.
- Suhaeni, C., Kurnia, A., & Ristiyanti. (2018). Perbandingan Hasil Pengelompokan menggunakan Analisis Cluster Berhirarki, K-Means Cluster, dan Cluster Ensemble (Studi Kasus Data Indikator Pelayanan Kesehatan Ibu Hamil). *Jurnal Media Infotama*, 14(1).
- Syahreni, Almais, A. T. W., Nugroho, F., Crysdian, C., Harini, S., Zarkoni, A., Ma'muri, & Sugiharto, T. I. (2025). Analisis Pola Data Klimatologi Potensial dengan Principal Component Analysis. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(6).
- Talakua, M. W., Leleury, Z. A., & Talluta, A. W. (2017). Analisis Cluster dengan Menggunakan Metode K-Means untuk Pengelompokkan Kabupaten/Kota di Provinsi Maluku Berdasarkan Indeks Pembangunan Tahun 2024. *Barekeng: Jurnal Ilmu Matematika Dan Terapan*, 11(2), 119–128.
- Telaumbanua, S. A. B., Setiadi, F., & Nurjanah, S. (2025). Analisis Clustering Menggunakan Metode Enhanced Fuzzy C-Means Clustering Dengan Algoritma Rock Pada Student Performance Dataset. *Bit-Tech*, 7(3), 984–994.
- Thamrin, N., & Wijayanto, A. W. (2021). Comparison of Soft and Hard Clustering: A Case Study on Welfare Level in Cities on Java Island. *Indonesian Journal of Statistics and Its Applications*, 5(1), 141–160.

- Ulfah, M., & Irtwaty, A. (2022). Penerapan Data Mining Clustering Menggunakan Metode K-Means Dalam Pengelompokan Buku Perpustakaan Politeknik Negeri Balikpapan. *Fidelity : Jurnal Teknik Elektro*, 3, 62–68.
- Wagner, F. H., Dalagnol, R., Sánchez, A. H., Hirye, M. C. M., Favrichon, S., Lee, J. H., Mauceri, S., Yang, Y., Saatchi, S., Ferreira, M. P., Carvalho, M., De Paulo, M., & Kattenborn, T. (2022). K-textures, a self-supervised hard clustering deep learning algorithm for satellite image segmentation. *Frontiers in Environmental Science*.
- World Health Organization. (2025). *Depressive disorder (depression)*.
- Yang, M.-S. (1993). A Survey of Fuzzy Clustering. *Mathl. Comput. Modelling*, 18(11), 1–16.
- Zadeh, L. A. (1965). Fuzzy Sets. *Information and Control*, 8, 338–353.

LAMPIRAN

Lampiran 1 *Dataset* Penelitian

<https://docs.google.com/spreadsheets/d/1mUhBNo-1mGVMMYnYB2XjkXpb9raIW0eF/edit?usp=sharing&ouid=109194483551456367490&rtpof=true&sd=true>



Lampiran 2 *Dataset* setelah *Preprocessing*

https://docs.google.com/spreadsheets/d/1bJ7PZWzNgqrbLDhI0JuhHttGzwsGRTT9/edit?usp=drive_link&ouid=109194483551456367490&rtpof=true&sd=true



Lampiran 3 Perhitungan Hasil *Clustering* dengan $m = 1,1$

https://drive.google.com/drive/folders/1OpwmJ-mYuvs19eP5x0xhkKXWYh7r3Xoy?usp=drive_link



Lampiran 4 Perhitungan Hasil *Clustering* dengan $m = 1,3$

https://drive.google.com/drive/folders/1a-CguZ4EEEnFBzRrIhKs-YqvLtcecLg_e?usp=drive_link



Lampiran 5 Perhitungan Hasil *Clustering* dengan $m = 1,5$

https://drive.google.com/drive/folders/1Tj7Fe-TgCKTRo9i0UgPToZNhZWlnbEyp?usp=drive_link



Lampiran 6 Perhitungan Hasil *Clustering* dengan $m = 1,7$

https://drive.google.com/drive/folders/1pa_W3Fagwxdp4RviUeJ7qcKmKpwUH9fZ?usp=drive_link



Lampiran 7 Perhitungan Hasil *Clustering* dengan $m = 1,9$

https://drive.google.com/drive/folders/1kDPCfygCXo5CAaXQs6_Kd5tXpGCx_hNi?usp=drive_link



