

**OPTIMASI ALGORITMA K-MEANS
MENGUNAKAN *PARTICLE SWARM OPTIMIZATION***

SKRIPSI

**OLEH:
MUTHIA AFIFI
NIM. 210601110107**



**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM MALANG
2025**

**OPTIMASI ALGORITMA *K-MEANS*
MENGGUNAKAN *PARTICLE SWARM OPTIMIZATION***

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
Muthia Afifi
NIM. 210601110107**

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM MALANG
2025**

OPTIMASI ALGORITMA *K-MEANS* MENGUNAKAN *PARTICLE SWARM OPTIMIZATION*

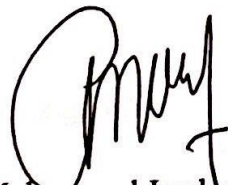
SKRIPSI

**Oleh
Muthia Afifi
NIM. 210601110107**

Telah Disetujui Untuk Diuji

Malang, 20 Mei 2025

Dosen Pembimbing I



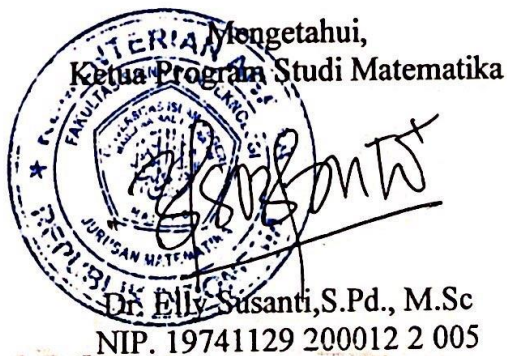
**Dr. Mohammad Jamhuri, M.Si
NIP. 19810502 200501 1 004**

Dosen Pembimbing II



**Erna Herawati, M.Pd
NIPPPK. 19760723 202321 2 006**

**Mengetahui,
Ketua Program Studi Matematika**



**Dr. Elly Susanti, S.Pd., M.Sc
NIP. 19741129 200012 2 005**

OPTIMASI ALGORITMA *K-MEANS* MENGUNAKAN *PARTICLE SWARM OPTIMIZATION*

SKRIPSI

Oleh
Muthia Afifi
NIM. 210601110107

Telah Dipertahankan di Depan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
untuk Memperoleh Gelar Sarjana Matematika (S.Mat)
Tanggal 13 Juni 2025

Ketua Penguji : Dr. Usman Pagalay, M.Si.

Anggota Penguji 1 : Juhari, M.Si.

Anggota Penguji 2 : Dr. Mohammad Jamhuri,
M.Si.

Anggota Penguji 3 : Erna Herawati, M.Pd.

.....
.....
.....
.....

Mengetahui,
Ketua Program Studi Matematika

Dr. Elly Susanti, S.Pd., M.Sc
NIP. 19741129 200012 2 005

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Muthia Afifi

NIM : 210601110107

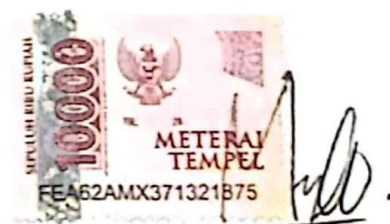
Program Studi : Matematika

Fakultas : Sains dan Teknologi

Judul Skripsi : Optimasi Algoritma *K-Means* Menggunakan *Particle Swarm Optimization*

Menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini merupakan hasil karya saya sendiri, bukan pengambilan tulisan atau pemikiran orang lain yang saya akui sebagai hasil tulisan dan pemikiran saya, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka. Apabila di kemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 13 Juni 2025



Muthia Afifi
NIM.210601110107

MOTO

“Maka, sesungguhnya beserta kesulitan ada kemudahan. Sesungguhnya beserta kesulitan ada kemudahan”

(QS. Al-Insyirah, 94: 5-6)

“Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya”

(QS. Al-Baqarah, 2: 286)

PERSEMBAHAN

Bismillaahirrahmaanirrahiim

Skripsi ini penulis persembahkan kepada seluruh keluarga penulis, terkhusus kepada kedua orang tua penulis, Bapak Mursal Pili dan Ibu Agustina yang selalu mendo'akan, memberi semangat dan memberi dukungan material untuk kelancaran penulis dalam menyelesaikan tugas akhir ini.

KATA PENGANTAR

Assalamu 'alaikum Warahmatullahi Wabarakatuh

Alhamdulillah, segala puji dan syukur senantiasa penulis panjatkan kepada Allah *subhanahu wa ta'ala* atas segala rahmat, taufik, dan hidayah yang telah diberikan, sehingga penulis dapat menyelesaikan Skripsi dengan judul “Optimasi Algoritma *K-Means* Menggunakan *Particle Swarm Optimization*”. Shalawat dan salam senantiasa tercurahkan kepada Nabi Muhammad *shallallahu 'alaihi wa sallam* yang telah membimbing manusia dalam jalan kebenaran, yakni agama Islam dan dinantikan syafaat beliau kelak di akhirat.

Dalam proses penyelesaian skripsi ini, penulis banyak mendapat bimbingan, dukungan, dan arahan serta sumbangsih dari berbagai pihak. Oleh karena itu penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Prof. Dr. H. M. Zainuddin, M.A., selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim.
2. Prof. Dr. Sri Harini, M.Si, selaku dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim.
3. Dr. Elly Susanti, S.Pd., M.Sc, selaku ketua Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim.
4. Dr.Mohammad Jamhuri, M.Si selaku dosen pembimbing I yang telah memberikan berbagai pengetahuan, pengalaman, arahan, nasihat, serta motivasi kepada penulis.
5. Erna Herawati, M.Pd selaku dosen pembimbing II yang telah memberikan ilmu, nasihat, bimbingan, pengalaman, serta motivasi kepada penulis.
6. Seluruh dosen Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim.
7. Bapak Mursal Pili dan Ibu Agustina serta seluruh keluarga yang senantiasa mendoakan, memberikan semangat, dukungan, nasihat, serta kasih sayang sehingga penulis dapat menyelesaikan tugas akhir.
8. Teman-teman *averous generation* di Malang yang selalu memberikan dukungan dan semangat kepada penulis.
9. Seluruh mahasiswa Program Studi Matematika angkatan 2021 yang senantiasa memberikan bantuan, dukungan dan semangat kepada penulis.

10. Pihak-pihak lain yang tidak bisa penulis sebutkan satu persatu.

Semoga Allah *subhanahu wa ta'ala* selalu memberikan balasan atas segala bantuan dan kebaikan yang telah diberikan kepada penulis. Penulis berharap agar laporan ini dapat bermanfaat bagi penulis serta pembaca untuk menambah wawasan keilmuan yang selalu berkembang.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Malang, 13 Juni 2025

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGAJUAN.....	ii
HALAMAN PERSETUJUAN	iii
PERNYATAAN KEASLIAN TULISAN	v
MOTO.....	vi
PERSEMBAHAN.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
ABSTRAK	xiv
ABSTRACT	xv
مستخلص البحث.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	5
1.3 Tujuan Penelitian	5
1.4 Manfaat Penelitian	6
1.5 Batasan Masalah	6
1.6 Definisi Istilah.....	6
BAB II KAJIAN PUSTAKA.....	8
2.1 Teori <i>Clustering</i>	8
2.2 Algoritma <i>K-Means</i>	9
2.3 <i>Particle Swarm Optimization</i> (PSO).....	12
2.4 Data Iris dan Data Wine.....	14
2.5 Kajian Integrasi Topik dengan Al-Qur'an/Hadits.....	16
2.6 Kajian Topik Dengan Teori Pendukung	17
BAB III METODE PENELITIAN	20
3.1 Jenis Penelitian.....	20
3.2 Data dan Sumber Data	20
3.3 Struktur Data.....	21
3.4 Tahapan Penelitian.....	21
3.4.1 <i>K-Means Clustering</i>	21
3.4.2 <i>Particle Swarm Optimization</i> (PSO)	23
3.4.3 Penggabungan <i>K-Means</i> dan PSO.....	26
3.5 Evaluasi dan Validasi Hasil	28
BAB IV HASIL DAN PEMBAHASAN	31
4.1 Deskripsi Dataset Penelitian	31
4.1.1 Dataset Iris.....	31
4.1.2 Dataset Wine.....	31
4.2 Implementasi <i>K-Means Clustering</i>	32
4.2.1 Implementasi <i>K-Means</i> Manual pada Dataset Iris (5 Data Pertama).....	32
4.2.2 <i>K-Means</i> pada Dataset Iris (Menggunakan Seluruh Data)	46

4.2.3 Implementasi <i>K-Means</i> Manual pada Dataset Wine (5 Data Pertama)	48
4.2.4 <i>K-Means</i> pada Dataset Wine (Menggunakan Seluruh Data)	53
4.3 Implementasi PSO	55
4.3.1 Dataset Iris (5 Data Pertama)	55
4.3.2 Dataset Iris (Menggunakan Seluruh Data)	105
4.3.3 Dataset Wine (5 Data Pertama)	105
4.3.4 Dataset Wine (Menggunakan Seluruh Data)	126
4.4 Implementasi <i>K-Means</i> - PSO	126
4.4.1 <i>K-Means</i> - PSO Pada Dataset Iris (5 Data Pertama)	126
4.4.2 <i>K-Means</i> - PSO Pada Dataset Iris (Seluruh Data)	127
4.4.3 <i>K-Means</i> - PSO Pada Dataset Wine (5 Data Pertama)	128
4.4.4 <i>K-Means</i> - PSO Pada Dataset Wine (Seluruh Data)	129
4.5 Evaluasi dan Validasi Hasil	130
4.5.1 Evaluasi Hasil <i>Clustering</i> pada Dataset Iris	130
4.5.2 Evaluasi Hasil <i>Clustering</i> pada Dataset Wine	131
4.5.3 Tabel Klasifikasi/ <i>Clustering</i> Data Iris	132
4.5.4 Tabel Klasifikasi/ <i>Clustering</i> Data Wine	134
4.6 Kajian Penelitian dalam Perspektif Islam	136
BAB V KESIMPULAN DAN SARAN	140
5.1 Kesimpulan	140
5.2 Saran untuk Penelitian Lanjutan	141
DAFTAR PUSTAKA	142
LAMPIRAN	145
RIWAYAT HIDUP	152

DAFTAR TABEL

Tabel 3.1	Dataset Penelitian.....	21
Tabel 4.1	Lima Data Pertama Dataset Iris	31
Tabel 4.2	Lima Data Pertama Dataset Wine.....	32
Tabel 4.3	Pengelompokan <i>Cluster</i> pada Iterasi 1	37
Tabel 4.4	Pengelompokan <i>Cluster</i> pada Iterasi 2	39
Tabel 4.5	Pengelompokan <i>Cluster</i> pada Iterasi 1	50
Tabel 4.6	Pengelompokan <i>Cluster</i> pada Iterasi 2	52
Tabel 4.7	Pengelompokan <i>Cluster</i> pada Iterasi 4	53
Tabel 4.8	Inisialisasi Posisi Partikel pada PSO.....	55
Tabel 4.9	Pengelompokan <i>Cluster</i> Partikel 1 pada Iterasi 1	58
Tabel 4.10	Pengelompokan <i>Cluster</i> Partikel 2 pada Iterasi 1	60
Tabel 4.11	Pengelompokan <i>Cluster</i> Partikel 3 pada Iterasi 1	63
Tabel 4.12	Hasil Iterasi 1 (<i>centroid</i> setelah <i>update</i>)	80
Tabel 4.13	Pengelompokan <i>Cluster</i> Partikel 1 pada Iterasi 2	82
Tabel 4.14	Pengelompokan <i>Cluster</i> Partikel 2 pada Iterasi 2	85
Tabel 4.15	Pengelompokan <i>Cluster</i> Partikel 3 pada Iterasi 2	87
Tabel 4.16	Inisialisasi Posisi Partikel pada PSO.....	105
Tabel 4.17	Pengelompokan <i>Cluster</i> Partikel 1 pada Iterasi 1	108
Tabel 4.18	Pengelompokan <i>Cluster</i> Partikel 2 pada Iterasi 1	110
Tabel 4.19	Pengelompokan <i>Cluster</i> Partikel 3 pada Iterasi 1	113
Tabel 4.20	Rata-rata Evaluasi <i>K-Means</i> pada Dataset Iris.....	130
Tabel 4.21	Rata-rata Evaluasi <i>K-Means</i> pada Dataset Wine	131
Tabel 4.22	Perbandingan Evaluasi pada Penelitian Sebelumnya.....	132
Tabel 4.23	Perbandingan Evaluasi pada Penelitian Sebelumnya.....	134

DAFTAR GAMBAR

Gambar 4.1 Hasil <i>K-Means Clustering</i> Dataset Iris (5 Data pertama)	46
Gambar 4.2 Hasil <i>K-Means Clustering</i> Dataset Iris.....	47
Gambar 4.3 Hasil <i>K-Means Clustering</i> Dataset Wine (5 Data pertama)	53
Gambar 4.4 Hasil <i>K-Means Clustering</i> Dataset Wine	55
Gambar 4.5 Hasil <i>K-Means</i> – PSO Dataset Iris (5 Data Pertama).....	127
Gambar 4.6 Hasil <i>K-Means</i> – PSO Dataset Iris	128
Gambar 4.7 Hasil <i>K-Means</i> – PSO Dataset Wine (5 Data Pertama)	129
Gambar 4.8 Hasil <i>K-Means</i> – PSO Dataset Wine.....	130

ABSTRAK

Afifi, Muthia. 2025. **Optimasi Algoritma *K-Means* Menggunakan *Particle Swarm Optimization***. Skripsi. Program Studi Matematika Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr. Mohammad Jamhuri, M.Si. (II) Erna Herawati, M.Pd.

Kata Kunci: *clustering*, *K-Means*, *Particle Swarm Optimization*, optimasi, *centroid*, iris, wine

Pengelompokan data (*clustering*) merupakan teknik penting dalam analisis data tidak berlabel, terutama ketika volume dan kompleksitas data terus meningkat. *K-Means* adalah salah satu algoritma *clustering* yang populer, namun memiliki kelemahan dalam proses inisialisasi *centroid* secara acak, yang dapat menyebabkan solusi sub-optimal. Penelitian ini bertujuan untuk meningkatkan performa *K-Means* melalui integrasi dengan algoritma *Particle Swarm Optimization* (PSO) dalam menentukan inisialisasi *centroid* awal. Pengujian dilakukan pada dua dataset klasik, yaitu Iris dan Wine. Evaluasi kinerja dilakukan menggunakan metrik *Sum of Squared Error* (SSE), *silhouette score*, dan akurasi klasifikasi. Hasil eksperimen menunjukkan bahwa metode *K-Means*+PSO menghasilkan peningkatan akurasi pada dataset Iris (89,3%) dan Wine (97,2%) dibandingkan *K-Means* standar (88,7% dan 96,6%). Peningkatan juga terlihat pada nilai *silhouette score*, meskipun perbedaan SSE relatif kecil. Temuan ini menunjukkan bahwa integrasi PSO dapat meningkatkan kualitas hasil pengelompokan dibandingkan metode *K-Means* standar.

ABSTRACT

Afifi, Muthia. 2025. **Optimization of the K-Means Algorithm Using Particle Swarm Optimization**. Thesis. Department of Mathematics, Faculty of Science and Technology, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Advisors: (I) Dr. Mohammad Jamhuri, M.Si. (II) Erna Herawati, M.Pd.

Keywords: clustering, K-Means, Particle Swarm Optimization, optimization, centroid, iris, wine

Data clustering is a fundamental technique in the analysis of unlabeled data, especially as the volume and complexity of data continue to grow. K-Means is one of the most widely used clustering algorithms; however, it suffers from a limitation in the form of random initialization of centroids, which can lead to sub-optimal solutions. This study aims to improve the performance of K-Means by integrating it with Particle Swarm Optimization (PSO) for better centroid initialization. Experiments were conducted on two classical datasets, namely Iris and Wine. The performance was evaluated using three metrics: Sum of Squared Error (SSE), silhouette score, and classification accuracy. The experimental results show that the K-Means + PSO method achieves higher accuracy on the Iris (89.3%) and Wine (97.2%) datasets compared to standard K-Means (88.7% and 96.6%). Improvements were also observed in silhouette scores, although the differences in SSE were relatively small. These findings indicate that integrating PSO can enhance the quality of clustering results compared to the standard K-Means method.

مستخلص البحث

عفيفي، مطيعة. ٢٠٢٥. تحسين خوارزمية *K-Means* باستخدام خوارزمية تحسين سرب الجسيمات (PSO). البحث الجامعي. قسم الرياضيات، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج. المشرف: (١) الدكتور محمد جمهوري، الماجستير في العلوم. (٢) إيرنا هيراواقي، الماجستير في تعليم اللغة العربية.

الكلمات الأساسية: التجميع، *K-Means*، تحسين سرب الجسيمات، التحسين، المركز، إيريس، واين

يُعدّ تجميع البيانات (*Clustering*) من التقنيات الأساسية في تحليل البيانات غير المصنّفة، خاصةً مع التزايد المستمر في حجم البيانات وتعقيدها. وتُعدّ خوارزمية *K-Means* من أشهر خوارزميات التجميع، إلا أنها تعاني من ضعف في اختيار النقاط المركزية (*centroids*) عشوائيًا في البداية، مما يؤدي إلى حلول غير مثالية. هدفت هذه الدراسة إلى تحسين أداء خوارزمية *K-Means* من خلال دمجها مع خوارزمية تحسين سرب الجسيمات (PSO) لاختيار المراكز الأولية. تم إجراء الاختبار على مجموعتي بيانات كلاسيكيتين: إيريس (Iris) و واين (Wine). تم تقييم الأداء باستخدام ثلاث مقاييس: مجموع مربعات الخطأ (*SSE*)، ودرجة الظل (*silhouette score*)، ودقة التصنيف. أظهرت النتائج أن خوارزمية *K-Means + PSO* حققت دقة أعلى في مجموعة بيانات Iris (٨٩,٣٪) و Wine (٩٧,٢٪) مقارنة بخوارزمية *K-Means* التقليدية (٨٨,٧٪ و ٩٦,٦٪). كما لوحظ تحسّن في درجة الظل، رغم أن الفرق في *SSE* كان طفيفًا. تُبيّن هذه النتائج أن دمج PSO يعزّز جودة نتائج التجميع مقارنةً بخوارزمية *K-Means* وحدها.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Seiring berjalannya kemajuan teknologi informasi saat ini, keperluan akan informasi dari sekumpulan data semakin berkembang pesat. Data menjadi salah satu aset penting bagi berbagai sektor, termasuk organisasi, perusahaan, hingga pemerintah dalam proses pengambilan keputusan yang lebih tepat dan efisien. Pengelolaan data yang tepat memungkinkan pengambilan keputusan berbasis fakta yang lebih optimal. Namun, semakin banyak data yang dihasilkan, semakin kompleks pula tantangan dalam menganalisis dan memahami pola-pola yang tersembunyi di dalamnya. Oleh karena itu, muncul kebutuhan akan teknik dan alat yang lebih canggih untuk mengolah, mengelompokkan, dan menginterpretasikan data tersebut.

Salah satu metode yang sering digunakan untuk memproses data adalah *clustering*, yaitu teknik untuk mengelompokkan objek-objek ke dalam grup berdasarkan kesamaan atau kedekatan karakteristik (Jain, 2010). *Clustering* merupakan alat penting dalam analisis data statistik untuk melakukan pengolahan data (Nur Afidah, 2023). Dengan demikian, *clustering* sangat berguna dalam menemukan grup atau kelompok yang sebelumnya tidak dikenal dalam data (Haddad, 2011).

Dalam kajian integrasi perspektif Islam, QS. Al-Hujurat ayat 13, Allah SWT menjelaskan bahwa manusia diciptakan dalam berbagai kelompok seperti bangsa dan suku, yang bertujuan untuk saling mengenal satu sama lain.

أَيُّهَا النَّاسُ إِنَّا خَلَقْنَاكُمْ مِنْ ذَكَرٍ وَأُنْثَىٰ وَجَعَلْنَاكُمْ شُعُوبًا وَقَبَائِلَ لِتَعَارَفُوا إِنَّ أَكْرَمَكُمْ عِنْدَ اللَّهِ أَتْقَىٰ إِنَّ اللَّهَ عَلِيمٌ خَبِيرٌ ﴿١٣﴾

(Kemenag, 2024)

Artinya: “Wahai manusia, sesungguhnya Kami menciptakan kamu dari seorang laki-laki dan perempuan. Kemudian, kami menjadikan kamu berbangsa-bangsa dan bersuku-suku agar kamu saling mengenal. Sesungguhnya yang paling mulia di antara kamu di sisi Allah adalah orang yang paling bertakwa. Sesungguhnya Allah Maha Mengetahui lagi Maha Teliti.”

Ayat diatas menjelaskan bahwa Allah SWT menciptakan manusia dari seorang laki-laki dan perempuan, kemudian menjadikannya berbangsa-bangsa dan bersuku-suku agar mereka saling mengenal. Meskipun setiap bangsa dan suku memiliki perbedaan dalam budaya, bahasa, dan kebiasaan, mereka tetap merupakan bagian dari satu kesatuan umat manusia. Ayat ini menekankan bahwa keberagaman bukan untuk saling memisahkan, tetapi untuk dipahami dan dikenali dalam konteks yang lebih luas.

Dalam konsep *clustering*, prinsip ini bisa diterapkan jika data yang dimiliki berbeda dengan berbagai karakteristik, dikelompokkan berdasarkan kesamaan tertentu untuk menemukan pola dan keteraturan. Meskipun terdapat perbedaan antar data, seperti halnya perbedaan antara bangsa dan suku, tujuan pengelompokan adalah untuk memahami keteraturan yang ada di balik perbedaan tersebut. Sebagai contoh, jika menerapkan algoritma K-Means pada data sosial ekonomi penduduk, kita dapat mengelompokkan masyarakat ke dalam beberapa kategori berdasarkan pendapatan, tingkat pendidikan, atau pola konsumsi. Hal ini mencerminkan konsep dalam ayat tersebut, di mana keberagaman dalam data tetap bisa dikelompokkan tanpa menghilangkan identitas masing-masing.

K-Means adalah salah satu algoritma *clustering* yang sering digunakan untuk mengelompokkan data berdasarkan kedekatan karakteristik antar data.

Algoritma ini bekerja dengan cara mengelompokkan data ke dalam *cluster* berdasarkan jarak ke pusat *cluster* (*centroid*). Namun, salah satu kelemahan utama dari *K-Means* adalah pemilihan *centroid* awal yang dilakukan secara acak, yang dapat menyebabkan algoritma terjebak dalam solusi lokal (J. A. Hartigan, 2012). Oleh karena itu, diperlukan suatu metode untuk mengoptimalkan pemilihan *centroid* sehingga algoritma dapat menghasilkan *clustering* yang lebih akurat dan stabil.

Dalam penelitian ini, penulis mengusulkan penggunaan *Particle Swarm Optimization* (PSO) sebagai upaya untuk mengatasi keterbatasan tersebut dengan mencari solusi global yang lebih baik, terutama dalam pemilihan *centroid*. PSO adalah algoritma optimasi yang terinspirasi oleh perilaku kawanan burung, di mana partikel-partikel dalam *swarm* bekerja sama untuk mencari solusi optimal dalam ruang pencarian. PSO dapat digunakan untuk mengatasi masalah lokal optimal pada *K-Means* dengan memperbaiki pemilihan *centroid*, sehingga menghasilkan *clustering* yang lebih efektif dan akurat (Yudi Agusta, 2007).

Penelitian ini bertujuan untuk menguji apakah optimasi dengan PSO dapat meningkatkan akurasi dan efektivitas *clustering* pada *K-Means*, yang akan dibandingkan dengan *K-Means* standar, khususnya dalam pengelompokan data Iris dan data Wine. Dataset ini dipilih karena keduanya sering digunakan sebagai contoh dalam pengujian berbagai algoritma *clustering*, termasuk *K-Means*.

Binti Muslimatin (2011) meneliti perbedaan antara *K-Means* dan FCM pada data saham harian PT.Astra, Tbk, menunjukkan bahwa *K-Means* secara tegas mengalokasikan data ke *cluster* terdekat. Sedangkan FCM menggunakan teori *fuzzy* untuk alokasi lebih fleksibel, dengan risiko kegagalan konvergensi di kedua metode

(Muslimatin, 2011). Yeni Dwi Mayangsari (2022) menggunakan *K-Means* untuk mengelompokkan kabupaten-kota di Provinsi Jawa Timur berdasarkan kasus COVID-19, menghasilkan empat *cluster* dengan tingkat kasus yang berbeda (YENI, 2022). Nanda Azzahrotun Nafisa (2023) membandingkan metode DBSCAN dan *K-Means* dalam menganalisis pola penyebaran tempat wisata di Malang Raya, dengan hasil menunjukkan bahwa *K-Means* lebih baik dalam hal validitas *cluster* dibandingkan DBSCAN berdasarkan nilai *Silhouette Coefficient* (Nafisa, 2023).