Lampiran 8 Perhitungan Hasil Clustering dengan $m = 2,1$

https://drive.google.com/drive/folders/1zVIUzbH1TsMaCXO3tnOhAZ9gDEQ7fnFb?usp=drive_link

**Lampiran 9** *Source Code* Algoritma Fuzzy Gustafson-Kessel

```
import numpy as np
import pandas as pd

def fgk_clustering(X, c, m, max_iter, epsilon, random_state, save_path):
    generator = np.random.default_rng(random_state)
    N, h = X.shape

    # Inisialisasi matriks partisi U (acak, normalisasi per baris)
    U = generator.random((N, c))
    U = U / U.sum(axis=1, keepdims=True)

    # Simpan matriks partisi awal
    with pd.ExcelWriter(save_path, engine="openpyxl") as writer:
        pd.DataFrame(U, columns=[f'Cluster {i+1}' for i in
range(c)]).to_excel(
            writer, sheet_name="Matriks_Partisi_Awal", index=False
        )

    for t in range(1, max_iter + 1):
        U_m = U ** m

        # Menghitung pusat cluster
        V = (U_m.T @ X) / np.sum(U_m.T, axis=1)[:, None]

        F_list, A_list = [], []
        for i in range(c):

            # Menghitung matriks kovarian
            diff = X - V[i]
            pembilang = (diff.T * U_m[:, i]) @ diff
            penyebut = np.sum(U_m[:, i])
            F = pembilang / penyebut
            F_list.append(F)

            # Menghitung matriks A_k (norm-inducing matrix)
            det_F = np.linalg.det(F)
            if det_F <= 1e-6:
                F += np.eye(h) * 1e-6
                det_F = np.linalg.det(F)
            A = (det_F ** (1.0 / h)) * np.linalg.inv(F)
```

```

A_list.append(A)

# Mengitung jarak D
D = np.zeros((N, c))
for i in range(c):
    diff = X - V[i]
    D[:, i] = np.sqrt(np.einsum('ij,jk,ik->i', diff, A_list[i], diff) + 1e-10)

# Update matriks partisi U
eps = 1e-10
U_new = np.zeros_like(U)
for i in range(N):
    for k in range(c):
        penyebut = np.sum(((D[i, k] + eps) / (D[i, :] + eps)) ** (2 / (m -
1))))
        U_new[i, k] = 1.0 / penyebut

# Fungsi objektif
J = np.sum((U ** m) * (D ** 2))
delta = np.linalg.norm(U_new - U)

# Menyimpan sheet per iterasi
with pd.ExcelWriter(save_path, mode='a', engine="openpyxl",
if_sheet_exists='replace') as writer:
    # Centroid
    pd.DataFrame(V, columns=[f'Var_{i+1}' for i in
range(h)]).to_excel(writer, sheet_name=f'Iter_{t}_Pusat_Cluster",
index=False)

    # Matriks Kovarian
    gabungan_kovarian = []
    for i, F in enumerate(F_list):
        label = pd.DataFrame([[f'Matriks Kovarian Klaster {i+1}"]]])
        F_df = pd.DataFrame(F)
        gabungan_kovarian.append(pd.concat([label, F_df],
ignore_index=True))
    pd.concat(gabungan_kovarian,
ignore_index=True).to_excel(writer,
sheet_name=f'Iter_{t}_Matriks_Kovarian", index=False,
header=False)

    # Matriks A
    gabungan_matriksA = []
    for i, A in enumerate(A_list):
        label = pd.DataFrame([[f"Matriks A Klaster {i+1}"]]])
        A_df = pd.DataFrame(A)
        gabungan_matriksA.append(pd.concat([label, A_df],
ignore_index=True))