Sementara itu, penelitian oleh Jamhuri dan Utomo (2024) menunjukkan bahwa PSO efektif dalam meningkatkan kinerja jaringan syaraf tiruan (JST) untuk klasifikasi sinyal radar. Dalam penelitian tersebut, PSO digunakan untuk menentukan jumlah neuron optimal dalam *hidden layer* JST, yang secara signifikan meningkatkan *accuracy*, *precision*, *recall*, dan *F1-score* dibandingkan metode baseline seperti SVM, Naive Bayes, dan Logistic Regression. Temuan ini menunjukkan bahwa PSO mampu menghindari solusi lokal dan mengeksplorasi ruang solusi secara efisien, memberikan hasil yang lebih stabil dan akurat (Jamhuri & Utomo, 2024).

Namun, pada penelitian-penelitian tersebut, belum ada yang membahas tentang pengoptimalan inisialisasi *centroid* pada *K-Means* untuk menghindari masalah inisialisasi acak yang sering menyebabkan hasil *clustering* terjebak pada solusi lokal. Selain itu, tidak ada yang melakukan perbandingan langsung antara *K-Means* standar dengan *K-Means* yang dioptimalkan dengan PSO.

Mengacu pada penelitian yang telah disebutkan, jelas bahwa setiap metode *clustering* memiliki kelebihan dan keterbatasannya masing-masing. Oleh karena

itu, penting untuk mengeksplorasi dan mengkombinasikan metode-metode ini untuk mengatasi berbagai tantangan yang dihadapi dalam analisis data. Dengan demikian ini sejalan dengan penelitian yang akan penulis lakukan yaitu Optimasi Algoritma *K-Means* Menggunakan PSO dan kebaruan penelitian ini terletak pada pengaplikasian PSO untuk optimasi inisialisasi *centroid*, yang belum dilakukan dalam penelitian terdahulu, serta pada pengaruh jumlah iterasi PSO terhadap kualitas *clustering*.

1.2 Rumusan Masalah

Rumusan masalah dalam penelitian ini yaitu:

1. Bagaimana optimalisasi algoritma *K-Means* dengan *Particle Swarm Optimization* (PSO)?
2. Bagaimana performa *clustering* algoritma *K-Means* yang dioptimalkan dengan PSO?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini yaitu:

1. Mengoptimalkan algoritma *K-Means* menggunakan metode PSO untuk meningkatkan efektivitas proses pengelompokan data.
2. Menganalisis dan membandingkan performa *clustering K-Means* dengan *K-Means* yang dioptimalkan menggunakan PSO.

1.4 Manfaat Penelitian

Manfaat dari penelitian ini berdasarkan penjelasan diatas terdapat dua aspek yaitu:

1. Manfaat akademis

Penelitian ini memberikan kontribusi dalam pengembangan metode optimasi *centroid*. Dengan hasil penelitian ini, diharapkan akademisi dan penulis dapat memahami bagaimana optimisasi *centroid* dapat memengaruhi akurasi dan efisiensi pengelompokan data.

2. Manfaat praktis

Secara praktis, hasil penelitian ini dapat meningkatkan akurasi pengelompokan, dan efisiensi analisis data berbasis *clustering*.

1.5 Batasan Masalah

Pada penelitian ini, penulis akan membatasi masalah dengan fokus pada penerapan PSO untuk meningkatkan kinerja algoritma *K-Means* dalam mengelompokkan data Iris dan data Wine. PSO akan digunakan untuk mengoptimalkan proses inisialisasi *centroid* pada algoritma *K-Means*. Algoritma *K-Means* yang digunakan dalam penelitian ini adalah versi standar, yang mengelompokkan data berdasarkan jarak *Euclidean* antara titik data dan *centroid*.

1.6 Definisi Istilah

1. *K-Means Clustering* : Algoritma pengelompokan yang digunakan untuk membagi sekumpulan data ke dalam kelompok-kelompok berdasarkan kesamaan. Setiap kelompok diwakili oleh sebuah titik pusat yang disebut

centroid, dan data dikelompokkan berdasarkan jarak terdekat ke *centroid* tersebut.

2. *Particle Swarm Optimization* (PSO) : Metode pengoptimalan yang meniru perilaku kelompok hewan, seperti kawanan burung saat mencari makanan. Dalam PSO, sejumlah partikel berinteraksi satu sama lain untuk menemukan solusi terbaik dengan mengikuti dua informasi utama: posisi terbaik yang telah dicapai oleh setiap partikel dan posisi terbaik yang ditemukan oleh seluruh kelompok.
3. *Centroid* : Titik pusat yang mewakili lokasi rata-rata dari semua data dalam sebuah kelompok. Dalam konteks *K-Means*, *centroid* digunakan untuk menentukan batasan kelompok dan memandu proses pengelompokan.
4. Data Iris : Dataset yang berisi informasi tentang tiga spesies bunga Iris. Setiap spesies terdiri dari 50 sampel, dengan empat fitur morfologis.
5. Data Wine : Dataset yang berisi informasi tentang tiga jenis anggur yang berbeda dan memiliki 13 fitur kimia yang berbeda untuk setiap sampel anggur.
6. *Clustering* : Teknik analisis data yang digunakan untuk mengelompokkan objek atau data ke dalam kelompok yang lebih kecil berdasarkan kesamaan atau kedekatan antar data tersebut. Teknik ini sering digunakan untuk menemukan pola atau struktur dalam data.

BAB II

KAJIAN PUSTAKA

2.1 Teori *Clustering*

Klasterisasi atau yang dikenal sebagai *clustering*, adalah teknik untuk mengelompokkan data. *Clustering* merupakan alat penting dalam analisis data statistik untuk melakukan pengolahan data (Nur Afidah, 2023). Proses ini membagi sekumpulan objek data menjadi beberapa *cluster*, di mana setiap objek dalam satu *cluster* memiliki kesamaan karakteristik.

Menurut Alashwal et al. (2019, seperti dikutip dalam Wu et al., 2021), analisis klaster atau *clustering* merupakan metode yang digunakan ketika algoritma klasifikasi tidak dapat diterapkan, yaitu saat informasi mengenai setiap kategori tidak diketahui sebelumnya atau data yang tersedia tidak memiliki label kategori yang sesuai. Dalam kondisi seperti ini, analisis klaster memungkinkan pengelompokan objek-objek yang memiliki kemiripan ke dalam kategori atau subset tertentu melalui proses klasifikasi statis. Objek-objek yang berada dalam satu subset diasumsikan memiliki sifat-sifat yang serupa.

Menurut (Benabdellah et al., 2019), klastering adalah teknik yang bertujuan untuk mengelompokkan data (objek) ke dalam klaster, di mana objek yang serupa dikelompokkan dalam satu klaster, sementara objek yang berbeda ditempatkan dalam klaster yang terpisah. Dari perspektif optimasi, tujuan utama klastering adalah memaksimalkan homogenitas internal dalam klaster dan heterogenitas eksternal antar klaster yang berbeda. Berbeda dengan klasifikasi, klastering tidak memerlukan data yang sudah memiliki label kelas, sehingga klastering dianggap

sebagai tugas unsupervised (tanpa pengawasan), yang mengukur kesamaan antar objek tanpa informasi mengenai distribusi yang benar atau *ground truth*.

Metode klasterisasi juga perlu mampu mengukur efektivitasnya dalam mencari pola tersembunyi dalam data yang sedang dianalisis. Salah satu algoritma *clustering* yang paling banyak digunakan adalah *K-Means*, yang dikembangkan oleh MacQueen (1967). Algoritma ini membagi data ke dalam sejumlah *cluster* yang sudah ditentukan sebelumnya, dengan meminimalkan jarak antara data dan pusat *cluster* (*centroid*). Setiap iterasi dalam *K-Means* bertujuan untuk memperbaiki pembagian *cluster* hingga mencapai konvergensi, di mana perpindahan data antar *cluster* tidak lagi signifikan.

2.2 Algoritma *K-Means*

K-Means adalah algoritma pengelompokan data non-hirarki yang mengelompokkan data ke dalam satu atau lebih *cluster*. Metode *K-Means* merupakan salah satu algoritma yang tergolong cepat dan efisien saat digunakan, karena tidak membutuhkan banyak waktu dan efisiensi komputasi yang tinggi. Hal ini menjadikan *K-Means* sebagai metode yang populer dalam pengelompokan data berukuran besar. Namun demikian, metode ini memiliki keterbatasan, terutama dalam menangani data berdimensi tinggi karena dapat mengurangi akurasi hasil klastering. Selain itu, *K-Means* tidak mampu mengenali klaster yang berbentuk tidak bulat atau tidak simetris (*nonspherical*), sehingga hasil pengelompokan bisa menjadi kurang representatif jika data memiliki pola bentuk yang kompleks (Wu et al., 2021).

Algoritma ini bertujuan untuk mengelompokkan data ke dalam beberapa *cluster*, di mana data yang memiliki kesamaan karakteristik berada dalam kelompok yang sama, sedangkan data yang berbeda akan dikelompokkan ke dalam kelompok lain. Algoritma ini termasuk dalam kategori *unsupervised learning* karena memiliki sifat iteratif dan numerik.

Algoritma ini cocok untuk data yang variabelnya bersifat kuantitatif, dan jarak antar titik data diukur menggunakan jarak *Euclidean* kuadrat.

Rumus untuk menghitung jarak *Euclidean* adalah sebagai berikut:

Diberikan $\bar{x}_i = (x_{i1}, x_{i2}, \dots, x_{iq})$ dan $\bar{\mu}_j = (\mu_{j1}, \mu_{j2}, \dots, \mu_{jq})$ dengan $i = 1, 2, \dots, m$, dan $j = 1, 2, \dots, K$.

$$d(\bar{x}_i, \bar{\mu}_j) = \|\bar{x}_i - \bar{\mu}_j\| \quad (2.1)$$

Persamaan (2.1) menyatakan jarak antara suatu data $\bar{x}_i$ dan *centroid* $\bar{\mu}_j$ menggunakan norm *Euclidean*. Secara lebih rinci, persamaan tersebut dapat dituliskan sebagai berikut:

$$\begin{aligned} \|\bar{x}_i - \bar{\mu}_j\| &= \|(x_{i1}, x_{i2}, \dots, x_{iq}) - (\mu_{j1}, \mu_{j2}, \dots, \mu_{jq})\| \\ &= \|(x_{i1} - \mu_{j1}, x_{i2} - \mu_{j2}, \dots, x_{iq} - \mu_{jq})\| \\ &= \sqrt{(x_{i1} - \mu_{j1})^2 + (x_{i2} - \mu_{j2})^2 + \dots + (x_{iq} - \mu_{jq})^2} \end{aligned} \quad (2.2)$$

di mana:

- $d(\bar{x}_i, \bar{\mu}_j)$ = Jarak antara titik data ke- i dan *centroid* ke- j
- q = Jumlah fitur yang akan diamati
- x_{il} = Nilai fitur ke- l dari titik data ke- i
- μ_{jl} = Nilai fitur ke- l dari *centroid* ke- j (pusat *cluster*)

(Hastie dkk., 2009)

Dengan demikian, persamaan (2.1) merupakan bentuk umum dari jarak *Euclidean*, sedangkan persamaan (2.2) merupakan bentuk eksplisitnya yang menunjukkan bahwa jarak dihitung dari akar kuadrat jumlah selisih kuadrat setiap fitur antara data dan *centroid*-nya.

Secara matematis, diberikan sebuah $C = \{C_1, C_2, C_3, \dots, C_K\}$. Tujuan dari algoritma K-Means adalah untuk meminimalkan fungsi $SSE(C)$.

Dengan menggunakan rumus jarak *Euclidean* yang telah didefinisikan diatas, maka fungsi SSE dapat ditulis sebagai:

$$SSE(C) = \sum_{j=1}^K \sum_{x_i \in C_j} \sum_{l=1}^q (x_{il} - \mu_{jl})^2 \quad (2.3)$$

di mana:

- K adalah jumlah *cluster*
- q adalah jumlah fitur
- C_j adalah *cluster* ke- j
- $\bar{x}_i$ adalah titik data yang termasuk dalam *cluster* C_j
- $\bar{\mu}_j$ adalah *centroid* (pusat) dari *cluster* C_j , yaitu rata-rata dari semua titik di *cluster* tersebut.

Algoritma K-Means bertujuan untuk menemukan *cluster* C^* yang meminimalkan $SSE(C)$:

$$C^* = \arg \min_c SSE(C) \quad (2.4)$$

Langkah-langkah algoritma:

1. Inisialisasi: Pilih secara acak k titik sebagai *centroid* awal.
2. Langkah iteratif:

- a. Penugasan *cluster*: Setiap titik data x_i ditugaskan ke *cluster* dengan *centroid* terdekat (berdasarkan jarak *Euclidean*).
 - b. Pembaruan *centroid*: Setelah semua titik ditugaskan, *centroid* diperbarui menjadi rata-rata dari semua titik yang ada dalam *cluster* tersebut.
3. Kriteria konvergensi: Iterasi dilanjutkan sampai perbedaan antara *centroid* pada iterasi sekarang dan iterasi sebelumnya kurang dari nilai ambang batas (ϵ). Secara matematis, ini dinyatakan sebagai:

$$\sum_{j=1}^K \|\bar{\mu}_j^{(t)} - \bar{\mu}_j^{(t-1)}\|^2 \leq \epsilon \quad (2.5)$$

di mana:

- $\epsilon > 0$ konvergen
- t adalah iterasi ke- t
- $\bar{\mu}_j^{(t)}$ adalah *centroid* ke- j pada iterasi t
- $\bar{\mu}_j^{(t-1)}$ adalah *centroid* ke- j pada iterasi sebelumnya (iterasi ke- $(t - 1)$).

(Zaki & Meira, 2020)

2.3 Particle Swarm Optimization (PSO)

PSO merupakan metode optimisasi global berbasis heuristik yang mensimulasikan perilaku sosial kawanan burung dalam bergerak menuju posisi yang menjanjikan untuk mencapai tujuan yang tepat dalam ruang multidimensi (Sheikhpour et al., 2016). Metode ini dikembangkan oleh James Kennedy dan Russell Eberhart pada tahun 1995. Kennedy dan Eberhart mempublikasikan makalah berjudul “Particle Swarm Optimization” di konferensi IEEE *International Conference on Neural Networks* pada tahun tersebut, yang membahas metode PSO

dan penggunaannya dalam optimisasi non-linier.

Konsep partikel diperkenalkan oleh Kennedy dan Eberhart sebagai partikel yang bergerak dalam area pencarian dan berinteraksi satu sama lain.. Prinsip dasar algoritma PSO menyatakan bahwa setiap partikel dalam populasi merujuk pada dua informasi utama yaitu lokasi terbaik yang pernah dicapai ($\bar{p}_{best}$) dan lokasi terbaik yang dicapai oleh seluruh populasi ($\bar{g}_{best}$) (Eberhart, 1995).

Dalam PSO, sekelompok partikel bergerak di ruang pencarian untuk menemukan solusi optimal. Setiap partikel i memiliki dua vektor utama yaitu:

1. Vektor kecepatan

$$\bar{v}_i = (v_{i1}, v_{i2}, \dots, v_{il})$$

yang menunjukkan kecepatan partikel i dalam ruang pencarian.

2. Vektor posisi

$$\bar{p}_i = (p_{i1}, p_{i2}, p_{i3}, \dots, p_{il})$$

yang menunjukkan posisi partikel i dalam ruang pencarian.

PSO mencari solusi optimal dengan memperbarui kecepatan dan posisi setiap partikel berdasarkan persamaan berikut:

$$\bar{v}_{il}^{t+1} = \omega \bar{v}_{il}^t + a_1 r_1 (\bar{p}_{best_{il}} - \bar{p}_{il}^t) + a_2 r_2 (\bar{g}_{best_{il}} - \bar{p}_{il}^t) \quad (2.5)$$

$$\bar{p}_{il}^{t+1} = \bar{p}_{il}^t + \bar{v}_{il}^{t+1} \quad (2.6)$$

dimana:

- t menunjukkan iterasi , dan l menunjukkan dimensi ke- l dalam ruang pencarian.
- ω adalah bobot inersia.
- a_1 dan a_2 adalah faktor pembelajaran individu dan kelompok.
- r_1 dan r_2 adalah nilai acak dengan rentang $[0,1]$.

Proses dasar algoritma PSO adalah sebagai berikut:

1. Inisialisasi: Partikel diinisialisasikan dengan posisi dan kecepatan yang acak.
2. Evaluasi: Nilai fungsi objektif dihitung untuk setiap partikel.
3. Menentukan $\bar{p}_{best}$: Jika nilai fungsi objektif dari partikel i lebih baik daripada $\bar{p}_{best}$ sebelumnya, maka nilai tersebut disimpan sebagai $\bar{p}_{best}$ yang baru.
4. Menentukan $\bar{g}_{best}$: Jika ada $\bar{p}_{best}$ yang lebih baik daripada $\bar{g}_{best}$, maka $\bar{g}_{best}$ diperbarui dengan nilai tersebut.
5. Perbarui kecepatan dan posisi: Kecepatan setiap partikel diperbarui berdasarkan persamaan (2.5), dan posisi diperbarui menggunakan persamaan (2.6).
6. Kriteria berhenti: Jika jumlah iterasi telah tercapai, algoritma akan berhenti, jika belum, proses kembali ke langkah 2.

(Sheikhpour et al., 2016).

2.4 Data Iris dan Data Wine

Dataset Iris adalah kumpulan data yang umum digunakan dalam penelitian. Dikenal juga sebagai dataset Iris Fisher, dataset bunga Iris adalah kumpulan data multivariat yang diperkenalkan oleh Ronald Fisher, seorang ahli statistik dan biologi asal Inggris, dalam makalahnya yang dipublikasikan pada tahun 1936 berjudul “*The Use of Multiple Measurements in Taxonomic Problem*” sebagai contoh analisis diskriminan linier (R. A. Fisher, Sc.D., 1936). Dataset ini kadang disebut sebagai dataset Iris Anderson karena data untuk mengukur variasi morfologi bunga Iris dari tiga spesies terkait dikumpulkan oleh Edgar Anderson (Anderson, 1936).

Dataset ini berisi informasi tentang variasi morfologi bunga Iris dari tiga spesies yaitu Iris setosa, Iris virginica, dan Iris versicolor. Setiap spesies terdiri dari 50 sampel dan terdapat empat fitur diukur dari setiap sampel yaitu panjang dan lebar sepal (bagian luar bunga) serta panjang dan lebar petal (bagian dalam bunga), yang diukur dalam sentimeter.

Dataset ini sering digunakan sebagai contoh dalam berbagai teknik klasifikasi dalam machine learning, seperti *support vector machine* (SVM), untuk membantu mengembangkan dan menguji algoritma klasifikasi. Dalam penelitian ini, data Iris akan digunakan sebagai objek untuk menguji penerapan algoritma *K-Means* yang dioptimalkan dengan PSO. Pengelompokan ini bertujuan untuk mengidentifikasi pola dan kelompok dalam data, serta mengevaluasi kinerja metode *clustering* yang dioptimalkan.

Selain dataset Iris, penelitian ini juga menggunakan dataset Wine yang merupakan hasil analisis kimiawi anggur yang ditanam di wilayah yang sama di Italia tetapi berasal dari tiga jenis anggur yang berbeda. Analisis ini mengidentifikasi 13 komponen kimia yang ada di setiap jenis anggur, seperti alkohol, asam sulfat, abu, magnesium, dan lainnya.

Dalam penelitian ini, dataset Wine akan digunakan untuk menguji performa algoritma *K-Means* yang dioptimalkan dengan PSO, dengan tujuan mengelompokkan sampel anggur berdasarkan fitur kimianya dan mengevaluasi efektivitas metode *clustering* yang ditingkatkan.

2.5 Kajian Integrasi Topik dengan Al-Qur'an/Hadits

Terdapat ayat Al-Qur'an yang relevan dengan konsep *clustering* dalam ilmu data. Pada QS. Al-Waqiah ayat 7-10, Allah SWT menjelaskan pembagian manusia pada hari kiamat menjadi beberapa golongan.

وَكُنْتُمْ أَزْوَاجًا ثَلَاثَةً ﴿٧﴾ فَأَصْحَابُ الْمَيْمَنَةِ مَا أَصْحَابُ الْمَيْمَنَةِ ﴿٨﴾ وَ أَصْحَابُ الْمَشْأَمَةِ مَا
أَصْحَابُ الْمَشْأَمَةِ ﴿٩﴾ وَالسَّابِقُونَ السَّابِقُونَ ﴿١٠﴾

(Kemenag, 2024)

Artinya: “Dan kamu menjadi tiga golongan. Yaitu golongan kanan, alangkah mulianya golongan kanan itu. Dan golongan kiri, alangkah sengsaranya golongan kiri itu. Selain itu, (golongan ketiga adalah) orang-orang yang paling dahulu (beriman). Merekalah yang paling dahulu (masuk surga).”

Berdasarkan tafsir para ulama, ayat ini menegaskan bahwa pada hari pembalasan (*Yaumul Qiyamah*), manusia akan dibagi menjadi tiga golongan besar yang mencerminkan amal dan perbuatan mereka selama hidup di dunia. Menurut Tafsir Ibnu Katsir (Al-Sheikh, 2005) golongan-golongan tersebut ialah golongan kanan, golongan kiri, dan golongan *as-sabiqun*. Pembagian ini mencerminkan proses pengelompokan manusia ke dalam kategori yang relevan dengan karakteristik mereka, yaitu amal kebaikan atau keburukan.

Penjelasan tentang pembagian manusia menjadi golongan kanan dan kiri juga diperkuat oleh hadits Rasulullah SAW. yang diriwayatkan oleh Al-Bukhari dan Muslim.

“Pada hari kiamat, orang-orang beriman akan diberi kitab catatan amal mereka dengan tangan kanan. Sedangkan orang-orang kafir dan munafik akan diberi kitab catatan amal mereka dengan tangan kiri.” (HR.Al-Bukhari, No.2443 dan Muslim, No.302).

Dalam ilmu data, *clustering* adalah metode pengelompokan data berdasarkan kemiripan karakteristik. Misalnya, algoritma *clustering* seperti K-

Means atau PSO akan mengelompokkan data ke dalam *cluster* tertentu berdasarkan fitur atau atribut yang dimiliki. Dalam konteks ayat ini, "atribut" manusia adalah amal perbuatan mereka, sedangkan "*cluster*" adalah golongan kanan, kiri, dan *as-sabiqun*.

Proses *clustering* dalam data menyerupai konsep pembagian manusia ke dalam golongan seperti yang disebutkan dalam Al-Qur'an. Dalam hal ini terdapat tiga golongan yang dianggap sebagai *cluster* dalam data. Karakteristik data berupa amal perbuatan manusia, seperti kejujuran, ketaatan, atau dosa, bertindak sebagai fitur yang menentukan keanggotaan seseorang dalam suatu *cluster*. Sama seperti algoritma *clustering* yang membagi data berdasarkan kemiripan, manusia dikelompokkan oleh Allah SWT berdasarkan amal perbuatan mereka yang serupa. Golongan kanan dihuni oleh orang-orang yang beriman dan taat, golongan kiri oleh orang-orang yang ingkar, dan golongan *as-sabiqun* oleh mereka yang memiliki amal paling unggul.

Hubungan ini mengajarkan bahwa kehidupan manusia pada akhirnya akan dievaluasi dan "dikelompokkan" oleh Allah SWT sesuai amal perbuatannya, mengingatkan manusia untuk terus memperbaiki diri agar masuk ke golongan yang mulia. Dengan demikian, konsep *clustering* secara tidak langsung menggambarkan proses pembagian manusia berdasarkan amal di sisi Allah SWT.

2.6 Kajian Topik Dengan Teori Pendukung

Algoritma *K-Means* merupakan salah satu metode yang sering digunakan dalam proses pengelompokan data atau *clustering*. Algoritma ini bekerja dengan membagi data berdasarkan kedekatan antara data satu dengan yang lain, dimana

setiap data dalam satu kelompok memiliki tingkat kesamaan yang tinggi. Namun, kelemahan utama dari *K-Means* terletak pada inisialisasi awal *centroid* atau pusat *cluster* yang dipilih secara acak. Pemilihan *centroid* yang tidak tepat dapat menyebabkan algoritma terjebak dalam solusi lokal, yang pada akhirnya memengaruhi hasil pengelompokan.