```

```

        pd.concat(gabungan_matriksA,
ignore_index=True).to_excel(writer,
        sheet_name=f"Iter_{t}_Matriks_A", index=False, header=False)

# Jarak
pd.DataFrame(D, columns=[f'Cluster {i+1}' for i in
range(c)]).to_excel(writer, sheet_name=f"Iter_{t}_Jarak",
index=False)

# Fungsi objektif
pd.DataFrame([[J]], columns=["Objective"]).to_excel(writer,
sheet_name=f"Iter_{t}_Fungsi_Objektif", index=False)

# Matriks partisi
pd.DataFrame(U_new, columns=[f'Cluster {i+1}' for i in
range(c)]).to_excel(writer,
sheet_name=f"Iter_{t}_Matriks_Partisi", index=False)

# Delta
pd.DataFrame([[delta]], columns=["Delta"]).to_excel(writer,
sheet_name=f"Iter_{t}_Delta_U", index=False)

if delta < epsilon:
    print(f'Konvergen pada iterasi {t}')
    U = U_new
    break

U = U_new

# Menyimpan hasil akhir
labels = np.argmax(U, axis=1) + 1
hasil_cluster = pd.DataFrame(X, columns=[f'Fitur_{i+1}' for i in
range(h)])
hasil_cluster['Cluster'] = labels

with pd.ExcelWriter(save_path, mode='a', engine="openpyxl",
if_sheet_exists='replace') as writer:
    hasil_cluster.to_excel(writer, index=False,
sheet_name="Hasil_Clustering")

print(f'Hasil FGK tersimpan di {save_path}')

return V, U, labels, t

# Penerapan
path_data = "/content/Dataset_Olah.xlsx"
df = pd.read_excel(path_data)
X = df.values

```

```

V, U, labels, iter_last = fgk_clustering(
    X,
    c=2,
    m=1.3,
    max_iter=1000,
    epsilon=1e-4,
    random_state=42,
    save_path="/content/drive/MyDrive/contohhasil_cluster_2-1,3.xlsx"
)

print("Iterasi terakhir:", iter_last)

```

Lampiran 10 *Source Code Xie-Beni Index*

```

import numpy as np
import pandas as pd

# Parameter
file_input = "/content/Dataset_evaluasiindex_m=1.3 - Copy.xlsx"
file_output = "hasil_XB_m=1,3.xlsx"
m = 1.3

xls = pd.ExcelFile(file_input)
sheet_names = xls.sheet_names

# Writer Excel
writer = pd.ExcelWriter(file_output, engine="openpyxl")

# Loop Semua Sheet
for sheet in sheet_names:

    df = pd.read_excel(file_input, sheet_name=sheet)

    # Data asli
    X = df.iloc[:, 0:7].values
    n = X.shape[0]

    # Tentukan jumlah cluster
    total_col = df.shape[1]
    c = total_col - 14

    # Matriks U
    U = df.iloc[:, 7:7+c].values

    # Pusat cluster (c baris, 7 kolom)
    V = df.iloc[0:c, 7+c:7+c+7].values

```

```

# Hitung S
S = 0.0
for k in range(c):
    for i in range(n):
        dist2 = np.linalg.norm(X[i] - V[k])**2
        S += (U[i, k]**m) * dist2
    S = S / n

# Hitung d_min
d_min = np.inf
for i in range(c):
    for k in range(c):
        if i != k:
            dist_vk = np.linalg.norm(V[i] - V[k])**2
            d_min = min(d_min, dist_vk)

# Hitung XB
XB = S / d_min

# Simpan Ke Sheet Baru
hasil_df = pd.DataFrame({
    "Komponen": ["Jumlah Cluster (c)", "Weighting Exponent (m)",
"Nilai S", "Nilai d_min", "Nilai XB"],
    "Nilai": [c, m, S, d_min, XB]
})

hasil_df.to_excel(writer, sheet_name=f"{sheet}_XB", index=False)

# Simpan file
writer.close()

print("Hasil Xie-Beni disimpan ke:", file_output)

```

Lampiran 11 *Source Code* Visualisasi Hasil *Clustering* dengan *Principal Component Analysis* (PCA)

```

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.decomposition import PCA
from scipy.spatial import ConvexHull
from IPython.display import Image, display

file_path = "/content/hasil fgk optimal.xlsx"

X = pd.read_excel(file_path, sheet_name="dataset").values
U = pd.read_excel(file_path, sheet_name="matriks_u").values

# Label Cluster
labels = np.argmax(U, axis=1)

```

```

# PCA
pca = PCA(n_components=2)
X_pca = pca.fit_transform(X)

# Plot
plt.figure(figsize=(8,6))

colors = ['red', 'green', 'blue', 'purple']
markers = ['o', '^', 's', '+']

for k in np.unique(labels):
    points = X_pca[labels == k]

    plt.scatter(points[:,0], points[:,1],
                color=colors[k],
                marker=markers[k],
                label=f'Cluster {k+1}',
                edgecolors='black')

    if len(points) >= 3:
        hull = ConvexHull(points)
        hull_pts = points[hull.vertices]
        plt.fill(hull_pts[:,0], hull_pts[:,1],
                color=colors[k], alpha=0.3)

plt.xlabel("PC1")
plt.ylabel("PC2")
plt.title("Visualisasi Hasil Clustering FGK (c=3, m=1,3)")
plt.legend()
plt.grid(True)
plt.tight_layout()

# Simpan & Tampilkan
plt.savefig("cluster_fgk.png", dpi=300)
plt.close()

display(Image("cluster_fgk.png"))