Sebagai contoh, pada dataset Iris, yang berisi 150 sampel bunga dari tiga spesies berbeda (Iris setosa, Iris versicolor, dan Iris virginica) dengan empat fitur utama (panjang dan lebar kelopak serta panjang dan lebar daun), pemilihan *centroid* yang tidak tepat dapat menyebabkan kesalahan dalam pengelompokan spesies. Misalnya, algoritma *K-Means* standar mungkin akan mengelompokkan Iris versicolor dan Iris virginica dalam satu *cluster* meskipun keduanya memiliki perbedaan yang cukup signifikan dalam karakteristik fitur mereka. Hal ini bisa terjadi karena keduanya memiliki kesamaan dalam panjang kelopak dan lebar kelopak, yang membuat algoritma sulit membedakan keduanya tanpa adanya pemilihan *centroid* yang optimal.

Begitu pula pada dataset Wine, yang terdiri dari 13 fitur kimiawi tentang karakteristik tiga jenis anggur (jenis 1, jenis 2, dan jenis 3), pemilihan *centroid* yang acak juga dapat menyebabkan masalah dalam pengelompokan. Dimensi data yang tinggi pada dataset Wine menambah tantangan tersendiri, karena *K-Means* standar sering kali gagal membedakan dengan jelas antara *cluster* jenis anggur yang berbeda. Sebagai contoh, dalam kasus ini, jenis anggur dengan kandungan kimia yang serupa dapat dikelompokkan bersama-sama, meskipun memiliki karakteristik yang berbeda dalam variabel lainnya, menyebabkan ketidakakuratan dalam hasil *clustering*.

Untuk mengatasi masalah tersebut, algoritma PSO dapat diimplementasikan untuk mengoptimalkan pemilihan *centroid* pada *K-Means*. PSO merupakan algoritma optimasi populasi yang meniru cara hewan, seperti burung yang terbang dalam kawanan. Dalam PSO, setiap solusi yang mungkin disebut partikel, bergerak di dalam ruang pencarian untuk menemukan solusi terbaik dengan mempertimbangkan pengalaman terbaik mereka sendiri dan seluruh kelompok. Dalam konteks *K-Means*, PSO membantu mencari posisi *centroid* yang optimal dengan cara meminimalkan jarak antar data dan *centroid* yang lebih relevan sehingga pengelompokan data bisa lebih akurat dan mencegah algoritma *K-Means* terjebak dalam solusi yang kurang optimal.

Algoritma *K-Means* yang dikombinasikan dengan PSO bekerja dengan cara menggabungkan kelebihan PSO dalam mencari solusi terbaik secara menyeluruh, dan kecepatan *K-Means* dalam membentuk kelompok data. Gabungan ini membuat hasil pengelompokan data menjadi lebih akurat dibandingkan dengan metode-metode sebelumnya (Dey & Chakraborty, 2014). Dengan demikian, penggunaan *K-Means* yang dioptimalkan dengan PSO pada kedua dataset ini tidak hanya meningkatkan akurasi *clustering*, tetapi juga menunjukkan keunggulan dari pendekatan gabungan ini dalam mengatasi masalah pengelompokan data yang kompleks dan realistis. Melalui penerapan PSO, dapat diperoleh hasil yang lebih optimal dan relevan, yang tentunya lebih mendekati solusi global, daripada terjebak pada solusi lokal seperti yang sering terjadi pada *K-Means* standar.

BAB III

METODE PENELITIAN

3.1 Jenis Penelitian

Metode yang digunakan dalam penelitian ini adalah kualitatif, untuk menganalisis bagaimana PSO dapat mengoptimalkan *K-Means Clustering* pada pengelompokan data Iris dan data Wine.

3.2 Data dan Sumber Data

Penelitian ini menggunakan dataset Iris dan dataset Wine yang tersedia di *UCI Machine Learning Repository*. Dataset Iris memiliki 150 sampel dengan 4 fitur: panjang petal, lebar petal, panjang sepal, dan lebar sepal. Dataset ini juga terbagi menjadi tiga kelas: Iris Setosa, Iris Versicolor, dan Iris Virginica, dengan masing-masing kelas terdiri dari 50 sampel. Sementara, dataset Wine memiliki tiga jenis anggur yang berbeda dan memiliki 13 fitur sampel yaitu alkohol, asam sulfat, abu, magnesium, alkalinitas abu, jumlah fenol, flavanoid, fenol nonflavanoid, proantosianidin, intensitas warna, warna, OD280/OD315 dari anggur yang diencerkan, dan prolin.

Kedua dataset yang digunakan dalam penelitian ini tersedia di Google Drive melalui tautan berikut: <https://tinyurl.com/datasetpenelitian>.

3.3 Struktur Data

Secara umum, struktur data untuk dataset yang digunakan dalam penelitian ini terdiri dari m dan q . Dataset iris terdiri dari $m = 150$ baris dan $q = 4$. Sementara dataset Wine terdiri dari $m = 178$ baris dan $q = 13$.

Tabel 3.1 Dataset Penelitian

Dataset	Total Observasi (m)	Fitur (q)	Atribut Kelas (K)
Iris	150	4	3
Wine	178	13	3

3.4 Tahapan Penelitian

3.4.1 K-Means Clustering

Langkah-langkah algoritma K-Means adalah sebagai berikut:

1. Menentukan jumlah *cluster*:

Dalam penelitian ini, karena dataset memiliki 3 kelas yang berbeda, maka jumlah *cluster* yang akan digunakan adalah $K = 3$.

2. Inisialisasi *centroid*:

Algoritma ini dimulai dengan menginisialisasi tiga *centroid* secara acak dari X , di mana $X = (\bar{x}_1, \bar{x}_2, \bar{x}_3, \dots, \bar{x}_m)$. $\bar{x}_i = (x_{i1}, x_{i2}, \dots, x_{iq})$ dengan $i = 1, 2, \dots, m$.

Centroid awal ini dinotasikan sebagai:

$$\mu = (\bar{\mu}_1, \bar{\mu}_2, \dots, \bar{\mu}_K)$$

di mana $\bar{\mu}_j \in X$ adalah *centroid* ke- j pada iterasi awal dan $j = 1, 2, \dots, K$.

3. Perhitungan jarak setiap data terhadap *centroid* dilakukan pada setiap cluster:

Untuk setiap data $\bar{x}_i$ (diambil dari X), dapat dihitung menggunakan jarak *Euclidean* ke setiap *centroid* $\bar{\mu}_j$ yang dirumuskan sebagai berikut:

$$d(\bar{x}_i, \bar{\mu}_j) = \|\bar{x}_i - \bar{\mu}_j\| \quad (3.1)$$

di mana:

- $d(\bar{x}_i, \bar{\mu}_j)$ = Jarak antara titik data ke- i dan *centroid* ke- j ,
- q = Jumlah fitur yang akan diamati,
- $\bar{x}_i$ = Data ke- i ,
- $\bar{\mu}_j$ = *Centroid* ke- j .

4. Data dikelompokkan ke dalam *cluster* berdasarkan jarak terpendek.

Setelah menghitung $d(\bar{x}_i, \bar{\mu}_j)$, langkah selanjutnya adalah mengelompokkan data $\bar{x}_i$ ke *cluster* dengan jarak terpendek. Dengan kata lain, mencari *centroid* $\bar{\mu}_j$ yang memberikan nilai jarak terkecil bagi data $\bar{x}_i$.

$$C_j = \{\bar{x}_i \mid d(\bar{x}_i, \bar{\mu}_j) = \min_j d(\bar{x}_i, \bar{\mu}_j)\} \quad (3.2)$$

di mana:

- C_j yaitu *cluster* ke- j , yaitu himpunan semua data $\bar{x}_i$ yang dikelompokkan ke dalam *cluster* C_j .
- $\bar{x}_i$ yaitu vektor data ke- i .
- $d(\bar{x}_i, \bar{\mu}_j)$ yaitu jarak *Euclidean* antara data $\bar{x}_i$ dan *centroid* $\bar{\mu}_j$ dari *cluster* C_j .
- $\min_j d(\bar{x}_i, \bar{\mu}_j)$ yaitu nilai jarak minimum antara data $\bar{x}_i$ dan semua *centroid* $\bar{\mu}_j$, (*centroid* dari semua *cluster* j').

5. Perhitungan *centroid* berikutnya atau rata-rata tiap *cluster*:

$$\bar{\mu}_j = \frac{1}{n_j} \sum_{\bar{x}_i \in C_j} \bar{x}_i \quad (3.3)$$

di mana:

- $\bar{\mu}_j$ = *Centroid* untuk *cluster* ke- j pada iterasi pertama,
- n_j = Jumlah titik data yang berada dalam *cluster* C_j ,
- $\sum_{\bar{x}_i \in C_j} \bar{x}_i$ = Jumlah semua titik data $\bar{x}_i$ yang berada dalam *cluster* C_j ,
- $\bar{x}_i$ = Titik data ke- i yang termasuk dalam *cluster* C_j .

6. Lakukan langkah-langkah 3 sampai 5 diulang hingga *centroid* tidak lagi berubah secara signifikan. Iterasi akan dihentikan jika perhitungan *centroid* pada iterasi terakhir tidak berbeda dari iterasi sebelumnya.

3.4.2 Particle Swarm Optimization (PSO)

PSO digunakan untuk mengoptimalkan posisi *centroid* yang diperoleh dari *K-Means*. Dalam PSO, partikel merepresentasikan posisi *centroid* dan dioptimalkan melalui iterasi. Langkah-langkah PSO:

1. Inisialisasi Partikel

a. Definisi Partikel

Setiap partikel dalam PSO mewakili sekumpulan *centroid*:

$$\mu = (\bar{\mu}_1, \bar{\mu}_2, \dots, \bar{\mu}_K)$$

dengan K adalah jumlah *cluster*. Jadi, setiap partikel adalah sebuah vektor dengan dimensi $K \times q$, dimana q adalah jumlah fitur.

b. Inisialisasi Posisi Partikel

Posisi awal partikel $\bar{p}_i$ (yang mewakili sekumpulan *centroid*) diinisialisasi secara acak dalam ruang fitur X :

$$\bar{p}_i = (p_{i1}, p_{i2}, p_{i3}, \dots, p_{il})$$

dengan $p_{il} \in X$.

c. Inisialisasi Kecepatan Partikel

Kecepatan awal dari setiap partikel $\bar{v}_i$, juga diinisialisasikan secara acak.

$$\bar{v}_i = (v_{i1}, v_{i2}, \dots, v_{il})$$

dengan $v_{il} \in X$.

d. Parameter PSO

Parameter PSO yang diinisialisasi meliputi:

- ω adalah koefisien inersia yang mengontrol pengaruh kecepatan sebelumnya.
- a_1 dan a_2 adalah konstanta posistif bobot yang mengatur pengalaman terbaik partikel sendiri (*local best*) dan pengalaman terbaik kawanan (*global best*).
- r_1 dan r_2 adalah nilai acak dalam interval $[0,1]$.

2. Perhitungan Fungsi Objektif (SSE)

Misalkan fungsi objektif $f(\mu)$ untuk *centroid* μ didefinisikan sebagai:

$$f(\mu) = SSE = \sum_{j=1}^K \sum_{x_i \in C_j} \sum_{l=1}^q (x_{il} - \mu_{jl})^2 \quad (3.4)$$

di mana:

- $f(\mu)$ adalah nilai fungsi tujuan untuk posisi *centroid* μ
- K adalah jumlah *cluster*
- C_j adalah *cluster* ke- j yang dibentuk oleh *centroid* μ_j
- $\bar{x}_i$ adalah vektor data ke- i
- $(x_{il} - \mu_{jl})^2$ adalah kuadrat dari selisih antara nilai fitur l dari titik data ke- i dan nilai fitur l dari *centroid* ke- j .

3. *Update Kecepatan dan Posisi*

a. *Update Kecepatan*

Setiap partikel memperbarui kecepatan pengalaman terbaik individu dan kelompok:

$$\bar{v}_{il}^{t+1} = \omega \bar{v}_{il}^t + a_1 r_1 (\bar{p}_{best_{il}} - \bar{p}_{il}^t) + a_2 r_2 (\bar{g}_{best_{il}} - \bar{p}_{il}^t) \quad (3.5)$$

di mana:

- r_1 dan r_2 adalah nilai acak dalam interval $[0,1]$
- a_1 dan a_2 adalah konstanta persistif bobot yang mengatur pengalaman terbaik partikel sendiri (*local best*) dan pengalaman terbaik kawan (*global best*)
- $\bar{p}_{best}$ adalah posisi terbaik yang pernah dicapai oleh partikel i
- $\bar{g}_{best}$ adalah posisi terbaik yang pernah dicapai oleh seluruh populasi.

b. *Update Posisi*

Setiap partikel memperbarui posisi *centroid*-nya:

$$\bar{p}_{il}^{t+1} = \bar{p}_{il}^t + \bar{v}_{il}^{t+1} \quad (3.6)$$

di mana $\bar{p}_i^{t+1}$ adalah posisi baru partikel i pada iterasi $t + 1$.

4. *Update Best Position*

Jika posisi baru $\bar{p}_i^{t+1}$ memberikan nilai fungsi tujuan yang lebih baik dari $\bar{p}_{best}$, maka *update* $\bar{p}_{best}$.

Jika ada posisi yang memberikan nilai fungsi tujuan terbaik secara keseluruhan, maka *update* $\bar{g}_{best}$.

5. Lakukan langkah 3 sampai 4 hingga jumlah iterasi maksimum tercapai atau fungsi tujuan tidak berubah secara signifikan.

6. Setelah proses iterasi selesai, posisi *centroid* terakhir dari partikel $\bar{g}_{best}$ dapat digunakan sebagai *centroid* yang dioptimalkan untuk *K-Means*.

3.4.3 Penggabungan *K-Means* dan PSO

Pada penelitian ini, PSO digunakan untuk mengoptimalkan *K-Means* dalam menentukan posisi pusat *cluster* (*centroid*) yang optimal. Langkah-langkah integrasi *K-Means* dan PSO dalam penelitian ini adalah sebagai berikut:

1. Inisialisasi awal dengan *K-Means*
 - a. Menentukan jumlah *cluster* K
 - b. Inisialisasi *centroid* secara acak tiga *centroid* awal dari data X

$$\mu = (\bar{\mu}_1, \bar{\mu}_2, \dots, \bar{\mu}_K)$$

di mana $\bar{\mu}_j \in X$ adalah *centroid* ke- j pada iterasi awal dan $j = 1, 2, \dots, K$.

- c. Perhitungan jarak setiap data terhadap *centroid* dilakukan pada setiap cluster menggunakan persamaan (3.1).
- d. Data dikelompokkan ke dalam cluster berdasarkan jarak terpendek.
- e. Perhitungan *centroid* berikutnya atau rata-rata tiap *cluster* menggunakan persamaan (3.3).
- f. Ulangi langkah di atas hingga *centroid* tidak lagi berubah secara signifikan.

Setelah beberapa iterasi *K-Means*, diperoleh *centroid* awal yang akan digunakan sebagai posisi awal untuk PSO.

2. Inisialisasi partikel PSO berdasarkan *centroid K-Means*

Setiap partikel dalam PSO mewakili sekumpulan *centroid* yang didapat dari *K-Means*:

Posisi awal partikel

$$\bar{p}_i = (p_{i1}, p_{i2}, p_{i3}, \dots, p_{il})$$

dengan $p_{il} \in X$.

Kecepatan awal partikel

$$\bar{v}_i = (v_{i1}, v_{i2}, \dots, v_{il})$$

dengan $v_{il} \in X$.

3. Perhitungan fungsi objektif untuk setiap partikel menggunakan persamaan (3.4).

4. *Update* kecepatan dan posisi dengan partikel PSO

Setiap partikel memperbarui kecepatan dan posisi *centroid* berdasarkan persamaan (3.5) dan persamaan (3.6).

5. Perbandingan dan *update* nilai terbaik ($\bar{p}_{best}$ dan $\bar{g}_{best}$)

Jika posisi baru $\bar{p}_i^{t+1}$ memberikan nilai fungsi objektif yang lebih kecil dari nilai terbaik sebelumnya $\bar{p}_{best}$, maka *update* $\bar{p}_{best}$ untuk partikel tersebut.

Jika nilai $\bar{p}_i^{t+1}$ memberikan nilai fungsi objektif global terbaik di antara seluruh partikel, maka *update* $\bar{g}_{best}$.

6. Iterasi hingga konvergensi atau iterasi maksimum

Langkah-langkah di atas akan di ulang hingga iterasi maksimum tercapai atau ketika perubahan posisi *centroid* tidak lagi signifikan.

7. Hasil Akhir

Setelah PSO mencapai konvergensi, posisi terbaik partikel yang mewakili *centroid* yang dioptimalkan digunakan untuk membagi data ke dalam *cluster* yang lebih baik dibandingkan dengan metode *K-Means* murni. *Centroid* hasil optimasi ini kemudian digunakan sebagai pusat *cluster* final untuk melakukan pengelompokan data pada dataset.

3.5 Evaluasi dan Validasi Hasil

Setelah melakukan pengelompokan data menggunakan *K-Means* yang dioptimalkan dengan PSO, hasil pengelompokan dievaluasi dan divalidasi menggunakan metrik dua metrik utama, yaitu *Sum of Squared Errors* (SSE), *Silhouette Score*, dan akurasi. Evaluasi ini bertujuan untuk menilai seberapa baik kualitas kluster yang terbentuk dibandingkan dengan metode *K-Means* standar serta penelitian sebelumnya.

1. *Sum of Squared Errors* (SSE)

SSE mengukur total variasi dalam setiap kluster dengan menghitung jumlah kuadrat jarak antara setiap titik data dengan *centroid* klasternya. Nilai SSE yang lebih kecil menunjukkan bahwa kluster lebih kompak dan data lebih dekat dengan *centroid*nya. Rumusnya adalah:

$$SSE = \sum_{j=1}^K \sum_{x_i \in C_j} \sum_{l=1}^q (x_{il} - \mu_{jl})^2 \quad (3.7)$$

dimana:

- x_{il} adalah nilai fitur ke- l dari titik data ke- i ,
- μ_{jl} adalah nilai fitur ke- l dari *centroid* ke- j ,

- K adalah jumlah klaster.

2. *Silhouette Score*

Silhouette Score digunakan untuk mengukur kualitas klaster dengan mengevaluasi seberapa baik suatu objek berada dalam klaster yang telah ditetapkan dibandingkan dengan klaster lainnya (Rousseeuw, 1987). Nilai ini menunjukkan sejauh mana sebuah objek cocok dengan klasternya sendiri dan seberapa jauh objek tersebut dari klaster lain.

Rumusnya adalah sebagai berikut:

$$S(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (3.8)$$

dengan:

- $a(i)$ adalah rata-rata disimilasi (ketidaksamaan) antara objek i dengan seluruh objek lainnya yang berada dalam klaster yang sama. Menurut Wikipedia, secara matematis dirumuskan sebagai berikut:

Untuk titik data $i \in C_I$, di mana C_I adalah klaster tempat titik data i berada, dengan jumlah anggota $|C_I|$.

Maka $a(i)$ didefinisikan sebagai:

$$a(i) = \frac{1}{|C_I| - 1} \sum_{j \in C_I, i \neq j} d(i, j),$$

merupakan rata-rata jarak dari data ke- i ke semua anggota lain dalam klasternya sendiri.

Kemudian, rata-rata jarak dari titik i ke klaster lain C_J ($C_J \neq C_I$)

$$d(i, C_J) = \frac{1}{|C_J|} \sum_{j \in C_J} d(i, j)$$

- $b(i)$ adalah rata-rata disimilasi (ketidaksamaan) dari objek i terhadap semua objek ke dalam klaster yang berbeda. Menurut Wikipedia, secara matematis dirumuskan sebagai berikut:

$$b(i) = \min_{j \neq i} \frac{1}{|C_j|} \sum_{j \in C_j} d(i, j),$$

merupakan nilai minimum dari rata-rata jarak dari data ke- i ke anggota klaster lain yang berbeda.

3. Akurasi (Accuracy)

Untuk menghitung akurasi, dapat menggunakan rumus berikut:

$$Akurasi = \frac{Jumlah\ Prediksi\ Benar}{Total\ Data} \times 100\% \quad (3.9)$$

4. Perbandingan dengan metode dasar dan penelitian sebelumnya

Untuk menilai efektivitas penggunaan algoritma *K-Means* yang dioptimalkan dengan PSO, dilakukan perbandingan dengan metode dasar dan penelitian sebelumnya.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Deskripsi Dataset Penelitian

4.1.1 Dataset Iris

Dataset Iris merupakan kumpulan data yang berisi 150 sampel data bunga yang terbagi ke dalam tiga kelas: Iris setosa, Iris versicolor, dan Iris virginica. Setiap kelas memiliki 50 sampel. Setiap data memiliki empat fitur yaitu panjang sepal, lebar sepal, panjang petal dan lebar petal.

Subset data untuk perhitungan manual (menggunakan 5 data pertama) yang dinyatakan pada tabel 4.1.

Tabel 4.1 Lima Data Pertama Dataset Iris

Data	<i>Sepal Length</i> (cm)	<i>Sepal Width</i> (cm)	<i>Petal Length</i> (cm)	<i>Petal Width</i> (cm)
1	5.1	3.5	1.4	0.2
2	4.9	3.0	1.4	0.2
3	4.7	3.2	1.3	0.2
4	4.6	3.1	1.5	0.2
5	5.0	3.6	1.4	0.2

4.1.2 Dataset Wine

Dataset Wine merupakan kumpulan data yang berasal dari tiga jenis berbeda. Dataset ini terdiri dari 178 sampel data anggur dengan 13 fitur kimia, seperti alkohol, asam sulfat, abu, magnesium, alkalinitas abu, jumlah fenol, flavanoid, fenol nonflavanoid, proantosianidin, intensitas warna, warna, OD280/OD315 dari anggur yang diencerkan, dan prolin.

Subset data untuk perhitungan manual (menggunakan 5 data pertama dan 4 fitur) dinyatakan pada tabel 4.2.

Tabel 4.2 Lima Data Pertama Dataset Wine

Data	<i>Alkohol</i>	<i>Malic Acid</i>	<i>Ash</i>	<i>Alcanitiy of Ash</i>
1	14.23	1.71	2.43	15.6
2	13.2	1.78	2.14	11.2
3	13.16	2.36	2.67	18.6
4	14.37	1.95	2.5	16.8
5	13.24	2.59	2.87	21

4.2 Implementasi *K-Means Clustering*

4.2.1 Implementasi *K-Means* Manual pada Dataset Iris (5 Data Pertama)

Berdasarkan tabel 4.1 subset data penelitian pada data Iris terdapat lima data pertama yang di notasikan sebagai berikut:

$\bar{x}_i = (x_{i1}, x_{i2}, \dots, x_{iq})$, q adalah jumlah fitur, maka:

$$\bar{x}_1 = (x_{11}, x_{12}, x_{13}, x_{14}) = (5.1, 3.5, 1.4, 0.2)$$

$$\bar{x}_2 = (x_{21}, x_{22}, x_{23}, x_{24}) = (4.9, 3.0, 1.4, 0.2)$$

$$\bar{x}_3 = (x_{31}, x_{32}, x_{33}, x_{34}) = (4.7, 3.2, 1.3, 0.2)$$

$$\bar{x}_4 = (x_{41}, x_{42}, x_{43}, x_{44}) = (4.6, 3.1, 1.5, 0.2)$$

$$\bar{x}_5 = (x_{51}, x_{52}, x_{53}, x_{54}) = (5, 3.6, 1.4, 0.2)$$

Berikut merupakan perhitungan manual beserta langkahnya:

Iterasi 1

1. Menentukan jumlah *cluster*

Karena pada dataset Iris memiliki 3 kelas yaitu Iris setosa, Iris versicolor, dan Iris virginica, maka:

$$K = 3$$

2. Inisialisasi *centroid*

Penentuan *centroid*/pusat awal *cluster* didapatkan secara acak/*random* dari himpunan data yang tersedia.

Dalam kasus ini, indeks data yang digunakan sebagai berikut:

Indeks data = {1,2,3,4,5}.

Untuk memilih tiga *centroid* secara acak dari indeks data dapat menggunakan fungsi `random` dalam `python`. Berikut contoh implementasinya:

```
import random
angka = [1,2,3,4,5]
centroid = random.sample(angka,3)
print(centroid)
```

Kode di atas menghasilkan tiga indeks acak dari himpunan data, yang digunakan sebagai acuan untuk menentukan pusat *cluster* awal. Karena bersifat acak, hasilnya dapat berbeda setiap kali dijalankan. Pada percobaan ini dilakukan tiga kali, dengan hasil: percobaan pertama {3, 2, 5}, kedua {5, 2, 4}, dan ketiga {1, 3, 5}. Penulis menggunakan hasil percobaan ketiga sebagai inisialisasi *centroid*.

Berikut merupakan *centroid*-nya:

$$\bar{\mu}_1 = (5.1, 3.5, 1.4, 0.2)$$

$$\bar{\mu}_2 = (4.7, 3.2, 1.3, 0.2)$$

$$\bar{\mu}_3 = (5, 3.6, 1.4, 0.2)$$

3. Perhitungan jarak *Euclidean*

Untuk mengukur jarak antara data dengan *centroid* digunakan jarak *Euclidean distance* yaitu dengan rumus pada persamaan (3.1) berikut:

$$d(\bar{x}_i, \bar{\mu}_j) = \|\bar{x}_i - \bar{\mu}_j\|$$

dimana $\bar{x}_i$ merupakan data ke- i dan $\bar{\mu}_j$ merupakan *centroid* ke- j .

Kemudian perhitungan jarak *Euclidean* dimulai antara data pertama ($\bar{x}_1$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$\begin{aligned}
d(\bar{x}_1, \bar{\mu}_1) &= \|(\bar{x}_{11} - \bar{\mu}_{11}, \bar{x}_{12} - \bar{\mu}_{12}, \bar{x}_{13} - \bar{\mu}_{13}, \bar{x}_{14} - \bar{\mu}_{14})\| \\
&= \|(5.1 - 5.1, 3.5 - 3.5, 1.4 - 1.4, 0.2 - 0.2)\| \\
&= \|(0, 0, 0, 0)\| \\
&= \sqrt{(0)^2 + (0)^2 + (0)^2 + (0)^2} = 0
\end{aligned}$$

$$\begin{aligned}
d(\bar{x}_1, \bar{\mu}_2) &= \|(\bar{x}_{11} - \bar{\mu}_{21}, \bar{x}_{12} - \bar{\mu}_{22}, \bar{x}_{13} - \bar{\mu}_{23}, \bar{x}_{14} - \bar{\mu}_{24})\| \\
&= \|(5.1 - 4.7, 3.5 - 3.2, 1.4 - 1.3, 0.2 - 0.2)\| \\
&= \|(0.4, 0.3, 0.1, 0)\| \\
&= \sqrt{(0.4)^2 + (0.3)^2 + (0.1)^2 + (0)^2} = 0.509
\end{aligned}$$

$$\begin{aligned}
d(\bar{x}_1, \bar{\mu}_3) &= \|(\bar{x}_{11} - \bar{\mu}_{31}, \bar{x}_{12} - \bar{\mu}_{32}, \bar{x}_{13} - \bar{\mu}_{33}, \bar{x}_{14} - \bar{\mu}_{34})\| \\
&= \|(5.1 - 5, 3.5 - 3.6, 1.4 - 1.4, 0.2 - 0.2)\| \\
&= \|(0.1, -0.1, 0, 0)\| \\
&= \sqrt{(0.1)^2 + (-0.1)^2 + (0)^2 + (0)^2} = 0.141
\end{aligned}$$

Kemudian dilanjut dengan perhitungan jarak *Euclidean* antara data kedua ($\bar{x}_2$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$\begin{aligned}
d(\bar{x}_2, \bar{\mu}_1) &= \|(\bar{x}_{21} - \bar{\mu}_{11}, \bar{x}_{22} - \bar{\mu}_{12}, \bar{x}_{23} - \bar{\mu}_{13}, \bar{x}_{24} - \bar{\mu}_{14})\| \\
&= \|(4.9 - 5.1, 3.0 - 3.5, 1.4 - 1.4, 0.2 - 0.2)\| \\
&= \|(-0.2, -0.5, 0, 0)\| \\
&= \sqrt{(-0.2)^2 + (-0.5)^2 + (0)^2 + (0)^2} = 0.538
\end{aligned}$$

$$\begin{aligned}
d(\bar{x}_2, \bar{\mu}_2) &= \|(\bar{x}_{21} - \bar{\mu}_{21}, \bar{x}_{22} - \bar{\mu}_{22}, \bar{x}_{23} - \bar{\mu}_{23}, \bar{x}_{24} - \bar{\mu}_{24})\| \\
&= \|(4.9 - 4.7, 3.0 - 3.2, 1.4 - 1.3, 0.2 - 0.2)\| \\
&= \|(0.2, -0.2, 0.1, 0)\| \\
&= \sqrt{(0.2)^2 + (-0.2)^2 + (0.1)^2 + (0)^2} = 0.3
\end{aligned}$$

$$d(\bar{x}_2, \bar{\mu}_3) = \|(\bar{x}_{21} - \bar{\mu}_{31}, \bar{x}_{22} - \bar{\mu}_{32}, \bar{x}_{23} - \bar{\mu}_{33}, \bar{x}_{24} - \bar{\mu}_{34})\|$$

$$\begin{aligned}
&= \|(4.9 - 5, 3.0 - 3.6, 1.4 - 1.4, 0.2 - 0.2)\| \\
&= \|(-0.1, -0.6, 0, 0)\| \\
&= \sqrt{(-0.1)^2 + (-0.6)^2 + (0)^2 + (0)^2} = 0.608
\end{aligned}$$

Kemudian dilanjut dengan perhitungan jarak *Euclidean* antara data ketiga ($\bar{x}_3$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$\begin{aligned}
d(\bar{x}_3, \bar{\mu}_1) &= \|(\bar{x}_{31} - \bar{\mu}_{11}, \bar{x}_{32} - \bar{\mu}_{12}, \bar{x}_{33} - \bar{\mu}_{13}, \bar{x}_{34} - \bar{\mu}_{14})\| \\
&= \|(4.7 - 5.1, 3.2 - 3.5, 1.3 - 1.4, 0.2 - 0.2)\| \\
&= \|(-0.4, -0.3, -0.1, 0)\| \\
&= \sqrt{(-0.4)^2 + (-0.3)^2 + (-0.1)^2 + (0)^2} = 0.509
\end{aligned}$$

$$\begin{aligned}
d(\bar{x}_3, \bar{\mu}_2) &= \|(\bar{x}_{31} - \bar{\mu}_{21}, \bar{x}_{32} - \bar{\mu}_{22}, \bar{x}_{33} - \bar{\mu}_{23}, \bar{x}_{34} - \bar{\mu}_{24})\| \\
&= \|(4.7 - 4.7, 3.2 - 3.2, 1.3 - 1.3, 0.2 - 0.2)\| \\
&= \|(0, 0, 0, 0)\| \\
&= \sqrt{(0)^2 + (0)^2 + (0)^2 + (0)^2} = 0
\end{aligned}$$

$$\begin{aligned}
d(\bar{x}_3, \bar{\mu}_3) &= \|(\bar{x}_{31} - \bar{\mu}_{31}, \bar{x}_{32} - \bar{\mu}_{32}, \bar{x}_{33} - \bar{\mu}_{33}, \bar{x}_{34} - \bar{\mu}_{34})\| \\
&= \|(4.7 - 5, 3.2 - 3.6, 1.3 - 1.4, 0.2 - 0.2)\| \\
&= \|(-0.3, -0.4, -0.1, 0)\| \\
&= \sqrt{(-0.3)^2 + (-0.4)^2 + (-0.1)^2 + (0)^2} = 0.509
\end{aligned}$$

Kemudian dilanjut dengan perhitungan jarak *Euclidean* antara data keempat ($\bar{x}_4$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$\begin{aligned}
d(\bar{x}_4, \bar{\mu}_1) &= \|(\bar{x}_{41} - \bar{\mu}_{11}, \bar{x}_{42} - \bar{\mu}_{12}, \bar{x}_{43} - \bar{\mu}_{13}, \bar{x}_{44} - \bar{\mu}_{14})\| \\
&= \|(4.6 - 5.1, 3.1 - 3.5, 1.5 - 1.4, 0.2 - 0.2)\| \\
&= \|(-0.4, -0.3, -0.1, 0)\| \\
&= \sqrt{(-0.5)^2 + (-0.4)^2 + (0.1)^2 + (0)^2} = 0.648
\end{aligned}$$

$$\begin{aligned}
d(\bar{x}_4, \bar{\mu}_2) &= \|(\bar{x}_{41} - \bar{\mu}_{21}, \bar{x}_{42} - \bar{\mu}_{22}, \bar{x}_{43} - \bar{\mu}_{23}, \bar{x}_{44} - \bar{\mu}_{24})\| \\
&= \|(4.6 - 4.7, 3.1 - 3.2, 1.5 - 1.3, 0.2 - 0.2)\| \\
&= \|(-0.1, -0.1, 0.2, 0)\| \\
&= \sqrt{(-0.1)^2 + (-0.1)^2 + (0.2)^2 + (0)^2} = 0.244 \\
d(\bar{x}_4, \bar{\mu}_3) &= \|(\bar{x}_{41} - \bar{\mu}_{31}, \bar{x}_{42} - \bar{\mu}_{32}, \bar{x}_{43} - \bar{\mu}_{33}, \bar{x}_{44} - \bar{\mu}_{34})\| \\
&= \|(4.6 - 5, 3.1 - 3.6, 1.5 - 1.4, 0.2 - 0.2)\| \\
&= \|(-0.4, -0.5, 0.1, 0)\| \\
&= \sqrt{(-0.4)^2 + (-0.5)^2 + (0.1)^2 + (0)^2} = 0.648
\end{aligned}$$

Terakhir dilanjut dengan perhitungan jarak *Euclidean* antara data kelima ($\bar{x}_5$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$\begin{aligned}
d(\bar{x}_5, \bar{\mu}_1) &= \|(\bar{x}_{51} - \bar{\mu}_{11}, \bar{x}_{52} - \bar{\mu}_{12}, \bar{x}_{53} - \bar{\mu}_{13}, \bar{x}_{54} - \bar{\mu}_{14})\| \\
&= \|(5 - 5.1, 3.6 - 3.5, 1.4 - 1.4, 0.2 - 0.2)\| \\
&= \|(-0.1, 0.1, 0, 0)\| \\
&= \sqrt{(-0.1)^2 + (0.1)^2 + (0)^2 + (0)^2} = 0.141 \\
d(\bar{x}_5, \bar{\mu}_2) &= \|(\bar{x}_{51} - \bar{\mu}_{21}, \bar{x}_{52} - \bar{\mu}_{22}, \bar{x}_{53} - \bar{\mu}_{23}, \bar{x}_{54} - \bar{\mu}_{24})\| \\
&= \|(5 - 4.7, 3.6 - 3.2, 1.4 - 1.3, 0.2 - 0.2)\| \\
&= \|(0.3, 0.4, 0.1, 0)\| \\
&= \sqrt{(0.3)^2 + (0.4)^2 + (0.1)^2 + (0)^2} = 0.509 \\
d(\bar{x}_5, \bar{\mu}_3) &= \|(\bar{x}_{51} - \bar{\mu}_{31}, \bar{x}_{52} - \bar{\mu}_{32}, \bar{x}_{53} - \bar{\mu}_{33}, \bar{x}_{54} - \bar{\mu}_{34})\| \\
&= \|(5 - 5, 3.6 - 3.6, 1.4 - 1.4, 0.2 - 0.2)\| \\
&= \|(0, 0, 0, 0)\| \\
&= \sqrt{(0)^2 + (0)^2 + (0)^2 + (0)^2} = 0
\end{aligned}$$

4. Pengelompokan data sesuai *cluster* berdasarkan jarak terpendek

Setelah menghitung jarak *Euclidean* langkah selanjutnya yaitu mengelompokkan data sesuai *cluster* berdasarkan jarak terpendek.

Tabel 4.3 Pengelompokan *Cluster* pada Iterasi 1

$\bar{x}_i$	$d(\bar{\mu}_1)$	$d(\bar{\mu}_2)$	$d(\bar{\mu}_3)$	<i>Cluster</i>
$\bar{x}_1$	0	0.509	0.141	1
$\bar{x}_2$	0.538	0.3	0.608	2
$\bar{x}_3$	0.509	0	0.509	2
$\bar{x}_4$	0.648	0.244	0.648	2
$\bar{x}_5$	0.141	0.509	0	3

5. Perhitungan *centroid* berikutnya atau rata-rata tiap *cluster*

Berdasarkan tabel 4.3 menunjukkan hasil pengelompokan *cluster* pada iterasi pertama. Kemudian langkah selanjutnya adalah memperbarui *centroid* untuk setiap *cluster*. Proses ini dilakukan dengan menghitung rata-rata menggunakan persamaan (3.3) dari data yang termasuk dalam setiap *cluster*.

$\bar{\mu}_1 = \bar{x}_1$, karena $\bar{x}_1$ termasuk ke dalam *cluster* 1.

$\bar{\mu}_2 = \bar{x}_2, \bar{x}_3, \bar{x}_4$, karena $\bar{x}_2, \bar{x}_3, \bar{x}_4$ termasuk ke dalam *cluster* 2.

$\bar{\mu}_3 = \bar{x}_5$, karena $\bar{x}_5$ termasuk ke dalam *cluster* 3.

Sehingga rata-rata untuk nilai *centroid* baru adalah

$$\bar{\mu}_1 = \frac{(5.1, 3.5, 1.4, 0.2)}{1} = (5.1, 3.5, 1.4, 0.2)$$

$$\bar{\mu}_2 = \frac{(4.9, 3.0, 1.4, 0.2) + (4.7, 3.2, 1.3, 0.2) + (4.6, 3.1, 1.5, 0.2)}{3}$$

$$= (4.7, 3.1, 1.4, 0.2)$$

$$\bar{\mu}_3 = \frac{(5, 3.6, 1.4, 0.2)}{1} = (5, 3.6, 1.4, 0.2)$$

Iterasi 2

1. Menggunakan *centroid* baru yang didapatkan dari iterasi pertama

$$\bar{\mu}_1 = (5.1, 3.5, 1.4, 0.2)$$

$$\bar{\mu}_2 = (4.7, 3.1, 1.4, 0.2)$$

$$\bar{\mu}_3 = (5, 3.6, 1.4, 0.2)$$

2. Perhitungan jarak *Euclidean* menggunakan persamaan (3.1)

Dimulai dengan perhitungan jarak *Euclidean* antara data pertama ($\bar{x}_1$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_1, \bar{\mu}_1) = \|\bar{x}_1 - \bar{\mu}_1\| = 0$$

$$d(\bar{x}_1, \bar{\mu}_2) = \|\bar{x}_1 - \bar{\mu}_2\| = 0.565$$

$$d(\bar{x}_1, \bar{\mu}_3) = \|\bar{x}_1 - \bar{\mu}_3\| = 0.509$$

Kemudian dilanjut perhitungan jarak *Euclidean* dimulai antara data kedua ($\bar{x}_2$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_2, \bar{\mu}_1) = \|\bar{x}_2 - \bar{\mu}_1\| = 0.538$$

$$d(\bar{x}_2, \bar{\mu}_2) = \|\bar{x}_2 - \bar{\mu}_2\| = 0.223$$

$$d(\bar{x}_2, \bar{\mu}_3) = \|\bar{x}_2 - \bar{\mu}_3\| = 0.608$$

Kemudian dilanjut perhitungan jarak *Euclidean* dimulai antara data ketiga ($\bar{x}_3$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_3, \bar{\mu}_1) = \|\bar{x}_3 - \bar{\mu}_1\| = 0.509$$

$$d(\bar{x}_3, \bar{\mu}_2) = \|\bar{x}_3 - \bar{\mu}_2\| = 0.141$$

$$d(\bar{x}_3, \bar{\mu}_3) = \|\bar{x}_3 - \bar{\mu}_3\| = 0.509$$

Kemudian dilanjut perhitungan jarak *Euclidean* dimulai antara data keempat ($\bar{x}_4$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_4, \bar{\mu}_1) = \|\bar{x}_4 - \bar{\mu}_1\| = 0.648$$

$$d(\bar{x}_4, \bar{\mu}_2) = \|\bar{x}_{41} - \bar{\mu}_{21}\| = 0.141$$

$$d(\bar{x}_4, \bar{\mu}_3) = \|\bar{x}_{41} - \bar{\mu}_{31}\| = 0.648$$

Terakhir dilanjut dengan perhitungan jarak *Euclidean* antara data kelima ($\bar{x}_5$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_5, \bar{\mu}_1) = \|\bar{x}_5 - \bar{\mu}_1\| = 0.141$$

$$d(\bar{x}_5, \bar{\mu}_2) = \|\bar{x}_{51} - \bar{\mu}_{21}\| = 0.583$$

$$d(\bar{x}_5, \bar{\mu}_3) = \|\bar{x}_5 - \bar{\mu}_3\| = 0$$

3. Pengelompokan data sesuai *cluster* berdasarkan jarak terpendek

Setelah menghitung jarak *Euclidean* langkah selanjutnya yaitu mengelompokkan data sesuai *cluster* berdasarkan jarak terpendek.

Tabel 4.4 Pengelompokan *Cluster* pada Iterasi 2

$\bar{x}_i$	$d(\bar{\mu}_1)$	$d(\bar{\mu}_2)$	$d(\bar{\mu}_3)$	Cluster (C)
$\bar{x}_1$	0	0.565	0.509	1
$\bar{x}_2$	0.538	0.223	0.608	2
$\bar{x}_3$	0.509	0.141	0.509	2
$\bar{x}_4$	0.648	0.141	0.648	2
$\bar{x}_5$	0.141	0.583	0	3

Berdasarkan hasil perhitungan dalam tabel di atas maka diperoleh:

$$C_1 = \{\bar{x}_1\}$$

$$C_2 = \{\bar{x}_2, \bar{x}_3, \bar{x}_4\}$$

$$C_3 = \{\bar{x}_5\}$$

4. Perhitungan *centroid* berikutnya atau rata-rata tiap *cluster*

Berdasarkan tabel 4.4 menunjukkan hasil pengelompokan *cluster* pada iterasi kedua. Dimana hasil nya menunjukkan pengelompokan *cluster* yang sama pada iterasi pertama. Artinya perhitungan ini sudah konvergen pada iterasi kedua. Untuk mendapatkan *centroid* akhirnya dilakukan dengan menghitung rata-rata menggunakan persamaan (3.3) dari data yang termasuk dalam setiap *cluster*.

$$\bar{\mu}_1 = \frac{(5.1, 3.5, 1.4, 0.2)}{1} = (5.1, 3.5, 1.4, 0.2)$$

$$\bar{\mu}_2 = \frac{(4.9, 3.0, 1.4, 0.2) + (4.7, 3.2, 1.3, 0.2) + (4.6, 3.1, 1.5, 0.2)}{3}$$

$$= (4.7, 3.1, 1.4, 0.2)$$

$$\bar{\mu}_3 = \frac{(5, 3.6, 1.4, 0.2)}{1} = (5, 3.6, 1.4, 0.2)$$

Metrik Evaluasi:

1. SSE

Perhitungan ini menggunakan persamaan (3.7) berikut:

$$SSE = \sum_{j=1}^K \sum_{\bar{x}_i \in C_j} \sum_{l=1}^q (x_{il} - \mu_{jl})^2$$

Dengan $K = 3$ (jumlah *cluster*) dan $q = 4$ (jumlah fitur), maka perhitungan

SSE dilakukan sebagai berikut:

- Untuk $j = 1$, maka diperoleh $C_j = C_1 = \{\bar{x}_1\}$, sedemikian hingga $\{\bar{x}_i\}_{i=1}^1 = \{\bar{x}_1\}$, dengan $\bar{x}_1 = (5.1, 3.5, 1.4, 0.2)$ dan *centroid* $\bar{\mu}_1 = (5.1, 3.5, 1.4, 0.2)$, diperoleh:

$$\begin{aligned} SE_1 &= \sum_{l=1}^4 (x_{1l} - \mu_{1l})^2 \\ &= (5.1 - 5.1)^2 + (3.5 - 3.5)^2 + (1.4 - 1.4)^2 + (0.2 - 0.2)^2 = 0 \end{aligned}$$

- Untuk $j = 2$, maka diperoleh $C_j = C_2 = \{\bar{x}_2, \bar{x}_3, \bar{x}_4\}$, sedemikian hingga $\{\bar{x}_i\}_{i=2}^4 = \{\bar{x}_2, \bar{x}_3, \bar{x}_4\}$, dengan $\bar{x}_2 = (4.9, 3.0, 1.4, 0.2)$ dan *centroid* $\bar{\mu}_2 = (4.7, 3.1, 1.4, 0.2)$, diperoleh:

$$SE_2 = \sum_{l=1}^4 (x_{2l} - \mu_{2l})^2$$

$$= (4.9 - 4.7)^2 + (3.0 - 3.1)^2 + (1.4 - 1.4)^2 + (0.2 - 0.2)^2 = 0.05$$

$\bar{x}_3 = (4.7, 3.2, 1.3, 0.2)$ dengan *centroid* $\bar{\mu}_2 = (4.7, 3.1, 1.4, 0.2)$,

diperoleh:

$$\begin{aligned} SE_3 &= \sum_{l=1}^4 (x_{3l} - \mu_{3l})^2 \\ &= (4.7 - 4.7)^2 + (3.2 - 3.1)^2 + (1.3 - 1.4)^2 + (0.2 - 0.2)^2 = 0.02 \end{aligned}$$

$\bar{x}_4 = (4.6, 3.1, 1.5, 0.2)$ dengan *centroid* $\bar{\mu}_2 = (4.7, 3.1, 1.4, 0.2)$,

diperoleh:

$$\begin{aligned} SE_4 &= \sum_{l=1}^4 (x_{4l} - \mu_{4l})^2 \\ &= (4.6 - 4.7)^2 + (3.1 - 3.1)^2 + (1.5 - 1.4)^2 + (0.2 - 0.2)^2 = 0.02 \end{aligned}$$

- Untuk $j = 2$, maka diperoleh $C_j = C_3 = \{\bar{x}_5\}$, sedemikian hingga $\{\bar{x}_i\}_{i=5}^5 = \{\bar{x}_5\}$, dengan $\bar{x}_5 = (5, 3.6, 1.4, 0.2)$ dan *centroid* $\bar{\mu}_3 = (5, 3.6, 1.4, 0.2)$,

diperoleh:

$$\begin{aligned} SE_5 &= \sum_{l=1}^4 (x_{5l} - \mu_{5l})^2 \\ &= (5 - 5)^2 + (3.6 - 3.6)^2 + (1.4 - 1.4)^2 + (0.2 - 0.2)^2 = 0 \end{aligned}$$

Dengan demikian,

$$\begin{aligned} total SSE &= SE_1 + SE_2 + SE_3 + SE_4 + SE_5 \\ &= 0 + 0.05 + 0.02 + 0.02 + 0 = 0.09 \end{aligned}$$

2. Silhoutte Score

Perhitungan ini menggunakan persamaan (3.8) berikut:

$$S(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

dengan:

$$a(i) = \frac{1}{|C_i| - 1} \sum_{j \in C_i, i \neq j} d(i, j),$$

merupakan rata-rata jarak dari data ke- i ke semua anggota lain dalam klasternya sendiri C_i .

$$b(i) = \min_{j \neq i} \frac{1}{|C_j|} \sum_{j \in C_j} d(i, j),$$

merupakan nilai minimum dari rata-rata jarak dari data ke- i ke anggota kluster lain yang berbeda. Perhitungan jarak ini menggunakan jarak *Euclidean* menggunakan persamaan (3.1).

- Untuk $\bar{x}_1$, dimana $\bar{x}_1$ termasuk ke dalam C_1

Karena $\bar{x}_1$ merupakan satu-satunya anggota *cluster* 1 atau C_1 , maka:

$$a(1) = 0$$

Kemudian menghitung $b(1)$, yaitu jarak dari $\bar{x}_1$ ke anggota dari kluster lain, yaitu C_2 dan C_3 :

$$b(1) = \min(C_2, C_3)$$

Jarak rata-rata $\bar{x}_1$ Ke C_2 , dimana $C_2 = \{\bar{x}_2, \bar{x}_3, \bar{x}_4\}$:

$$\begin{aligned} \frac{d(\bar{x}_1, \bar{x}_2) + d(\bar{x}_1, \bar{x}_3) + d(\bar{x}_1, \bar{x}_4)}{3} &= \frac{0.5385 + 0.5099 + 0.60487}{3} \\ &= 0.5655 \end{aligned}$$

Jarak rata-rata $\bar{x}_1$ Ke C_3 , dimana $C_3 = \{\bar{x}_5\}$

$$d(\bar{x}_1, \bar{x}_5) = 0.1414$$

Karena $b(1) = \min(C_2, C_3)$, maka $b(1) = 0.1414$.