```

RIWAYAT HIDUP



Amira Adelia Putri, lahir di Malang pada 24 Mei 2004. Penulis merupakan anak kedua dari dua bersaudara. Penulis menempuh pendidikan dimulai dari RA Munir Ismail dan lulus pada tahun 2010. Selanjutnya, penulis melanjutkan pendidikan sekolah dasar di SDN Sepanjang 5 dan lulus pada tahun 2016. Pendidikan kemudian dilanjutkan ke SMPN 1 Gondanglegi dan diselesaikan pada tahun 2019. Setelah itu, penulis melanjutkan ke jenjang sekolah menengah atas di MAN 1 Malang dan lulus pada tahun 2022. Pada tahun yang sama, penulis melanjutkan pendidikan di Universitas Islam Negeri Maulana Malik Ibrahim Malang pada Program Studi Matematika, Fakultas Sains dan Teknologi. Selama menempuh pendidikan di perguruan tinggi, penulis telah melaksanakan Praktik Kerja Lapangan (PKL) pada bidang logistik di PT PLN (Persero) UID Jawa Timur UP3 Malang. Selain itu, penulis juga pernah mengikuti magang mandiri pada bidang pengelolaan informasi dan administrasi kependudukan di Dinas Kependudukan dan Pencatatan Sipil Kabupaten Malang sebagai bentuk pengembangan pengalaman praktik kerja.



BUKTI KONSULTASI SKRIPSI

Nama : Amira Adelia Putri
NIM : 220601110051
Fakultas / Jurusan : Sains dan Teknologi / Matematika
Judul Skripsi : Optimalisasi Algoritma *Fuzzy Gustafson-Kessel* dengan Variasi Parameter *Weighting Exponent* untuk *Clustering* Data Risiko Depresi Pelajar
Dosen Pembimbing I : Intan Nisfulaila, M.Si
Dosen Pembimbing II : Erna Herawati, M.Pd

No	Tanggal	Hal	Tanda Tangan
1.	27 Agustus 2025	Konsultasi Topik dan Data	1.
2.	12 September 2025	Konsultasi Topik dan Data	2.
3.	23 September 2025	Konsultasi Bab I, II, dan III	3.
4.	29 September 2025	Konsultasi Bab I, II, dan III	4.
5.	07 Oktober 2025	Konsultasi Bab I, II, dan III	5.
6.	09 Oktober 2025	Konsultasi Kajian Agama Bab I dan II	6.
7.	13 Oktober 2025	Konsultasi Bab I, II, dan III	7.
8.	14 Oktober 2025	ACC Bab I, II, dan III	8.
9.	16 Oktober 2025	Konsultasi Kajian Agama Bab I dan II	9.
10.	16 Oktober 2025	ACC Kajian Agama Bab I dan II	10.
11.	04 November 2025	ACC Seminar Proposal	11.
12.	20 November 2025	Konsultasi Revisi Seminar Proposal	12.
13.	24 Desember 2025	Konsultasi Bab IV dan V	13.
14.	14 Januari 2026	ACC Bab IV dan V	14.
15.	19 Januari 2026	Konsultasi Kajian Agama Bab IV	15.
16.	21 Januari 2026	Konsultasi Kajian Agama Bab IV	16.



KEMENTERIAN AGAMA RI
UNIVERSITAS ISLAM NEGERI
MAULANA MALIK IBRAHIM MALANG
FAKULTAS SAINS DAN TEKNOLOGI
Jl. Gajayana No. 50 Malang 65144 Telp. / Fax. (0341) 558933

17.	28 Januari 2026	ACC Kajian Agama Bab IV	17.
18.	20 Februari 2026	ACC Seminar Hasil	18.
19.	13 Maret 2026	Konsultasi Revisi Seminar Hasil	19.
20.	06 April 2026	Konsultasi Revisi Seminar Hasil	20.
21.	14 April 2026	ACC Sidang Skripsi	21.
22.	27 April 2026	ACC Keseluruhan	22.

Malang, 27 April 2026

Mengetahui,

Ketua Program Studi Matematika



Dr. Fachrur Rozi, M.Si

NIP. 19800527 200801 1 012