Sehingga *Silhouette Score* untuk $\bar{x}_1$:

$$S(1) = \frac{0.1414 - 0}{\max(0, 0.1414)} = \frac{0.1414}{0.1414} = 1$$

- Untuk $\bar{x}_2$, dimana $\bar{x}_2$ termasuk ke dalam C_2

Karena $\bar{x}_2$ memiliki jarak terdekat dengan $\bar{x}_3$ dan $\bar{x}_4$, maka:

$$a(i) = \frac{1}{|C_i| - 1} \sum_{j \in C_i, i \neq j} d(i, j)$$

$$a(2) = \frac{1}{3 - 1} [d(\bar{x}_2, \bar{x}_3) + d(\bar{x}_2, \bar{x}_4)] = \frac{1}{2} (0.3 + 0.3317) = 0.3158$$

Kemudian menghitung $b(2)$, yaitu jarak rata-jarak dari $\bar{x}_2$ ke anggota dari kluster lain, yaitu C_1 dan C_3 :

$$b(2) = \min(C_1, C_3)$$

Jarak rata-rata $\bar{x}_2$ Ke C_1 , dimana $C_1 = \{\bar{x}_1\}$:

$$d(\bar{x}_2, \bar{x}_1) = 0.5385$$

Jarak rata-rata $\bar{x}_2$ Ke C_3 , dimana $C_3 = \{\bar{x}_5\}$

$$d(\bar{x}_2, \bar{x}_5) = 0.6083$$

Karena $b(2) = \min(C_1, C_3)$, maka $b(2) = 0.5385$.

Sehingga *Silhouette Score* untuk $\bar{x}_2$:

$$S(2) = \frac{0.5385 - 0.3158}{\max(0.3158, 0.5385)} = \frac{0.5385 - 0.3158}{0.5385} = 0.4135$$

- Untuk $\bar{x}_3$, dimana $\bar{x}_3$ termasuk ke dalam C_2

Karena $\bar{x}_3$ memiliki jarak terdekat dengan $\bar{x}_2$ dan $\bar{x}_4$, maka:

$$a(3) = \frac{1}{3 - 1} [d(\bar{x}_3, \bar{x}_2) + d(\bar{x}_3, \bar{x}_4)] = \frac{1}{2} (0.3 + 0.2449) = 0.2724$$

Kemudian menghitung $b(3)$, yaitu jarak rata-jarak dari $\bar{x}_3$ ke anggota dari kluster lain, yaitu C_1 dan C_3 :

$$b(3) = \min(C_1, C_3)$$

Jarak rata-rata $\bar{x}_3$ Ke C_1 , dimana $C_1 = \{\bar{x}_1\}$:

$$d(\bar{x}_3, \bar{x}_1) = 0.510$$

Jarak rata-rata $\bar{x}_3$ Ke C_3 , dimana $C_3 = \{\bar{x}_5\}$

$$d(\bar{x}_3, \bar{x}_5) = 0.510$$

Karena $b(3) = \min(C_1, C_3)$, maka $b(3) = 0.510$.

Sehingga *Silhouette Score* untuk $\bar{x}_3$:

$$S(3) = \frac{0.510 - 0.2724}{\max(0.2724, 0.510)} = \frac{0.510 - 0.2725}{0.510} = 0.4657$$

- Untuk $\bar{x}_4$, dimana $\bar{x}_4$ termasuk ke dalam C_2

Karena $\bar{x}_4$ memiliki jarak terdekat dengan $\bar{x}_2$ dan $\bar{x}_3$, maka:

$$a(4) = \frac{1}{3-1} [d(\bar{x}_4, \bar{x}_2) + d(\bar{x}_4, \bar{x}_3)] = \frac{1}{2} (0.3317 + 0.2449) = 0.2883$$

Kemudian menghitung $b(4)$, yaitu jarak rata-jarak dari $\bar{x}_4$ ke anggota dari klaster lain, yaitu C_1 dan C_3 :

$$b(4) = \min(C_1, C_3)$$

Jarak rata-rata $\bar{x}_4$ Ke C_1 , dimana $C_1 = \{\bar{x}_1\}$:

$$d(\bar{x}_4, \bar{x}_1) = 0.6495$$

Jarak rata-rata $\bar{x}_4$ Ke C_3 , dimana $C_3 = \{\bar{x}_5\}$

$$d(\bar{x}_4, \bar{x}_5) = 0.6495$$

Karena $b(4) = \min(C_1, C_3)$, maka $b(4) = 0.6495$.

Sehingga *Silhouette Score* untuk $\bar{x}_4$:

$$S(4) = \frac{0.6495 - 0.2883}{\max(0.2883, 0.6495)} = \frac{0.6495 - 0.2883}{0.6495} = 0.5561$$

- Untuk $\bar{x}_5$, dimana $\bar{x}_5$ termasuk ke dalam C_3

Karena $\bar{x}_5$ merupakan satu-satunya anggota *cluster* 3 atau C_3 , maka:

$$a(5) = 0$$

Kemudian menghitung $b(5)$, yaitu jarak dari $\bar{x}_5$ ke anggota dari kluster lain, yaitu C_1 dan C_2 :

$$b(5) = \min(C_1, C_2)$$

Jarak rata-rata $\bar{x}_5$ Ke C_1 , dimana $C_1 = \{\bar{x}_1\}$:

$$d(\bar{x}_5, \bar{x}_1) = 0.1414$$

Jarak rata-rata $\bar{x}_5$ Ke C_2 , dimana $C_2 = \{\bar{x}_2, \bar{x}_3, \bar{x}_4\}$

$$\frac{d(\bar{x}_5, \bar{x}_2) + d(\bar{x}_5, \bar{x}_3) + d(\bar{x}_5, \bar{x}_4)}{3} = \frac{0.510 + 0.538 + 0.3}{3} = 0.4493$$

Karena $b(5) = \min(C_1, C_2)$, maka $b(5) = 0.1414$.

Sehingga *Silhouette Score* untuk $\bar{x}_5$:

$$S(5) = \frac{0.1414 - 0}{\max(0, 0.1414)} = \frac{0.1414 - 0}{0.1414} = 1$$

$$\begin{aligned} S_{total} &= \frac{S(1) + S(2) + S(3) + S(4) + S(5)}{5} \\ &= \frac{1 + 0.4135 + 0.4675 + 0.5561 + 1}{5} = 0.687 \end{aligned}$$

3. Akurasi

Untuk menghitung akurasi digunakan persamaan (3.9) berikut:

$$Akurasi = \frac{Jumlah\ Prediksi\ Benar}{Total\ Data} \times 100\%$$

Hasil *clustering* tidak memiliki label yang tetap, maka dilakukan pemetaan antara label hasil *clustering* dengan label asli berdasarkan jumlah kemunculan terbanyak di masing-masing *cluster*. Sehingga, pada perhitungan ini diperoleh jumlah prediksi benar sama dengan tiga dan jumlah prediksi salah sama dengan dua. Maka akurasi dapat dihitung sebagai berikut:

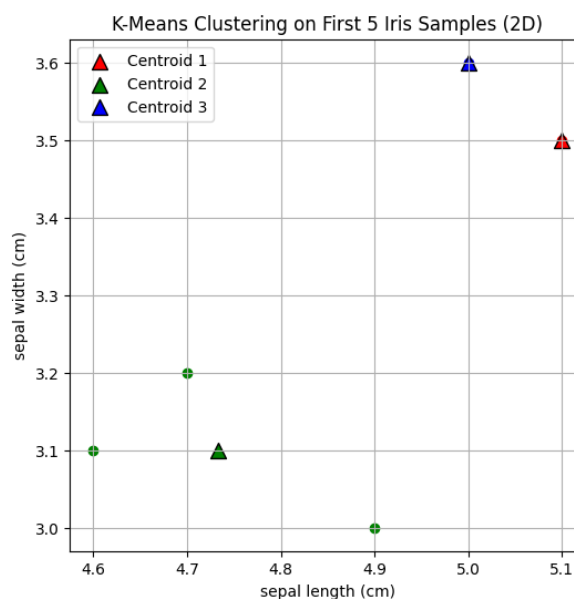
$$Akurasi = \frac{3}{5} \times 100\% = 60\%$$

Berdasarkan hasil perhitungan manual *K-Means* pada data Iris di atas, didapatkan hasil *clustering* hingga iterasi ke-2, dimana titik pusat tidak lagi berubah dan tidak ada data yang berpindah antar *cluster*. Diperoleh metrik evaluasi dengan SSE 0.09, *Silhouette Score* 0.687, dan akurasi 60% dengan *centroid* akhirnya:

$$\bar{\mu}_1 = (5.1, 3.5, 1.4, 0.2),$$

$$\bar{\mu}_2 = (4.7, 3.1, 1.4, 0.2),$$

$$\bar{\mu}_3 = (5, 3.6, 1.4, 0.2).$$



Gambar 4.1 Hasil *K-Means Clustering* Dataset Iris (5 Data pertama)

4.2.2 *K-Means* pada Dataset Iris (Menggunakan Seluruh Data)

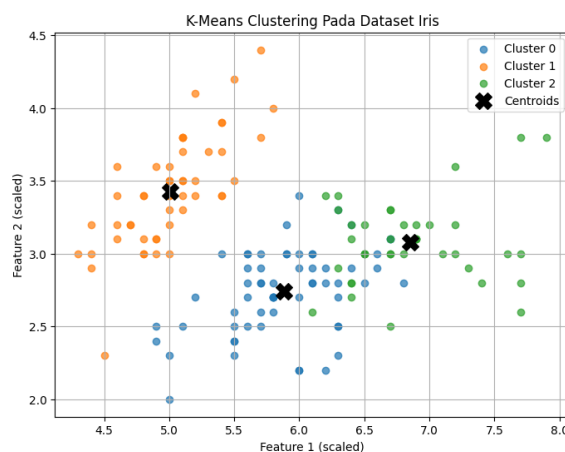
Pada subbab ini, dilakukan implementasi algoritma *K-Means* pada dataset Iris secara keseluruhan menggunakan pemrograman *Python*. Perhitungan ini bisa dilihat melalui tautan berikut ini: <https://tinyurl.com/Dataset-Iris>.

Pada percobaan ini, *K-Means* diterapkan tanpa dilakukan proses *scaling* terhadap fitur-fitur. Keputusan ini diambil setelah dilakukan pengujian terhadap dua kondisi percobaan, yaitu menggunakan data yang telah di *scale* (menggunakan *StandardScaler*) dan data yang tidak di *scale*. Berdasarkan hasil

pengujian, diperoleh bahwa akurasi *clustering* pada data tanpa *scaling* lebih tinggi dibandingkan dengan data yang telah di *scale*.

Dengan demikian, pemilihan data tanpa *scaling* dalam penelitian ini dilakukan, berdasarkan hasil percobaan yaitu didapatkan hasil *clustering* terbaik yang ditunjukkan oleh nilai akurasi sebesar 88.7%, lebih tinggi dibandingkan ketika data di standarisasi yaitu 85.3%. Berdasarkan lima kali percobaan yang dilakukan oleh penulis, diperoleh hasil *clustering* yang konsisten, baik dari sisi jumlah iterasi, nilai SSE, nilai *Silhouette Score*, maupun akurasi. Jumlah iterasi yang dihasilkan untuk mencapai konvergen pada setiap percobaan adalah sebanyak 11 iterasi. Nilai SSE yang dihasilkan sebesar 78.86, menunjukkan bahwa data dalam tiap klaster relatif dekat terhadap *centroid* masing-masing.

Selain itu, nilai *Silhouette Score* yang diperoleh adalah 0.551, dan akurasi nya sebesar 88.7%, yang berarti sebagian besar data berhasil dikelompokkan ke klaster yang sesuai dengan kelas aslinya. Dengan demikian, hasil *clustering* ini menunjukkan bahwa algoritma *K-Means* efektif dalam mengelompokkan dataset Iris dalam percobaan ini.



Gambar 4.2 Hasil *K-Means Clustering* Dataset Iris

4.2.3 Implementasi *K-Means* Manual pada Dataset Wine (5 Data Pertama)

Berdasarkan tabel 4.2 subset data penelitian pada data Wine terdapat lima data pertama yang di notasikan sebagai berikut:

$\bar{x}_i = (x_{i1}, x_{i2}, \dots, x_{iq})$, q adalah jumlah fitur, maka:

$$\bar{x}_1 = (x_{11}, x_{12}, x_{13}, x_{14}) = (14.23, 1.71, 2.43, 15.6)$$

$$\bar{x}_2 = (x_{21}, x_{22}, x_{23}, x_{24}) = (13.2, 1.78, 2.14, 11.2)$$

$$\bar{x}_3 = (x_{31}, x_{32}, x_{33}, x_{34}) = (13.16, 2.36, 2.67, 18.6)$$

$$\bar{x}_4 = (x_{41}, x_{42}, x_{43}, x_{44}) = (14.37, 1.95, 2.5, 16.8)$$

$$\bar{x}_5 = (x_{51}, x_{52}, x_{53}, x_{54}) = (13.24, 2.59, 2.87, 21)$$

Berikut merupakan perhitungan manual beserta langkahnya:

Iterasi 1

1. Menentukan jumlah *cluster*

Karena pada dataset Wine memiliki 3 jenis anggur yang berbeda, maka:

$$K = 3$$

2. Inisialisasi *centroid*

Penentuan *centroid*/pusat awal *cluster* didapatkan secara acak/*random* menggunakan metode *randbetween* pada Microsoft Excel.

$$\bar{\mu}_1 = (14.23, 1.71, 2.43, 15.6)$$

$$\bar{\mu}_2 = (13.16, 2.36, 2.67, 18.6)$$

$$\bar{\mu}_3 = (13.24, 2.59, 2.87, 21)$$

3. Perhitungan jarak *Euclidean*

Untuk mengukur jarak antara data dengan *centroid* digunakan jarak *Euclidean distance* yaitu dengan rumus pada persamaan (3.1). Kemudian

perhitungan jarak *Euclidean* dimulai antara data pertama ($\bar{x}_1$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_1, \bar{\mu}_1) = \|\bar{x}_1 - \bar{\mu}_1\| = 0$$

$$d(\bar{x}_1, \bar{\mu}_2) = \|\bar{x}_1 - \bar{\mu}_2\| = 3.26$$

$$d(\bar{x}_1, \bar{\mu}_3) = \|\bar{x}_1 - \bar{\mu}_3\| = 5.58$$

Kemudian dilanjutkan dengan perhitungan jarak *Euclidean* antara data kedua ($\bar{x}_2$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_2, \bar{\mu}_1) = \|\bar{x}_2 - \bar{\mu}_1\| = 4.53$$

$$d(\bar{x}_2, \bar{\mu}_2) = \|\bar{x}_2 - \bar{\mu}_2\| = 7.44$$

$$d(\bar{x}_2, \bar{\mu}_3) = \|\bar{x}_2 - \bar{\mu}_3\| = 9.86$$

Kemudian dilanjutkan dengan perhitungan jarak *Euclidean* antara data ketiga ($\bar{x}_3$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_3, \bar{\mu}_1) = \|\bar{x}_3 - \bar{\mu}_1\| = 3.26$$

$$d(\bar{x}_3, \bar{\mu}_2) = \|\bar{x}_3 - \bar{\mu}_2\| = 0$$

$$d(\bar{x}_3, \bar{\mu}_3) = \|\bar{x}_3 - \bar{\mu}_3\| = 2.42$$

Kemudian dilanjutkan dengan perhitungan jarak *Euclidean* antara data keempat ($\bar{x}_4$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_4, \bar{\mu}_1) = \|\bar{x}_4 - \bar{\mu}_1\| = 1.23$$

$$d(\bar{x}_4, \bar{\mu}_2) = \|\bar{x}_4 - \bar{\mu}_2\| = 2.21$$

$$d(\bar{x}_4, \bar{\mu}_3) = \|\bar{x}_4 - \bar{\mu}_3\| = 4.41$$

Terakhir dilanjutkan dengan perhitungan jarak *Euclidean* antara data kelima ($\bar{x}_5$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_5, \bar{\mu}_1) = \|\bar{x}_5 - \bar{\mu}_1\| = 5.58$$

$$d(\bar{x}_5, \bar{\mu}_2) = \|\bar{x}_5 - \bar{\mu}_2\| = 2.42$$

$$d(\bar{x}_5, \bar{\mu}_3) = \|\bar{x}_5 - \bar{\mu}_3\| = 0$$

4. Pengelompokan data sesuai *cluster* berdasarkan jarak terpendek

Setelah menghitung jarak *Euclidean* langkah selanjutnya yaitu mengelompokkan data sesuai *cluster* berdasarkan jarak terpendek.

Tabel 4.5 Pengelompokan *Cluster* pada Iterasi 1

$\bar{x}_i$	$d(\bar{\mu}_1)$	$d(\bar{\mu}_2)$	$d(\bar{\mu}_3)$	Cluster
$\bar{x}_1$	0	3.26	5.58	1
$\bar{x}_2$	4.53	7.44	9.86	1
$\bar{x}_3$	3.26	0	2.42	2
$\bar{x}_4$	1.23	2.21	4.41	1
$\bar{x}_5$	5.58	2.42	0	3

5. Perhitungan *centroid* berikutnya atau rata-rata tiap *cluster*

Berdasarkan tabel 4.5 menunjukkan hasil pengelompokan *cluster* pada iterasi pertama. Kemudian langkah selanjutnya adalah memperbarui *centroid* untuk setiap *cluster*. Proses ini dilakukan dengan menghitung rata-rata menggunakan persamaan (3.3) dari data yang termasuk dalam setiap *cluster*.

$\bar{\mu}_1 = \bar{x}_1, \bar{x}_2, \bar{x}_4$, karena $\bar{x}_1, \bar{x}_2, \bar{x}_4$ termasuk ke dalam *cluster* 1.

$\bar{\mu}_2 = \bar{x}_3$, karena $\bar{x}_3$ termasuk ke dalam *cluster* 2.

$\bar{\mu}_3 = \bar{x}_5$, karena $\bar{x}_5$ termasuk ke dalam *cluster* 3.

Sehingga rata-rata untuk nilai *centroid* baru adalah

$$\bar{\mu}_1 = \frac{(14.23, 1.71, 2.43, 15.6) + (13.2, 1.78, 2.14, 11.2) + (14.37, 1.95, 2.5, 16.8)}{3}$$

$$= (13.93, 1.81, 2.36, 14.53)$$

$$\bar{\mu}_2 = (13.16, 2.36, 2.67, 18.6)$$

$$\bar{\mu}_3 = (13.24, 2.59, 2.87, 21)$$

Iterasi 2

1. Menggunakan *centroid* baru yang didapatkan dari iterasi pertama

$$\bar{\mu}_1 = (13.93, 1.81, 2.36, 14.53)$$

$$\bar{\mu}_2 = (13.16, 2.36, 2.67, 18.6)$$

$$\bar{\mu}_3 = (13.24, 2.59, 2.87, 21)$$

2. Perhitungan jarak *Euclidean* menggunakan persamaan (3.1)

Dimulai dengan perhitungan jarak *Euclidean* antara data pertama ($\bar{x}_1$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_1, \bar{\mu}_1) = \|\bar{x}_1 - \bar{\mu}_1\| = 1.12$$

$$d(\bar{x}_1, \bar{\mu}_2) = \|\bar{x}_1 - \bar{\mu}_2\| = 3.26$$

$$d(\bar{x}_1, \bar{\mu}_3) = \|\bar{x}_1 - \bar{\mu}_3\| = 5.58$$

Kemudian dilanjut perhitungan jarak *Euclidean* dimulai antara data kedua ($\bar{x}_2$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_2, \bar{\mu}_1) = \|\bar{x}_2 - \bar{\mu}_1\| = 3.42$$

$$d(\bar{x}_2, \bar{\mu}_2) = \|\bar{x}_2 - \bar{\mu}_2\| = 7.44$$

$$d(\bar{x}_2, \bar{\mu}_3) = \|\bar{x}_2 - \bar{\mu}_3\| = 9.86$$

Kemudian dilanjut perhitungan jarak *Euclidean* dimulai antara data ketiga ($\bar{x}_3$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_3, \bar{\mu}_1) = \|\bar{x}_3 - \bar{\mu}_1\| = 4.18$$

$$d(\bar{x}_3, \bar{\mu}_2) = \|\bar{x}_3 - \bar{\mu}_2\| = 0$$

$$d(\bar{x}_3, \bar{\mu}_3) = \|\bar{x}_3 - \bar{\mu}_3\| = 2.42$$

Kemudian dilanjut perhitungan jarak *Euclidean* dimulai antara data keempat ($\bar{x}_4$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_4, \bar{\mu}_1) = \|\bar{x}_4 - \bar{\mu}_1\| = 2.32$$

$$d(\bar{x}_4, \bar{\mu}_2) = \|\bar{x}_4 - \bar{\mu}_2\| = 2.21$$

$$d(\bar{x}_4, \bar{\mu}_3) = \|\bar{x}_4 - \bar{\mu}_3\| = 4.41$$

Terakhir dilanjut dengan perhitungan jarak *Euclidean* antara data kelima ($\bar{x}_5$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_5, \bar{\mu}_1) = \|\bar{x}_5 - \bar{\mu}_1\| = 6.57$$

$$d(\bar{x}_5, \bar{\mu}_2) = \|\bar{x}_5 - \bar{\mu}_2\| = 2.42$$

$$d(\bar{x}_5, \bar{\mu}_3) = \|\bar{x}_5 - \bar{\mu}_3\| = 0$$

3. Pengelompokan data sesuai *cluster* berdasarkan jarak terpendek

Setelah menghitung jarak *Euclidean* langkah selanjutnya yaitu mengelompokkan data sesuai *cluster* berdasarkan jarak terpendek.

Tabel 4.6 Pengelompokan *Cluster* pada Iterasi 2

$\bar{x}_i$	$d(\bar{\mu}_1)$	$d(\bar{\mu}_2)$	$d(\bar{\mu}_3)$	<i>Cluster</i>
$\bar{x}_1$	1.12	3.26	5.58	1
$\bar{x}_2$	3.42	8.55	9.86	1
$\bar{x}_3$	4.19	0	2.42	2
$\bar{x}_4$	2.32	2.21	4.41	2
$\bar{x}_5$	6.57	2.42	0	3

4. Perhitungan *centroid* berikutnya atau rata-rata tiap *cluster*

$$\bar{\mu}_1 = \frac{(14.23, 1.71, 2.43, 15.6) + (13.2, 1.78, 2.14, 11.2)}{2}$$

$$= (13.71, 1.74, 2.28, 13.4)$$

$$\bar{\mu}_2 = \frac{(13.16, 2.36, 2.67, 18.6) + (14.37, 1.95, 2.5, 16.8)}{2} = (13.76, 2.15, 2.58, 17.7)$$

$$\bar{\mu}_3 = (13.24, 2.59, 2.87, 21)$$

Berdasarkan hasil perhitungan manual *K-Means* pada data Wine di atas, dilakukan hasil *clustering* hingga 2 iterasi saja, tetapi pada hitungan *python* mencapai konvergensi pada iterasi ke-4. Diperoleh metrik evaluasi dengan SSE 5.6827,

Silhouette Score 0.5995, SSE 5.6827, akurasi 60%. Hasil iterasi ke-4 sebagai berikut:

Tabel 4.7 Pengelompokan *Cluster* pada Iterasi 4

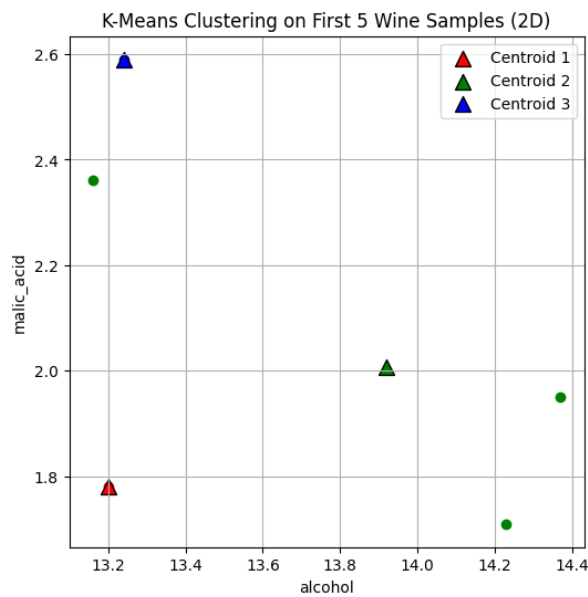
$\bar{x}_i$	$d(\bar{\mu}_1)$	$d(\bar{\mu}_2)$	$d(\bar{\mu}_3)$	<i>Cluster</i>
$\bar{x}_1$	4.53	1.47	5.58	2
$\bar{x}_2$	0	5.86	9.86	1
$\bar{x}_3$	7.44	1.81	2.42	2
$\bar{x}_4$	5.73	0.50	4.41	2
$\bar{x}_5$	9.86	4.11	0	3

Centroid akhirnya:

$$\bar{\mu}_1 = (13.2, 1.78, 2.14, 11.2),$$

$$\bar{\mu}_2 = (13.92, 2.01, 2.53, 17.00),$$

$$\bar{\mu}_3 = (13.24, 2.59, 2.87, 21.00).$$



Gambar 4.3 Hasil *K-Means Clustering* Dataset Wine (5 Data pertama)

4.2.4 *K-Means* pada Dataset Wine (Menggunakan Seluruh Data)

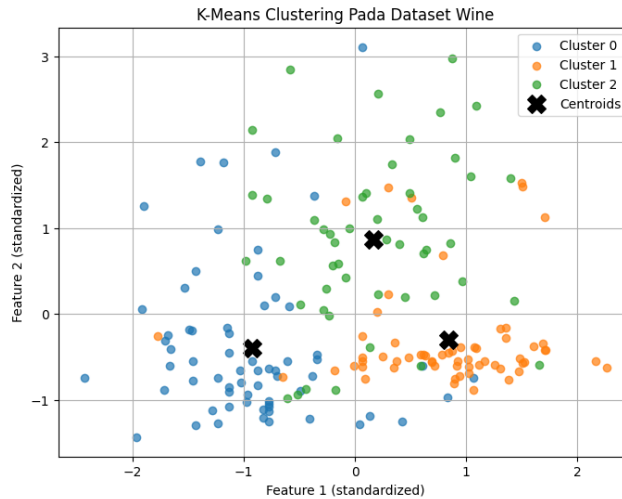
Pada subbab ini, dilakukan implementasi algoritma *K-Means* pada dataset Wine secara keseluruhan menggunakan bahasa pemrograman *Python*. Untuk melihat perhitungan menggunakan *Python* bisa dilihat melalui tautan berikut ini:

<https://tinyurl.com/Dataset-Wine>.

Pada percobaan ini, *K-Means* diterapkan dengan dilakukannya proses *scaling* terhadap fitur-fitur. Keputusan ini diambil setelah dilakukan pengujian terhadap dua kondisi percobaan, yaitu menggunakan data yang telah di *scale* (menggunakan *StandardScaler*) dan data yang tidak di *scale*. Berdasarkan hasil pengujian, diperoleh bahwa akurasi *clustering* pada data yang telah di *scale* lebih tinggi dibandingkan dengan data tanpa di *scale*.

Dengan demikian, pemilihan data yang telah di *scale* dalam penelitian ini dilakukan, berdasarkan hasil percobaan yaitu didapatkan hasil *clustering* terbaik yang ditunjukkan oleh nilai akurasi sebesar 96.6%, lebih tinggi dibandingkan dengan data tanpa di *scale* yaitu 70.22%. Berdasarkan lima kali percobaan yang dilakukan oleh penulis, diperoleh hasil *clustering* yang konsisten, baik dari sisi jumlah iterasi, nilai SSE, nilai *Silhouette Score*, maupun akurasi. Jumlah iterasi yang dihasilkan untuk mencapai konvergen pada setiap percobaan adalah sebanyak 3 iterasi. Nilai SSE yang dihasilkan sebesar 1277.93, menunjukkan bahwa data dalam tiap kluster relatif dekat terhadap *centroid* masing-masing.

Selain itu, nilai *Silhouette Score* yang diperoleh adalah 0.285, dan akurasi nya sebesar 96.6%, yang berarti sebagian besar data berhasil dikelompokkan ke kluster yang sesuai dengan kelas aslinya.



Gambar 4.4 Hasil *K-Means Clustering* Dataset Wine

4.3 Implementasi PSO

4.3.1 Dataset Iris (5 Data Pertama)

Setelah didapatkan *centroid* awal dari K-Means, PSO digunakan untuk mengoptimalkan posisi *centroid*.

1. Inisialisasi Partikel PSO

a. Inisialisasi Posisi Partikel

$$\bar{p}_i = (p_{i1}, p_{i2}, p_{i3}, \dots, p_{il})$$

Tabel 4.8 Inisialisasi Posisi Partikel pada PSO

Partikel ($\bar{p}_i$)	$\bar{\mu}_1$	$\bar{\mu}_2$	$\bar{\mu}_3$
$\bar{p}_1$	(5.1, 3.5, 1.4, 0.2)	(4.7, 3.1, 1.4, 0.2)	(5.0, 3.6, 1.4, 0.2)
$\bar{p}_2$	(5.2, 3.4, 1.3, 0.2)	(4.8, 3.2, 1.5, 0.2)	(4.9, 3.5, 1.3, 0.2)
$\bar{p}_3$	(5.0, 3.6, 1.4, 0.2)	(4.6, 3.0, 1.3, 0.2)	(5.1, 3.4, 1.5, 0.2)

Tabel 4.8 menunjukkan hasil inisialisasi posisi awal partikel pada algoritma PSO. Setiap partikel terdiri dari tiga *centroid* dengan nilai fitur masing-masing.

b. Inisialisasi Kecepatan Partikel

$$\bar{v}_i = (v_{i1}, v_{i2}, \dots, v_{il})$$

Pada tahap awal, kecepatan partikel diinisialisasi secara acak untuk setiap dimensi. Nilai kecepatan awal untuk masing-masing partikel sebagai berikut:

- $\bar{v}_1 = (-0.1, 0.2, -0.5, 0.0)$
- $\bar{v}_2 = (0.2, -0.1, 0.1, 0.0)$
- $\bar{v}_3 = (-0.2, 0.1, -0.5, 0.0)$

c. Parameter PSO

Parameter pada PSO juga ditentukan dengan nilai sebagai berikut:

- $\omega = 0.5$
- $a_1 = 2.0$
- $a_2 = 2.0$
- $r_1 = 0.7$
- $r_2 = 0.9$

2. Perhitungan SSE (Fungsi Objektif) Menggunakan Persamaan (3.4)

a. Hitung SSE untuk Partikel 1

Centroid awal (partikel 1) diambil dari tabel 4.8.

$$\bar{\mu}_1 = (5.1, 3.5, 1.4, 0.2)$$

$$\bar{\mu}_2 = (4.7, 3.1, 1.4, 0.2)$$

$$\bar{\mu}_3 = (5.0, 3.6, 1.4, 0.2)$$

Misalkan menggunakan 5 data pertama dari dataset Iris sesuai pada data tabel 4.1, kemudian menghitung jarak *Euclidean* dimulai antara data pertama ($\bar{x}_1$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_1, \bar{\mu}_1) = \|\bar{x}_1 - \bar{\mu}_1\| = 0$$

$$d(\bar{x}_1, \bar{\mu}_2) = \|\bar{x}_1 - \bar{\mu}_2\| = 0.566$$

$$d(\bar{x}_1, \bar{\mu}_3) = \|\bar{x}_1 - \bar{\mu}_3\| = 0.141$$

Kemudian dilanjut dengan perhitungan jarak *Euclidean* antara data kedua ($\bar{x}_2$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_2, \bar{\mu}_1) = \|\bar{x}_2 - \bar{\mu}_1\| = 0.538$$

$$d(\bar{x}_2, \bar{\mu}_2) = \|\bar{x}_2 - \bar{\mu}_2\| = 0.224$$

$$d(\bar{x}_2, \bar{\mu}_3) = \|\bar{x}_2 - \bar{\mu}_3\| = 0.608$$

Kemudian dilanjut dengan perhitungan jarak *Euclidean* antara data ketiga ($\bar{x}_3$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_3, \bar{\mu}_1) = \|\bar{x}_3 - \bar{\mu}_1\| = 0.51$$

$$d(\bar{x}_3, \bar{\mu}_2) = \|\bar{x}_3 - \bar{\mu}_2\| = 0.141$$

$$d(\bar{x}_3, \bar{\mu}_3) = \|\bar{x}_3 - \bar{\mu}_3\| = 0.510$$

Kemudian dilanjut dengan perhitungan jarak *Euclidean* antara data keempat ($\bar{x}_4$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_4, \bar{\mu}_1) = \|\bar{x}_4 - \bar{\mu}_1\| = 0.649$$

$$d(\bar{x}_4, \bar{\mu}_2) = \|\bar{x}_4 - \bar{\mu}_2\| = 0.141$$

$$d(\bar{x}_4, \bar{\mu}_3) = \|\bar{x}_4 - \bar{\mu}_3\| = 0.649$$

Terakhir dilanjut dengan perhitungan jarak *Euclidean* antara data kelima ($\bar{x}_5$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_5, \bar{\mu}_1) = \|\bar{x}_5 - \bar{\mu}_1\| = 0.141$$

$$d(\bar{x}_5, \bar{\mu}_2) = \|\bar{x}_5 - \bar{\mu}_2\| = 0.583$$

$$d(\bar{x}_5, \bar{\mu}_3) = \|\bar{x}_5 - \bar{\mu}_3\| = 0$$

Setelah dihitung jarak *Euclidean*, kemudian dikelompokkan ke dalam *cluster* berdasarkan jarak terkecil.

Tabel 4.9 Pengelompokan *Cluster* Partikel 1 pada Iterasi 1

$\bar{x}_i$	$d(\bar{\mu}_1)$	$d(\bar{\mu}_2)$	$d(\bar{\mu}_3)$	<i>Cluster</i>
$\bar{x}_1$	0	0.566	0.141	1
$\bar{x}_2$	0.538	0.224	0.608	2
$\bar{x}_3$	0.51	0.141	0.510	2
$\bar{x}_4$	0.649	0.141	0.649	2
$\bar{x}_5$	0.141	0.583	0	3

Setelah semua data masuk ke *cluster* masing-masing, kemudian menghitung SSE dengan persamaan (3.4).

- Untuk *cluster* 1 yang terdiri dari data $C_1 = \{\bar{x}_1\}$,

$\bar{x}_1 = (5.1, 3.5, 1.4, 0.2)$ dengan *centroid* $\bar{\mu}_1 = (5.1, 3.5, 1.4, 0.2)$,

diperoleh:

$$SE_1: (5.1 - 5.1)^2 + (3.5 - 3.5)^2 + (1.4 - 1.4)^2 + (0.2 - 0.2)^2 = 0$$

- Untuk *cluster* 2 yang terdiri dari data $C_2 = \{\bar{x}_2, \bar{x}_3, \bar{x}_4\}$,

$\bar{x}_2 = (4.9, 3.0, 1.4, 0.2)$ dengan *centroid* $\bar{\mu}_2 = (4.7, 3.1, 1.4, 0.2)$,

diperoleh:

$$SE_2: (4.9 - 4.7)^2 + (3.0 - 3.1)^2 + (1.4 - 1.4)^2 + (0.2 - 0.2)^2 = 0.05$$

$\bar{x}_3 = (4.7, 3.2, 1.3, 0.2)$ dengan *centroid* $\bar{\mu}_2 = (4.7, 3.1, 1.4, 0.2)$,

diperoleh:

$$SE_3: (4.7 - 4.7)^2 + (3.2 - 3.1)^2 + (1.3 - 1.4)^2 + (0.2 - 0.2)^2 = 0.02$$

$\bar{x}_4 = (4.6, 3.1, 1.5, 0.2)$ dengan *centroid* $\bar{\mu}_2 = (4.7, 3.1, 1.4, 0.2)$,

diperoleh:

$$SE_4: (4.6 - 4.7)^2 + (3.1 - 3.1)^2 + (1.5 - 1.4)^2 + (0.2 - 0.2)^2 = 0.02$$

- Untuk *cluster* 3 yang terdiri dari data $C_3 = \{\bar{x}_5\}$,

$\bar{x}_5 = (5, 3.6, 1.4, 0.2)$ dengan *centroid* $\bar{\mu}_3 = (5, 3.6, 1.4, 0.2)$, diperoleh:

$$SE_5: (5 - 5)^2 + (3.6 - 3.6)^2 + (1.4 - 1.4)^2 + (0.2 - 0.2)^2 = 0$$

Dengan demikian,

$$\begin{aligned} total\ SSE &= SE_1 + SE_2 + SE_3 + SE_4 + SE_5 \\ &= 0 + 0.05 + 0.02 + 0.02 + 0 = 0.09 \end{aligned}$$

b. Hitung SSE untuk Partikel 2

Centroid awal (partikel 2) diambil dari tabel 4.8.

$$\bar{\mu}_1 = (5.2, 3.4, 1.3, 0.2)$$

$$\bar{\mu}_2 = (4.8, 3.2, 1.5, 0.2)$$

$$\bar{\mu}_3 = (4.9, 3.5, 1.3, 0.2)$$

Misalkan menggunakan 5 data pertama dari dataset Iris sesuai pada data tabel 4.1, kemudian menghitung jarak *Euclidean* dimulai antara data pertama ($\bar{x}_1$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_1, \bar{\mu}_1) = \|\bar{x}_1 - \bar{\mu}_1\| = 0.173$$

$$d(\bar{x}_1, \bar{\mu}_2) = \|\bar{x}_1 - \bar{\mu}_2\| = 0.436$$

$$d(\bar{x}_1, \bar{\mu}_3) = \|\bar{x}_1 - \bar{\mu}_3\| = 0.224$$

Kemudian dilanjut dengan perhitungan jarak *Euclidean* antara data kedua ($\bar{x}_2$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_2, \bar{\mu}_1) = \|\bar{x}_2 - \bar{\mu}_1\| = 0.510$$

$$d(\bar{x}_2, \bar{\mu}_2) = \|\bar{x}_2 - \bar{\mu}_2\| = 0.245$$

$$d(\bar{x}_2, \bar{\mu}_3) = \|\bar{x}_2 - \bar{\mu}_3\| = 0.510$$

Kemudian dilanjut dengan perhitungan jarak *Euclidean* antara data ketiga ($\bar{x}_3$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_3, \bar{\mu}_1) = \|\bar{x}_3 - \bar{\mu}_1\| = 0.538$$

$$d(\bar{x}_3, \bar{\mu}_2) = \|\bar{x}_3 - \bar{\mu}_2\| = 0.223$$

$$d(\bar{x}_3, \bar{\mu}_3) = \|\bar{x}_3 - \bar{\mu}_3\| = 0.360$$

Kemudian dilanjut dengan perhitungan jarak *Euclidean* antara data keempat ($\bar{x}_4$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_4, \bar{\mu}_1) = \|\bar{x}_4 - \bar{\mu}_1\| = 0.7$$

$$d(\bar{x}_4, \bar{\mu}_2) = \|\bar{x}_4 - \bar{\mu}_2\| = 0.223$$

$$d(\bar{x}_4, \bar{\mu}_3) = \|\bar{x}_4 - \bar{\mu}_3\| = 0.538$$

Terakhir dilanjut dengan perhitungan jarak *Euclidean* antara data kelima ($\bar{x}_5$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_5, \bar{\mu}_1) = \|\bar{x}_5 - \bar{\mu}_1\| = 0.3$$

$$d(\bar{x}_5, \bar{\mu}_2) = \|\bar{x}_5 - \bar{\mu}_2\| = 0.458$$

$$d(\bar{x}_5, \bar{\mu}_3) = \|\bar{x}_5 - \bar{\mu}_3\| = 0.173$$

Setelah dihitung jarak *Euclidean*, kemudian dikelompokkan ke dalam *cluster* berdasarkan jarak terkecil.

Tabel 4.10 Pengelompokan *Cluster* Partikel 2 pada Iterasi 1

$\bar{x}_i$	$d(\bar{\mu}_1)$	$d(\bar{\mu}_2)$	$d(\bar{\mu}_3)$	<i>Cluster</i>
$\bar{x}_1$	0.173	0.436	0.224	1
$\bar{x}_2$	0.510	0.245	0.510	2
$\bar{x}_3$	0.538	0.223	0.360	2
$\bar{x}_4$	0.7	0.223	0.538	2
$\bar{x}_5$	0.3	0.458	0.173	3

Setelah semua data masuk ke *cluster* masing-masing, kemudian menghitung SSE dengan persamaan (3.4).

- Untuk *cluster* 1 yang terdiri dari data $C_1 = \{\bar{x}_1\}$,

$\bar{x}_1 = (5.1, 3.5, 1.4, 0.2)$ dengan *centroid* $\bar{\mu}_1 = (5.2, 3.4, 1.3, 0.2)$,

diperoleh:

$$SE_1: (5.1 - 5.2)^2 + (3.5 - 3.4)^2 + (1.4 - 1.3)^2 + (0.2 - 0.2)^2 =$$

$$0.03$$

- Untuk *cluster* 2 yang terdiri dari data $C_2 = \{\bar{x}_2, \bar{x}_3, \bar{x}_4\}$,

$\bar{x}_2 = (4.9, 3.0, 1.4, 0.2)$ dengan *centroid* $\bar{\mu}_2 = (4.8, 3.2, 1.5, 0.2)$,

diperoleh:

$$SE_2: (4.9 - 4.8)^2 + (3.0 - 3.2)^2 + (1.4 - 1.5)^2 + (0.2 - 0.2)^2 = 0.06$$

$\bar{x}_3 = (4.7, 3.2, 1.3, 0.2)$ dengan *centroid* $\bar{\mu}_2 = (4.8, 3.2, 1.5, 0.2)$,

diperoleh:

$$SE_3: (4.7 - 4.8)^2 + (3.2 - 3.2)^2 + (1.3 - 1.5)^2 + (0.2 - 0.2)^2 = 0.05$$

$\bar{x}_4 = (4.6, 3.1, 1.5, 0.2)$ dengan *centroid* $\bar{\mu}_2 = (4.8, 3.2, 1.5, 0.2)$,

diperoleh:

$$SE_4: (4.6 - 4.8)^2 + (3.1 - 3.2)^2 + (1.5 - 1.5)^2 + (0.2 - 0.2)^2 = 0.05$$

- Untuk *cluster* 3 yang terdiri dari data $C_3 = \{\bar{x}_4\}$,

$\bar{x}_5 = (5, 3.6, 1.4, 0.2)$ dengan *centroid* $\bar{\mu}_3 = (4.9, 3.5, 1.3, 0.2)$,

diperoleh:

$$SE_5: (5 - 4.9)^2 + (3.6 - 3.5)^2 + (1.4 - 1.3)^2 + (0.2 - 0.2)^2 = 0.03$$

Dengan demikian,

$$\begin{aligned} total\ SSE &= SE_1 + SE_2 + SE_3 + SE_4 + SE_5 \\ &= 0.03 + (0.06 + 0.05 + 0.05) + 0.03 = 0.22 \end{aligned}$$

c. Hitung SSE untuk Partikel 3

Centroid awal (partikel 3) diambil dari tabel 4.8.

$$\bar{\mu}_1 = (5.0, 3.6, 1.4, 0.2)$$

$$\bar{\mu}_2 = (4.6, 3.0, 1.3, 0.2)$$

$$\bar{\mu}_3 = (5.1, 3.4, 1.5, 0.2)$$

Misalkan menggunakan 5 data pertama dari dataset Iris sesuai pada data tabel 4.1, kemudian menghitung jarak *Euclidean* dimulai antara data pertama ($\bar{x}_1$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_1, \bar{\mu}_1) = \|\bar{x}_1 - \bar{\mu}_1\| = 0.141$$

$$d(\bar{x}_1, \bar{\mu}_2) = \|\bar{x}_1 - \bar{\mu}_2\| = 0.714$$

$$d(\bar{x}_1, \bar{\mu}_3) = \|\bar{x}_1 - \bar{\mu}_3\| = 0.141$$

Kemudian dilanjut dengan perhitungan jarak *Euclidean* antara data kedua ($\bar{x}_2$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_2, \bar{\mu}_1) = \|\bar{x}_2 - \bar{\mu}_1\| = 0.608$$

$$d(\bar{x}_2, \bar{\mu}_2) = \|\bar{x}_2 - \bar{\mu}_2\| = 0.316$$

$$d(\bar{x}_2, \bar{\mu}_3) = \|\bar{x}_2 - \bar{\mu}_3\| = 0.458$$

Kemudian dilanjut dengan perhitungan jarak *Euclidean* antara data ketiga ($\bar{x}_3$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_3, \bar{\mu}_1) = \|\bar{x}_3 - \bar{\mu}_1\| = 0.509$$

$$d(\bar{x}_3, \bar{\mu}_2) = \|\bar{x}_3 - \bar{\mu}_2\| = 0.223$$

$$d(\bar{x}_3, \bar{\mu}_3) = \|\bar{x}_3 - \bar{\mu}_3\| = 0.489$$

Kemudian dilanjut dengan perhitungan jarak *Euclidean* antara data keempat ($\bar{x}_4$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_4, \bar{\mu}_1) = \|\bar{x}_4 - \bar{\mu}_1\| = 0.648$$

$$d(\bar{x}_4, \bar{\mu}_2) = \|\bar{x}_4 - \bar{\mu}_2\| = 0.223$$

$$d(\bar{x}_4, \bar{\mu}_3) = \|\bar{x}_4 - \bar{\mu}_3\| = 0.583$$

Terakhir dilanjut dengan perhitungan jarak *Euclidean* antara data kelima ($\bar{x}_5$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_5, \bar{\mu}_1) = \|\bar{x}_5 - \bar{\mu}_1\| = 0$$

$$d(\bar{x}_5, \bar{\mu}_2) = \|\bar{x}_5 - \bar{\mu}_2\| = 0.728$$

$$d(\bar{x}_5, \bar{\mu}_3) = \|\bar{x}_5 - \bar{\mu}_3\| = 0.244$$

Setelah dihitung jarak *Euclidean*, kemudian dikelompokkan ke dalam *cluster* berdasarkan jarak terkecil.

Tabel 4.11 Pengelompokan *Cluster* Partikel 3 pada Iterasi 1

$\bar{x}_i$	$d(\bar{\mu}_1)$	$d(\bar{\mu}_2)$	$d(\bar{\mu}_3)$	Cluster
$\bar{x}_1$	0.141	0.714	0.141	3
$\bar{x}_2$	0.608	0.316	0.458	2
$\bar{x}_3$	0.509	0.223	0.489	2
$\bar{x}_4$	0.648	0.223	0.583	2
$\bar{x}_5$	0	0.728	0.244	1

Setelah semua data masuk ke *cluster* masing-masing, kemudian menghitung SSE dengan persamaan (3.4).

- Untuk *cluster* 1 yang terdiri dari data $C_1 = \{\bar{x}_5\}$,

$\bar{x}_5 = 5, 3.6, 1.4, 0.2$ dengan *centroid* $\bar{\mu}_1 = (5.0, 3.6, 1.4, 0.2)$, diperoleh:

$$SE_5: (5 - 5.0)^2 + (3.6 - 3.6)^2 + (1.4 - 1.4)^2 + (0.2 - 0.2)^2 = 0$$

- Untuk *cluster* 2 yang terdiri dari data $C_2 = \{\bar{x}_2, \bar{x}_3, \bar{x}_4\}$,

$\bar{x}_2 = (4.9, 3.0, 1.4, 0.2)$ dengan *centroid* $\bar{\mu}_2 = (4.6, 3.0, 1.3, 0.2)$,

diperoleh:

$$SE_2: (4.9 - 4.6)^2 + (3.0 - 3.0)^2 + (1.4 - 1.4)^2 + (0.2 - 0.2)^2 = 0.09$$

$\bar{x}_3 = (4.7, 3.2, 1.3, 0.2)$ dengan *centroid* $\bar{\mu}_2 = (4.6, 3.0, 1.3, 0.2)$,

diperoleh:

$$SE_3: (4.7 - 4.6)^2 + (3.2 - 3.0)^2 + (1.3 - 1.4)^2 + (0.2 - 0.2)^2 = 0.06$$

$\bar{x}_4 = (4.6, 3.1, 1.5, 0.2)$ dengan *centroid* $\bar{\mu}_2 = (4.6, 3.0, 1.3, 0.2)$,

diperoleh:

$$SE_4: (4.6 - 4.6)^2 + (3.1 - 3.0)^2 + (1.5 - 1.4)^2 + (0.2 - 0.2)^2 = 0.02$$

- Untuk *cluster* 3 yang terdiri dari data $C_3 = \{\bar{x}_1\}$,

$\bar{x}_1 = 5.1, 3.5, 1.4, 0.2$ dengan *centroid* $\bar{\mu}_3 = (5.1, 3.4, 1.5, 0.2)$, diperoleh:

$$SE_1: (5.1 - 4.9)^2 + (3.5 - 3.5)^2 + (1.4 - 1.3)^2 + (0.2 - 0.2)^2 = 0.05$$

Dengan demikian,

$$\begin{aligned} total\ SSE &= SE_1 + SE_2 + SE_3 + SE_4 + SE_5 \\ &= 0.05 + (0.09 + 0.06 + 0.02) + 0 = 0.22 \end{aligned}$$

3. *Update* Kecepatan dan Posisi

Karena ini adalah iterasi pertama dan belum ada riwayat pergerakan partikel, maka:

- $\bar{p}_{best}$ untuk setiap partikel adalah posisi awal partikel tersebut.
- $\bar{g}_{best}$ adalah posisi dari partikel dengan nilai SSE terkecil, yaitu partikel 1 ($\bar{p}_1$).

Proses *update* partikel 1 ($\bar{p}_1$) adalah sebagai berikut:

Cluster 1:

Posisi *centroid* saat ini: (5.1, 3.5, 1.4, 0.2)

$\bar{p}_{best}$: (5.1, 3.5, 1.4, 0.2)

$\bar{g}_{best}$: (5.1, 3.5, 1.4, 0.2) diambil dari $\bar{p}_1$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6) untuk partikel ke- i , fitur ke- l , pada iterasi ke- t :

$$\bar{v}_{il}^{t+1} = \omega \bar{v}_{il}^t + a_1 r_1 (\bar{p}_{best_{il}} - \bar{p}_{il}^t) + a_2 r_2 (\bar{g}_{best_{il}} - \bar{p}_{il}^t)$$

$$\bar{p}_{il}^{t+1} = \bar{p}_{il}^t + \bar{v}_{il}^{t+1}$$

Untuk $t = 1, i = 1, l = 1$, maka:

$$\bar{v}_{11}^2 = \omega \bar{v}_{11}^1 + a_1 r_1 (\bar{p}_{best_{11}} - \bar{p}_{11}^1) + a_2 r_2 (\bar{g}_{best_{11}} - \bar{p}_{11}^1)$$

$$\bar{p}_{11}^2 = \bar{p}_{11}^1 + \bar{v}_{11}^2$$

Dengan substitusi nilai:

$$\bar{v}_{11}^2 = \omega \bar{v}_{11}^1 + a_1 r_1 (\bar{p}_{best_{11}} - \bar{p}_{11}^1) + a_2 r_2 (\bar{g}_{best_{11}} - \bar{p}_{11}^1)$$

$$\bar{p}_{11}^2 = \bar{p}_{11}^1 + \bar{v}_{11}^2$$

$$\bar{v}_{11}^2 = 0.5 \cdot (-0.1) + 2.0 \cdot 0.7(5.1 - 5.1) + 2.0 \cdot 0.9(5.1 - 5.1)$$

$$\bar{p}_{11}^2 = 5.1 + \bar{v}_{11}^2$$

$$\bar{v}_{11}^2 = -0.05$$

$$\bar{p}_{11}^2 = 5.1 + (-0.05) = 5.05$$

Untuk $t = 1, i = 1, l = 2$, maka:

$$\bar{v}_{12}^2 = \omega \bar{v}_{12}^1 + a_1 r_1 (\bar{p}_{best_{12}} - \bar{p}_{12}^1) + a_2 r_2 (\bar{g}_{best_{12}} - \bar{p}_{12}^1)$$

$$\bar{p}_{12}^2 = \bar{p}_{12}^1 + \bar{v}_{12}^2$$

Dengan substitusi nilai:

$$\bar{v}_{12}^2 = 0.5 \cdot (0.2) + 2.0 \cdot 0.7(3.5 - 3.5) + 2.0 \cdot 0.9(3.5 - 3.5)$$

$$\bar{p}_{12}^2 = 3.5 + \bar{v}_{12}^2$$

$$\bar{v}_{12}^2 = 0.1$$

$$\bar{p}_{12}^2 = 3.5 + 0.1 = 3.6$$

Untuk $t = 1, i = 1, l = 3$, maka:

$$\bar{v}_{13}^2 = \omega \bar{v}_{13}^1 + a_1 r_1 (\bar{p}_{best_{13}} - \bar{p}_{13}^1) + a_2 r_2 (\bar{g}_{best_{13}} - \bar{p}_{13}^1)$$

$$\bar{p}_{13}^2 = \bar{p}_{13}^1 + \bar{v}_{13}^2$$

Dengan substitusi nilai:

$$\bar{v}_{13}^2 = 0.5 \cdot (-0.5) + 2.0 \cdot 0.7(1.4 - 1.4) + 2.0 \cdot 0.9(1.4 - 1.4)$$

$$\bar{p}_{13}^2 = 1.4 + \bar{v}_{13}^2$$

$$\bar{v}_{13}^2 = -0.25$$

$$\bar{p}_{13}^2 = 1.4 + (-0.25) = 1.15$$

Untuk $t = 1, i = 1, l = 4$, maka:

$$\bar{v}_{14}^2 = \omega \bar{v}_{14}^1 + a_1 r_1 (\bar{p}_{best_{14}} - \bar{p}_{14}^1) + a_2 r_2 (\bar{g}_{best_{14}} - \bar{p}_{14}^1)$$

$$\bar{p}_{14}^2 = \bar{p}_{14}^1 + \bar{v}_{14}^2$$

Dengan substitusi nilai:

$$\bar{v}_{14}^2 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(0.2 - 0.2) + 2.0 \cdot 0.9(0.2 - 0.2)$$

$$\bar{p}_{14}^2 = 0.2 + \bar{v}_{14}^2$$

$$\bar{v}_{14}^2 = 0$$

$$\bar{p}_{14}^2 = 0.2 + 0 = 0.2$$

Maka, posisi *centroid* baru untuk *cluster* 1 adalah

$$(5.05, 3.6, 1.15, 0.2)$$

Cluster 2:

Posisi *centroid* saat ini: (4.7, 3.1, 1.4, 0.2)

$$\bar{p}_{best}: (4.7, 3.1, 1.4, 0.2)$$

$$\bar{g}_{best}: (4.7, 3.1, 1.4, 0.2) \text{ diambil dari } \bar{p}_1$$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 1, i = 1, l = 1$, maka:

$$\bar{v}_{11}^2 = \omega \bar{v}_{11}^1 + a_1 r_1 (\bar{p}_{best_{11}} - \bar{p}_{11}^1) + a_2 r_2 (\bar{g}_{best_{11}} - \bar{p}_{11}^1)$$

$$\bar{p}_{11}^2 = \bar{p}_{11}^1 + \bar{v}_{11}^2$$

Dengan substitusi nilai:

$$\bar{v}_{11}^2 = 0.5 \cdot (0.2) + 2.0 \cdot 0.7(4.7 - 4.7) + 2.0 \cdot 0.9(4.7 - 4.7)$$

$$\bar{p}_{11}^2 = 4.7 + \bar{v}_{11}^2$$

$$\bar{v}_{11}^2 = 0.1$$

$$\bar{p}_{11}^2 = 4.7 + 0.1 = 4.8$$

Untuk $t = 1, i = 1, l = 2$, maka:

$$\bar{v}_{12}^2 = \omega \bar{v}_{12}^1 + a_1 r_1 (\bar{p}_{best_{12}} - \bar{p}_{12}^1) + a_2 r_2 (\bar{g}_{best_{12}} - \bar{p}_{12}^1)$$

$$\bar{p}_{12}^2 = \bar{p}_{12}^1 + \bar{v}_{12}^2$$

Dengan substitusi nilai:

$$\bar{v}_{12}^2 = 0.5 \cdot (-0.1) + 2.0 \cdot 0.7(3.1 - 3.1) + 2.0 \cdot 0.9(3.1 - 3.1)$$

$$\bar{p}_{12}^2 = 3.1 + \bar{v}_{12}^2$$

$$\bar{v}_{12}^2 = -0.05$$

$$\bar{p}_{12}^2 = 3.1 + (-0.05) = 3.05$$

Untuk $t = 1, i = 1, l = 3$, maka:

$$\bar{v}_{13}^2 = \omega \bar{v}_{13}^1 + a_1 r_1 (\bar{p}_{best_{13}} - \bar{p}_{13}^1) + a_2 r_2 (\bar{g}_{best_{13}} - \bar{p}_{13}^1)$$

$$\bar{p}_{13}^2 = \bar{p}_{13}^1 + \bar{v}_{13}^2$$

Dengan substitusi nilai:

$$\bar{v}_{13}^2 = 0.5 \cdot (0.1) + 2.0 \cdot 0.7(1.4 - 1.4) + 2.0 \cdot 0.9(1.4 - 1.4)$$

$$\bar{p}_{13}^2 = 1.4 + \bar{v}_{13}^2$$

$$\bar{v}_{13}^2 = 0.05$$

$$\bar{p}_{13}^2 = 1.4 + 0.05 = 1.45$$

Untuk $t = 1, i = 1, l = 4$, maka:

$$\bar{v}_{14}^2 = \omega \bar{v}_{14}^1 + a_1 r_1 (\bar{p}_{best_{14}} - \bar{p}_{14}^1) + a_2 r_2 (\bar{g}_{best_{14}} - \bar{p}_{14}^1)$$

$$\bar{p}_{14}^2 = \bar{p}_{14}^1 + \bar{v}_{14}^2$$

Dengan substitusi nilai:

$$\bar{v}_{14}^2 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(0.2 - 0.2) + 2.0 \cdot 0.9(0.2 - 0.2)$$

$$\bar{p}_{14}^2 = 0.2 + \bar{v}_{14}^2$$

$$\bar{v}_{14}^2 = 0$$

$$\bar{p}_{14}^2 = 0.2 + 0 = 0.2$$

Maka, posisi *centroid* baru untuk *cluster* 2 adalah

$$(4.8, 3.05, 1.45, 0.2)$$

Cluster 3:

Posisi *centroid* saat ini: (5.0, 3.6, 1.4, 0.2)

$$\bar{p}_{best}: (5.0, 3.6, 1.4, 0.2)$$

$$\bar{g}_{best}: (5.0, 3.6, 1.4, 0.2) \text{ diambil dari } \bar{p}_1$$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 1, i = 1, l = 1$, maka:

$$\bar{v}_{11}^2 = \omega \bar{v}_{11}^1 + a_1 r_1 (\bar{p}_{best_{11}} - \bar{p}_{11}^1) + a_2 r_2 (\bar{g}_{best_{11}} - \bar{p}_{11}^1)$$

$$\bar{p}_{11}^2 = \bar{p}_{11}^1 + \bar{v}_{11}^2$$

Dengan substitusi nilai:

$$\bar{v}_{11}^2 = 0.5 \cdot (-0.2) + 2.0 \cdot 0.7(5.0 - 5.0) + 2.0 \cdot 0.9(5.0 - 5.0)$$

$$\bar{p}_{11}^2 = 5.0 + \bar{v}_{11}^2$$

$$\bar{v}_{11}^2 = -0.1$$

$$\bar{p}_{11}^2 = 5.0 + (-0.1) = 4.9$$

Untuk $t = 1, i = 1, l = 2$, maka:

$$\bar{v}_{12}^2 = \omega \bar{v}_{12}^1 + a_1 r_1 (\bar{p}_{best_{12}} - \bar{p}_{12}^1) + a_2 r_2 (\bar{g}_{best_{12}} - \bar{p}_{12}^1)$$

$$\bar{p}_{12}^2 = \bar{p}_{12}^1 + \bar{v}_{12}^2$$

Dengan substitusi nilai:

$$\bar{v}_{12}^2 = 0.5 \cdot (0.1) + 2.0 \cdot 0.7(3.6 - 3.6) + 2.0 \cdot 0.9(3.6 - 3.6)$$

$$\bar{p}_{12}^2 = 3.6 + \bar{v}_{12}^2$$

$$\bar{v}_{12}^2 = 0.05$$

$$\bar{p}_{12}^2 = 3.6 + 0.05 = 3.65$$

Untuk $t = 1, i = 1, l = 3$, maka:

$$\bar{v}_{13}^2 = \omega \bar{v}_{13}^1 + a_1 r_1 (\bar{p}_{best_{13}} - \bar{p}_{13}^1) + a_2 r_2 (\bar{g}_{best_{13}} - \bar{p}_{13}^1)$$

$$\bar{p}_{13}^2 = \bar{p}_{13}^1 + \bar{v}_{13}^2$$

Dengan substitusi nilai:

$$\bar{v}_{13}^2 = 0.5 \cdot (-0.5) + 2.0 \cdot 0.7(1.4 - 1.4) + 2.0 \cdot 0.9(1.4 - 1.4)$$

$$\bar{p}_{13}^2 = 1.4 + \bar{v}_{13}^2$$

$$\bar{v}_{13}^2 = -0.25$$

$$\bar{p}_{13}^2 = 1.4 + (-0.25) = 1.15$$

Untuk $t = 1, i = 1, l = 4$, maka:

$$\bar{v}_{14}^2 = \omega \bar{v}_{14}^1 + a_1 r_1 (\bar{p}_{best_{14}} - \bar{p}_{14}^1) + a_2 r_2 (\bar{g}_{best_{14}} - \bar{p}_{14}^1)$$

$$\bar{p}_{14}^2 = \bar{p}_{14}^1 + \bar{v}_{14}^2$$

Dengan substitusi nilai:

$$\bar{v}_{14}^2 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(0.2 - 0.2) + 2.0 \cdot 0.9(0.2 - 0.2)$$

$$\bar{p}_{14}^2 = 0.2 + \bar{v}_{14}^2$$

$$\bar{v}_{14}^2 = 0$$

$$\bar{p}_{14}^2 = 0.2 + 0 = 0.2$$

Maka, posisi *centroid* baru untuk *cluster* 3 adalah

$$(4.9, 3.65, 1.15, 0.2)$$

Proses *update* untuk partikel 2 ($\bar{p}_2$) adalah sebagai berikut:

Cluster 1:

Posisi *centroid* saat ini: (5.2, 3.4, 1.3, 0.2)

$\bar{p}_{best}$: (5.2, 3.4, 1.3, 0.2)

$\bar{g}_{best}$: (5.1, 3.5, 1.4, 0.2) diambil dari $\bar{p}_1$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 1, i = 2, l = 1$, maka:

$$\bar{v}_{21}^2 = \omega \bar{v}_{21}^1 + a_1 r_1 (\bar{p}_{best_{21}} - \bar{p}_{21}^1) + a_2 r_2 (\bar{g}_{best_{21}} - \bar{p}_{21}^1)$$

$$\bar{p}_{21}^2 = \bar{p}_{21}^1 + \bar{v}_{21}^2$$

Dengan substitusi nilai:

$$\bar{v}_{21}^2 = 0.5 \cdot (-0.1) + 2.0 \cdot 0.7(5.2 - 5.2) + 2.0 \cdot 0.9(5.1 - 5.2)$$

$$\bar{p}_{21}^2 = 5.2 + \bar{v}_{21}^2$$

$$\bar{v}_{21}^2 = -0.23$$

$$\bar{p}_{21}^2 = 5.2 + (-0.23) = 4.97$$

Untuk $t = 1, i = 2, l = 2$, maka:

$$\bar{v}_{22}^2 = \omega \bar{v}_{22}^1 + a_1 r_1 (\bar{p}_{best_{22}} - \bar{p}_{22}^1) + a_2 r_2 (\bar{g}_{best_{22}} - \bar{p}_{22}^1)$$

$$\bar{p}_{22}^2 = \bar{p}_{22}^1 + \bar{v}_{22}^2$$

Dengan substitusi nilai:

$$\bar{v}_{22}^2 = 0.5 \cdot (0.2) + 2.0 \cdot 0.7(3.4 - 3.4) + 2.0 \cdot 0.9(3.5 - 3.4)$$

$$\bar{p}_{22}^2 = 3.4 + \bar{v}_{22}^2$$

$$\bar{v}_{22}^2 = 0.28$$

$$\bar{p}_{22}^2 = 3.4 + 0.28 = 3.68$$

Untuk $t = 1, i = 2, l = 3$, maka:

$$\bar{v}_{23}^2 = \omega \bar{v}_{23}^1 + a_1 r_1 (\bar{p}_{best_{23}} - \bar{p}_{23}^1) + a_2 r_2 (\bar{g}_{best_{23}} - \bar{p}_{23}^1)$$

$$\bar{p}_{23}^2 = \bar{p}_{23}^1 + \bar{v}_{23}^2$$

Dengan substitusi nilai:

$$\bar{v}_{23}^2 = 0.5 \cdot (-0.5) + 2.0 \cdot 0.7(1.4 - 1.4) + 2.0 \cdot 0.9(1.4 - 1.4)$$

$$\bar{p}_{23}^2 = 1.3 + \bar{v}_{23}^2$$

$$\bar{v}_{23}^2 = -0.07$$

$$\bar{p}_{23}^2 = 1.3 + (-0.07) = 1.23$$

Untuk $t = 1, i = 2, l = 4$, maka:

$$\bar{v}_{24}^2 = \omega \bar{v}_{24}^1 + a_1 r_1 (\bar{p}_{best_{24}} - \bar{p}_{24}^1) + a_2 r_2 (\bar{g}_{best_{24}} - \bar{p}_{24}^1)$$

$$\bar{p}_{24}^2 = \bar{p}_{24}^1 + \bar{v}_{24}^2$$

Dengan substitusi nilai:

$$\bar{v}_{24}^2 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(0.2 - 0.2) + 2.0 \cdot 0.9(0.2 - 0.2)$$

$$\bar{p}_{24}^2 = 0.2 + \bar{v}_{24}^2$$

$$\bar{v}_{24}^2 = 0$$

$$\bar{p}_{24}^2 = 0.2 + 0 = 0.2$$

Maka, posisi *centroid* baru untuk *cluster* 1 adalah

$$(4.97, 3.68, 1.23, 0.2)$$

Cluster 2:

Posisi *centroid* saat ini: (4.8, 3.2, 1.5, 0.2)

$$\bar{p}_{best}: (4.8, 3.2, 1.5, 0.2)$$

$$\bar{g}_{best}: (4.7, 3.1, 1.4, 0.2) \text{ diambil dari } \bar{p}_1$$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan

persamaan (3.6):

Untuk $t = 1, i = 2, l = 1$, maka:

$$\bar{v}_{21}^2 = \omega \bar{v}_{21}^1 + a_1 r_1 (\bar{p}_{best_{21}} - \bar{p}_{21}^1) + a_2 r_2 (\bar{g}_{best_{21}} - \bar{p}_{21}^1)$$

$$\bar{p}_{21}^2 = \bar{p}_{21}^1 + \bar{v}_{21}^2$$

Dengan substitusi nilai:

$$\bar{v}_{21}^2 = 0.5 \cdot (0.2) + 2.0 \cdot 0.7(4.8 - 4.8) + 2.0 \cdot 0.9(4.7 - 4.8)$$

$$\bar{p}_{21}^2 = 4.8 + \bar{v}_{21}^2$$

$$\bar{v}_{21}^2 = -0.08$$

$$\bar{p}_{21}^2 = 4.8 + (-0.08) = 4.72$$

Untuk $t = 1, i = 2, l = 2$, maka:

$$\bar{v}_{22}^2 = \omega \bar{v}_{22}^1 + a_1 r_1 (\bar{p}_{best_{22}} - \bar{p}_{22}^1) + a_2 r_2 (\bar{g}_{best_{22}} - \bar{p}_{22}^1)$$

$$\bar{p}_{22}^2 = \bar{p}_{22}^1 + \bar{v}_{22}^2$$

Dengan substitusi nilai:

$$\bar{v}_{22}^2 = 0.5 \cdot (-0.1) + 2.0 \cdot 0.7(3.2 - 3.2) + 2.0 \cdot 0.9(3.1 - 3.2)$$

$$\bar{p}_{22}^2 = 3.2 + \bar{v}_{22}^2$$

$$\bar{v}_{22}^2 = -0.23$$

$$\bar{p}_{22}^2 = 3.2 + (-0.23) = 2.97$$

Untuk $t = 1, i = 2, l = 3$, maka:

$$\bar{v}_{23}^2 = \omega \bar{v}_{23}^1 + a_1 r_1 (\bar{p}_{best_{23}} - \bar{p}_{23}^1) + a_2 r_2 (\bar{g}_{best_{23}} - \bar{p}_{23}^1)$$

$$\bar{p}_{23}^2 = \bar{p}_{23}^1 + \bar{v}_{23}^2$$

Dengan substitusi nilai:

$$\bar{v}_{23}^2 = 0.5 \cdot (0.1) + 2.0 \cdot 0.7(1.5 - 1.5) + 2.0 \cdot 0.9(1.4 - 1.5)$$

$$\bar{p}_{23}^2 = 1.5 + \bar{v}_{23}^2$$

$$\bar{v}_{23}^2 = -0.13$$

$$\bar{p}_{23}^2 = 1.5 + (-0.13) = 1.37$$

Untuk $t = 1, i = 2, l = 4$, maka:

$$\bar{v}_{24}^2 = \omega \bar{v}_{24}^1 + a_1 r_1 (\bar{p}_{best_{24}} - \bar{p}_{24}^1) + a_2 r_2 (\bar{g}_{best_{24}} - \bar{p}_{24}^1)$$

$$\bar{p}_{24}^2 = \bar{p}_{24}^1 + \bar{v}_{24}^2$$

Dengan substitusi nilai:

$$\bar{v}_{24}^2 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(0.2 - 0.2) + 2.0 \cdot 0.9(0.2 - 0.2)$$

$$\bar{p}_{24}^2 = 0.2 + \bar{v}_{24}^2$$

$$\bar{v}_{24}^2 = 0$$

$$\bar{p}_{24}^2 = 0.2 + 0 = 0.2$$

Maka, posisi *centroid* baru untuk *cluster* 2 adalah

$$(4.72, 2.97, 1.37, 0.2)$$

Cluster 3:

Posisi *centroid* saat ini: (4.9, 3.5, 1.3, 0.2)

$$\bar{p}_{best}: (4.9, 3.5, 1.3, 0.2)$$

$$\bar{g}_{best}: (5.0, 3.6, 1.4, 0.2) \text{ diambil dari } \bar{p}_1$$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 1, i = 2, l = 1$, maka:

$$\bar{v}_{21}^2 = \omega \bar{v}_{21}^1 + a_1 r_1 (\bar{p}_{best_{21}} - \bar{p}_{21}^1) + a_2 r_2 (\bar{g}_{best_{21}} - \bar{p}_{21}^1)$$

$$\bar{p}_{21}^2 = \bar{p}_{21}^1 + \bar{v}_{21}^2$$

Dengan substitusi nilai:

$$\bar{v}_{21}^2 = 0.5 \cdot (-0.2) + 2.0 \cdot 0.7(4.9 - 4.9) + 2.0 \cdot 0.9(5.0 - 4.9)$$

$$\bar{p}_{21}^2 = 4.9 + \bar{v}_{21}^2$$

$$\bar{v}_{21}^2 = 0.08$$

$$\bar{p}_{21}^2 = 4.9 + 0.08 = 4.98$$

Untuk $t = 1, i = 2, l = 2$, maka:

$$\bar{v}_{22}^2 = \omega \bar{v}_{22}^1 + a_1 r_1 (\bar{p}_{best_{22}} - \bar{p}_{22}^1) + a_2 r_2 (\bar{g}_{best_{22}} - \bar{p}_{22}^1)$$

$$\bar{p}_{22}^2 = \bar{p}_{22}^1 + \bar{v}_{22}^2$$

Dengan substitusi nilai:

$$\bar{v}_{22}^2 = 0.5 \cdot (0.1) + 2.0 \cdot 0.7(3.5 - 3.5) + 2.0 \cdot 0.9(3.6 - 3.5)$$

$$\bar{p}_{22}^2 = 3.5 + \bar{v}_{22}^2$$

$$\bar{v}_{22}^2 = 0.23$$

$$\bar{p}_{22}^2 = 3.5 + 0.23 = 3.73$$

Untuk $t = 1, i = 2, l = 3$, maka:

$$\bar{v}_{23}^2 = \omega \bar{v}_{23}^1 + a_1 r_1 (\bar{p}_{best_{23}} - \bar{p}_{23}^1) + a_2 r_2 (\bar{g}_{best_{23}} - \bar{p}_{23}^1)$$

$$\bar{p}_{23}^2 = \bar{p}_{23}^1 + \bar{v}_{23}^2$$

Dengan substitusi nilai:

$$\bar{v}_{23}^2 = 0.5 \cdot (-0.5) + 2.0 \cdot 0.7(1.3 - 1.3) + 2.0 \cdot 0.9(1.4 - 1.3)$$

$$\bar{p}_{23}^2 = 1.3 + \bar{v}_{23}^2$$

$$\bar{v}_{23}^2 = -0.07$$

$$\bar{p}_{23}^2 = 1.3 + (-0.07) = 1.23$$

Untuk $t = 1, i = 2, l = 4$, maka:

$$\bar{v}_{24}^2 = \omega \bar{v}_{24}^1 + a_1 r_1 (\bar{p}_{best_{24}} - \bar{p}_{24}^1) + a_2 r_2 (\bar{g}_{best_{24}} - \bar{p}_{24}^1)$$

$$\bar{p}_{24}^2 = \bar{p}_{24}^1 + \bar{v}_{24}^2$$

Dengan substitusi nilai:

$$\bar{v}_{24}^2 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(0.2 - 0.2) + 2.0 \cdot 0.9(0.2 - 0.2)$$

$$\bar{p}_{24}^2 = 0.2 + \bar{v}_{24}^2$$

$$\bar{v}_{24}^2 = 0$$

$$\bar{p}_{24}^2 = 0.2 + 0 = 0.2$$

Maka, posisi *centroid* baru untuk *cluster* 3 adalah

$$(4.98, 3.73, 1.23, 0.2)$$

Proses *update* untuk partikel 3 ($\bar{p}_3$) adalah sebagai berikut:

Cluster 1:

Posisi *centroid* saat ini: (5.2, 3.4, 1.3, 0.2)

$$\bar{p}_{best}: (5.2, 3.4, 1.3, 0.2)$$

$$\bar{g}_{best}: (5.1, 3.5, 1.4, 0.2) \text{ diambil dari } \bar{p}_1$$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 1, i = 3, l = 1$, maka:

$$\bar{v}_{31}^2 = \omega \bar{v}_{31}^1 + a_1 r_1 (\bar{p}_{best_{31}} - \bar{p}_{31}^1) + a_2 r_2 (\bar{g}_{best_{31}} - \bar{p}_{31}^1)$$

$$\bar{p}_{31}^2 = \bar{p}_{31}^1 + \bar{v}_{31}^2$$

Dengan substitusi nilai:

$$\bar{v}_{31}^2 = 0.5 \cdot (-0.1) + 2.0 \cdot 0.7(5.2 - 5.2) + 2.0 \cdot 0.9(5.1 - 5.2)$$

$$\bar{p}_{31}^2 = 5.2 + \bar{v}_{31}^2$$

$$\bar{v}_{31}^2 = -0.23$$

$$\bar{p}_{31}^2 = 5.2 + (-0.23) = 4.97$$

Untuk $t = 1, i = 3, l = 2$, maka:

$$\bar{v}_{32}^2 = \omega \bar{v}_{32}^1 + a_1 r_1 (\bar{p}_{best_{32}} - \bar{p}_{32}^1) + a_2 r_2 (\bar{g}_{best_{32}} - \bar{p}_{32}^1)$$

$$\bar{p}_{32}^2 = \bar{p}_{32}^1 + \bar{v}_{32}^2$$

Dengan substitusi nilai:

$$\bar{v}_{32}^2 = 0.5 \cdot (0.2) + 2.0 \cdot 0.7(3.4 - 3.4) + 2.0 \cdot 0.9(3.5 - 3.4)$$

$$\bar{p}_{32}^2 = 3.2 + \bar{v}_{32}^2$$

$$\bar{v}_{32}^2 = -0.23$$

$$\bar{p}_{32}^2 = 3.2 + (-0.23) = 2.97$$

Untuk $t = 1, i = 3, l = 3$, maka:

$$\bar{v}_{33}^2 = \omega \bar{v}_{33}^1 + a_1 r_1 (\bar{p}_{best_{33}} - \bar{p}_{33}^1) + a_2 r_2 (\bar{g}_{best_{33}} - \bar{p}_{33}^1)$$

$$\bar{p}_{33}^2 = \bar{p}_{33}^1 + \bar{v}_{33}^2$$

Dengan substitusi nilai:

$$\bar{v}_{33}^2 = 0.5 \cdot (-0.5) + 2.0 \cdot 0.7(1.3 - 1.3) + 2.0 \cdot 0.9(1.4 - 1.3)$$

$$\bar{p}_{33}^2 = 1.3 + \bar{v}_{33}^2$$

$$\bar{v}_{33}^2 = -0.07$$

$$\bar{p}_{33}^2 = 1.3 + (-0.07) = 1.23$$

Untuk $t = 1, i = 3, l = 4$, maka:

$$\bar{v}_{34}^2 = \omega \bar{v}_{34}^1 + a_1 r_1 (\bar{p}_{best_{34}} - \bar{p}_{34}^1) + a_2 r_2 (\bar{g}_{best_{34}} - \bar{p}_{34}^1)$$

$$\bar{p}_{34}^2 = \bar{p}_{34}^1 + \bar{v}_{34}^2$$

Dengan substitusi nilai:

$$\bar{v}_{34}^2 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(0.2 - 0.2) + 2.0 \cdot 0.9(0.2 - 0.2)$$

$$\bar{p}_{24}^2 = 0.2 + \bar{v}_{34}^2$$

$$\bar{v}_{34}^2 = 0$$

$$\bar{p}_{24}^2 = 0.2 + 0 = 0.2$$

Maka, posisi *centroid* baru untuk *cluster* 1 adalah

$$(4.97, 3.68, 1.23, 0.2)$$

Cluster 2:

Posisi *centroid* saat ini: (4.6, 3.0, 1.3, 0.2)

$$\bar{p}_{best}: (4.6, 3.0, 1.3, 0.2)$$

$$\bar{g}_{best}: (4.7, 3.1, 1.4, 0.2) \text{ diambil dari } \bar{p}_1$$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 1, i = 3, l = 1$, maka:

$$\bar{v}_{31}^2 = \omega \bar{v}_{31}^1 + a_1 r_1 (\bar{p}_{best_{31}} - \bar{p}_{31}^1) + a_2 r_2 (\bar{g}_{best_{31}} - \bar{p}_{31}^1)$$

$$\bar{p}_{31}^2 = \bar{p}_{31}^1 + \bar{v}_{31}^2$$

Dengan substitusi nilai:

$$\bar{v}_{31}^2 = 0.5 \cdot (0.2) + 2.0 \cdot 0.7(4.6 - 4.6) + 2.0 \cdot 0.9(4.7 - 4.6)$$

$$\bar{p}_{31}^2 = 4.6 + \bar{v}_{31}^2$$

$$\bar{v}_{31}^2 = 0.28$$

$$\bar{p}_{31}^2 = 4.6 + 0.28 = 4.88$$

Untuk $t = 1, i = 3, l = 2$, maka:

$$\bar{v}_{32}^2 = \omega \bar{v}_{32}^1 + a_1 r_1 (\bar{p}_{best_{32}} - \bar{p}_{32}^1) + a_2 r_2 (\bar{g}_{best_{32}} - \bar{p}_{32}^1)$$

$$\bar{p}_{32}^2 = \bar{p}_{32}^1 + \bar{v}_{32}^2$$

Dengan substitusi nilai:

$$\bar{v}_{32}^2 = 0.5 \cdot (-0.1) + 2.0 \cdot 0.7(3.0 - 3.0) + 2.0 \cdot 0.9(3.1 - 3.0)$$

$$\bar{p}_{32}^2 = 3.0 + \bar{v}_{32}^2$$

$$\bar{v}_{32}^2 = 0.13$$

$$\bar{p}_{32}^2 = 3.0 + 0.13 = 3.13$$

Untuk $t = 1, i = 3, l = 3$, maka:

$$\bar{v}_{33}^2 = \omega \bar{v}_{33}^1 + a_1 r_1 (\bar{p}_{best_{33}} - \bar{p}_{33}^1) + a_2 r_2 (\bar{g}_{best_{33}} - \bar{p}_{33}^1)$$

$$\bar{p}_{33}^2 = \bar{p}_{33}^1 + \bar{v}_{33}^2$$

Dengan substitusi nilai:

$$\bar{v}_{33}^2 = 0.5 \cdot (0.1) + 2.0 \cdot 0.7(1.3 - 1.3) + 2.0 \cdot 0.9(1.4 - 1.3)$$

$$\bar{p}_{33}^2 = 1.3 + \bar{v}_{33}^2$$

$$\bar{v}_{33}^2 = 0.23$$

$$\bar{p}_{33}^2 = 1.3 + 0.23 = 1.53$$

Untuk $t = 1, i = 3, l = 4$, maka:

$$\bar{v}_{34}^2 = \omega \bar{v}_{34}^1 + a_1 r_1 (\bar{p}_{best_{34}} - \bar{p}_{34}^1) + a_2 r_2 (\bar{g}_{best_{34}} - \bar{p}_{34}^1)$$

$$\bar{p}_{34}^2 = \bar{p}_{34}^1 + \bar{v}_{34}^2$$

Dengan substitusi nilai:

$$\bar{v}_{34}^2 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(0.2 - 0.2) + 2.0 \cdot 0.9(0.2 - 0.2)$$

$$\bar{p}_{24}^2 = 0.2 + \bar{v}_{34}^2$$

$$\bar{v}_{34}^2 = 0$$

$$\bar{p}_{24}^2 = 0.2 + 0 = 0.2$$

Maka, posisi *centroid* baru untuk *cluster 2* adalah

$$(4.88, 3.13, 1.53, 0.2)$$

Cluster 3:

Posisi *centroid* saat ini: $(5.1, 3.4, 1.5, 0.2)$

$$\bar{p}_{best} = (5.1, 3.4, 1.5, 0.2)$$

$$\bar{g}_{best} = (5.0, 3.6, 1.4, 0.2) \text{ dari } \bar{p}_1$$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 1, i = 3, l = 1$, maka:

$$\bar{v}_{31}^2 = \omega \bar{v}_{31}^1 + a_1 r_1 (\bar{p}_{best_{31}} - \bar{p}_{31}^1) + a_2 r_2 (\bar{g}_{best_{31}} - \bar{p}_{31}^1)$$

$$\bar{p}_{31}^2 = \bar{p}_{31}^1 + \bar{v}_{31}^2$$

Dengan substitusi nilai:

$$\bar{v}_{31}^2 = 0.5 \cdot (-0.2) + 2.0 \cdot 0.7(5.1 - 5.1) + 2.0 \cdot 0.9(5.0 - 5.1)$$

$$\bar{p}_{31}^2 = 5.1 + \bar{v}_{31}^2$$

$$\bar{v}_{31}^2 = -0.28$$

$$\bar{p}_{31}^2 = 5.1 + (-0.28) = 4.82$$

Untuk $t = 1, i = 3, l = 2$, maka:

$$\bar{v}_{32}^2 = \omega \bar{v}_{32}^1 + a_1 r_1 (\bar{p}_{best_{32}} - \bar{p}_{32}^1) + a_2 r_2 (\bar{g}_{best_{32}} - \bar{p}_{32}^1)$$

$$\bar{p}_{32}^2 = \bar{p}_{32}^1 + \bar{v}_{32}^2$$

Dengan substitusi nilai:

$$\bar{v}_{32}^2 = 0.5 \cdot (0.1) + 2.0 \cdot 0.7(3.4 - 3.4) + 2.0 \cdot 0.9(3.6 - 3.4)$$

$$\bar{p}_{32}^2 = 3.4 + \bar{v}_{32}^2$$

$$\bar{v}_{32}^2 = 0.41$$

$$\bar{p}_{32}^2 = 3.4 + 0.41 = 3.81$$

Untuk $t = 1, i = 3, l = 3$, maka:

$$\bar{v}_{33}^2 = \omega \bar{v}_{33}^1 + a_1 r_1 (\bar{p}_{best_{33}} - \bar{p}_{33}^1) + a_2 r_2 (\bar{g}_{best_{33}} - \bar{p}_{33}^1)$$

$$\bar{p}_{33}^2 = \bar{p}_{33}^1 + \bar{v}_{33}^2$$

Dengan substitusi nilai:

$$\bar{v}_{33}^2 = 0.5 \cdot (-0.5) + 2.0 \cdot 0.7(1.5 - 1.5) + 2.0 \cdot 0.9(1.4 - 1.5)$$

$$\bar{p}_{33}^2 = 1.5 + \bar{v}_{33}^2$$

$$\bar{v}_{33}^2 = -0.43$$

$$\bar{p}_{33}^2 = 1.5 + (-0.43) = 1.07$$

Untuk $t = 1, i = 3, l = 4$, maka:

$$\bar{v}_{34}^2 = \omega \bar{v}_{34}^1 + a_1 r_1 (\bar{p}_{best_{34}} - \bar{p}_{34}^1) + a_2 r_2 (\bar{g}_{best_{34}} - \bar{p}_{34}^1)$$

$$\bar{p}_{34}^2 = \bar{p}_{34}^1 + \bar{v}_{34}^2$$

Dengan substitusi nilai:

$$\bar{v}_{34}^2 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(0.2 - 0.2) + 2.0 \cdot 0.9(0.2 - 0.2)$$

$$\bar{p}_{24}^2 = 0.2 + \bar{v}_{34}^2$$

$$\bar{v}_{34}^2 = 0$$

$$\bar{p}_{24}^2 = 0.2 + 0 = 0.2$$

Maka, posisi *centroid* baru untuk *cluster* 3 adalah

$$(4.82, 3.81, 1.07, 0.2)$$

Hasil pembaruan posisi partikel setelah iterasi 1

Setelah dilakukan proses perhitungan *update* kecepatan dan posisi pada masing-masing partikel sesuai dengan persamaan perhitungan PSO, diperoleh hasil *centroid* terbaru untuk setiap partikel pada tabel 4.12 sebagai berikut:

Tabel 4.12 Hasil Iterasi 1 (*centroid* setelah *update*)

$(\bar{p}_i)$	$\bar{\mu}_1$	$\bar{\mu}_2$	$\bar{\mu}_3$
$\bar{p}_1$	(5.05, 3.6, 1.15, 0.2)	(4.8, 3.05, 1.45, 0.2)	(4.9, 3.65, 1.15, 0.2)
$\bar{p}_2$	(4.97, 3.68, 1.23, 0.2)	(4.72, 2.97, 1.37, 0.2)	(4.98, 3.73, 1.23, 0.2)
$\bar{p}_3$	(4.97, 3.68, 1.23, 0.2)	(4.88, 3.13, 1.53, 0.2)	(4.82, 3.81, 1.07, 0.2)

Centroid-centroid baru ini akan digunakan untuk proses evaluasi fungsi objektif (SSE) pada iterasi selanjutnya.

Iterasi 2

Pada iterasi ini di mulai dengan menghitung nilai SSE menggunakan jarak *Euclidean*.

1. Menghitung SSE (Fungsi Objektif)

a. Hitung SSE untuk Partikel 1

Centroid awal (partikel 1) diambil dari tabel 4.8.

$$\bar{\mu}_1 = (5.05, 3.6, 1.15, 0.2)$$

$$\bar{\mu}_2 = (4.8, 3.05, 1.45, 0.2)$$

$$\bar{\mu}_3 = (4.9, 3.65, 1.15, 0.2)$$

Misalkan menggunakan 5 data pertama dari dataset Iris sesuai pada data tabel 4.1, kemudian menghitung jarak *Euclidean* dimulai antara data pertama ($\bar{x}_1$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_1, \bar{\mu}_1) = \|\bar{x}_1 - \bar{\mu}_1\| = 0.273$$

$$d(\bar{x}_1, \bar{\mu}_2) = \|\bar{x}_1 - \bar{\mu}_2\| = 0.542$$

$$d(\bar{x}_1, \bar{\mu}_3) = \|\bar{x}_1 - \bar{\mu}_3\| = 0.353$$

Kemudian dilanjut dengan perhitungan jarak *Euclidean* antara data kedua ($\bar{x}_2$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_2, \bar{\mu}_1) = \|\bar{x}_2 - \bar{\mu}_1\| = 0.667$$

$$d(\bar{x}_2, \bar{\mu}_2) = \|\bar{x}_2 - \bar{\mu}_2\| = 0.122$$

$$d(\bar{x}_2, \bar{\mu}_3) = \|\bar{x}_2 - \bar{\mu}_3\| = 0.696$$

Kemudian dilanjut dengan perhitungan jarak *Euclidean* antara data ketiga ($\bar{x}_3$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_3, \bar{\mu}_1) = \|\bar{x}_3 - \bar{\mu}_1\| = 0.552$$

$$d(\bar{x}_3, \bar{\mu}_2) = \|\bar{x}_3 - \bar{\mu}_2\| = 0.234$$

$$d(\bar{x}_3, \bar{\mu}_3) = \|\bar{x}_3 - \bar{\mu}_3\| = 0.514$$

Kemudian dilanjut dengan perhitungan jarak *Euclidean* antara data keempat ($\bar{x}_4$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_4, \bar{\mu}_1) = \|\bar{x}_4 - \bar{\mu}_1\| = 0.758$$

$$d(\bar{x}_4, \bar{\mu}_2) = \|\bar{x}_4 - \bar{\mu}_2\| = 0.212$$

$$d(\bar{x}_4, \bar{\mu}_3) = \|\bar{x}_4 - \bar{\mu}_3\| = 0.717$$

Terakhir dilanjut dengan perhitungan jarak *Euclidean* antara data kelima ($\bar{x}_5$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_5, \bar{\mu}_1) = \|\bar{x}_5 - \bar{\mu}_1\| = 0.254$$

$$d(\bar{x}_5, \bar{\mu}_2) = \|\bar{x}_5 - \bar{\mu}_2\| = 0.587$$

$$d(\bar{x}_5, \bar{\mu}_3) = \|\bar{x}_5 - \bar{\mu}_3\| = 0.273$$

Setelah dihitung jarak *Euclidean*, kemudian dikelompokkan ke dalam *cluster* berdasarkan jarak terkecil.

Tabel 4.13 Pengelompokan *Cluster* Partikel 1 pada Iterasi 2

$\bar{x}_i$	$d(\bar{\mu}_1)$	$d(\bar{\mu}_2)$	$d(\bar{\mu}_3)$	Cluster
$\bar{x}_1$	0.273	0.542	0.353	1
$\bar{x}_2$	0.667	0.122	0.696	2
$\bar{x}_3$	0.552	0.234	0.514	2
$\bar{x}_4$	0.758	0.212	0.717	2
$\bar{x}_5$	0.254	0.587	0.273	1

Setelah semua data masuk ke *cluster* masing-masing, kemudian menghitung SSE dengan persamaan (3.4).

- Untuk *cluster* 1 yang terdiri dari data $C_1 = \{\bar{x}_1, \bar{x}_5\}$,

$$\bar{x}_1 = (5.1, 3.5, 1.4, 0.2) \text{ dengan centroid } \bar{\mu}_1 = \{5.05, 3.6, 1.15, 0.2\},$$

diperoleh:

$$SE_1: (5.1 - 5.05)^2 + (3.5 - 3.6)^2 + (1.4 - 1.15)^2 + (0.2 - 0.2)^2 = 0.075$$

$$\bar{x}_5 = (5, 3.6, 1.4, 0.2) \text{ dengan centroid } \bar{\mu}_1 = \{5.05, 3.6, 1.15, 0.2\},$$

diperoleh:

$$SE_5: (5 - 5.05)^2 + (3.6 - 3.6)^2 + (1.4 - 1.15)^2 + (0.2 - 0.2)^2 = 0.065$$

- Untuk *cluster* 2 yang terdiri dari data $C_2 = \{\bar{x}_2, \bar{x}_3, \bar{x}_4\}$,

$$\bar{x}_2 = (4.9, 3.0, 1.4, 0.2) \text{ dengan centroid } \bar{\mu}_2 = \{4.8, 3.05, 1.45, 0.2\},$$

diperoleh:

$$SE_2: (4.9 - 4.8)^2 + (3.0 - 3.05)^2 + (1.4 - 1.45)^2 + (0.2 - 0.2)^2 = 0.015$$

$\bar{x}_3 = (4.7, 3.2, 1.3, 0.2)$ dengan *centroid* $\bar{\mu}_2 = \{4.8, 3.05, 1.45, 0.2\}$,

diperoleh:

$$SE_3: (4.7 - 4.8)^2 + (3.2 - 3.05)^2 + (1.3 - 1.45)^2 + (0.2 - 0.2)^2 = 0.055$$

$\bar{x}_4 = (4.6, 3.1, 1.5, 0.2)$ dengan *centroid* $\bar{\mu}_2 = \{4.8, 3.05, 1.45, 0.2\}$,

diperoleh:

$$SE_4: (4.6 - 4.8)^2 + (3.1 - 3.05)^2 + (1.5 - 1.45)^2 + (0.2 - 0.2)^2 = 0.045$$

- Untuk *cluster* 3 bernilai 0 karena tidak ada data yang masuk ke dalam

C_3

Dengan demikian,

$$\begin{aligned} total\ SSE &= SE_1 + SE_2 + SE_3 + SE_4 + SE_5 \\ &= 0.075 + 0.015 + 0.055 + 0.045 + 0.065 = 0.255 \end{aligned}$$

b. Hitung SSE untuk Partikel 2

Centroid awal (partikel 2) diambil dari tabel 4.8.

$$\bar{\mu}_1 = (4.97, 3.68, 1.23, 0.2)$$

$$\bar{\mu}_2 = (4.72, 2.97, 1.37, 0.2)$$

$$\bar{\mu}_3 = (4.98, 3.73, 1.23, 0.2)$$

Misalkan menggunakan 5 data pertama dari dataset Iris sesuai pada data

tabel 4.1, kemudian menghitung jarak *Euclidean*:

$$d(\bar{x}_1, \bar{\mu}_1) = \|\bar{x}_1 - \bar{\mu}_1\| = 0.279$$

$$d(\bar{x}_1, \bar{\mu}_2) = \|\bar{x}_1 - \bar{\mu}_2\| = 0.658$$

$$d(\bar{x}_1, \bar{\mu}_3) = \|\bar{x}_1 - \bar{\mu}_3\| = 0.310$$

Kemudian dilanjutkan dengan perhitungan jarak *Euclidean* antara data kedua ($\bar{x}_2$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_2, \bar{\mu}_1) = \|\bar{x}_2 - \bar{\mu}_1\| = 0.704$$

$$d(\bar{x}_2, \bar{\mu}_2) = \|\bar{x}_2 - \bar{\mu}_2\| = 0.184$$

$$d(\bar{x}_2, \bar{\mu}_3) = \|\bar{x}_2 - \bar{\mu}_3\| = 0.753$$

Kemudian dilanjutkan dengan perhitungan jarak *Euclidean* antara data ketiga ($\bar{x}_3$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_3, \bar{\mu}_1) = \|\bar{x}_3 - \bar{\mu}_1\| = 0.555$$

$$d(\bar{x}_3, \bar{\mu}_2) = \|\bar{x}_3 - \bar{\mu}_2\| = 0.241$$

$$d(\bar{x}_3, \bar{\mu}_3) = \|\bar{x}_3 - \bar{\mu}_3\| = 0.603$$

Kemudian dilanjutkan dengan perhitungan jarak *Euclidean* antara data keempat ($\bar{x}_4$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_4, \bar{\mu}_1) = \|\bar{x}_4 - \bar{\mu}_1\| = 0.738$$

$$d(\bar{x}_4, \bar{\mu}_2) = \|\bar{x}_4 - \bar{\mu}_2\| = 0.219$$

$$d(\bar{x}_4, \bar{\mu}_3) = \|\bar{x}_4 - \bar{\mu}_3\| = 0.783$$

Terakhir dilanjutkan dengan perhitungan jarak *Euclidean* antara data kelima ($\bar{x}_5$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_5, \bar{\mu}_1) = \|\bar{x}_5 - \bar{\mu}_1\| = 0.190$$

$$d(\bar{x}_5, \bar{\mu}_2) = \|\bar{x}_5 - \bar{\mu}_2\| = 0.690$$

$$d(\bar{x}_5, \bar{\mu}_3) = \|\bar{x}_5 - \bar{\mu}_3\| = 0.215$$

Setelah dihitung jarak *Euclidean*, kemudian dikelompokkan ke dalam *cluster* berdasarkan jarak terkecil.

Tabel 4.14 Pengelompokan *Cluster* Partikel 2 pada Iterasi 2

$\bar{x}_i$	$d(\bar{\mu}_1)$	$d(\bar{\mu}_2)$	$d(\bar{\mu}_3)$	<i>Cluster</i>
$\bar{x}_1$	0.279	0.658	0.310	1
$\bar{x}_2$	0.704	0.184	0.753	2
$\bar{x}_3$	0.555	0.241	0.603	2
$\bar{x}_4$	0.738	0.219	0.783	2
$\bar{x}_5$	0.190	0.690	0.215	1

Setelah semua data masuk ke *cluster* masing-masing, kemudian menghitung SSE dengan persamaan (3.4).

- Untuk *cluster* 1 yang terdiri dari data $C_1 = \{\bar{x}_1, \bar{x}_5\}$,

$\bar{x}_1 = (5.1, 3.5, 1.4, 0.2)$ dengan *centroid* $\bar{\mu}_1 = \{4.97, 3.68, 1.23, 0.2\}$,

diperoleh:

$$SE_1: (5.1 - 4.97)^2 + (3.5 - 3.68)^2 + (1.4 - 1.23)^2 + (0.2 - 0.2)^2 = 0.078$$

$\bar{x}_5 = (5, 3.6, 1.4, 0.2)$ dengan *centroid* $\bar{\mu}_1 = \{4.97, 3.68, 1.23, 0.2\}$,

diperoleh:

$$SE_5: (5 - 4.97)^2 + (3.6 - 3.68)^2 + (1.4 - 1.23)^2 + (0.2 - 0.2)^2 = 0.036$$

- Untuk *cluster* 2 yang terdiri dari data $C_2 = \{\bar{x}_2, \bar{x}_3, \bar{x}_4\}$,

$\bar{x}_2 = (4.9, 3.0, 1.4, 0.2)$ dengan *centroid* $\bar{\mu}_2 = \{4.72, 2.97, 1.37, 0.2\}$,

diperoleh:

$$SE_2: (4.9 - 4.72)^2 + (3.0 - 2.97)^2 + (1.4 - 1.37)^2 + (0.2 - 0.2)^2 = 0.034$$

$\bar{x}_3 = (4.7, 3.2, 1.3, 0.2)$ dengan *centroid* $\bar{\mu}_2 = \{4.72, 2.97, 1.37, 0.2\}$,

diperoleh:

$$SE_3: (4.7 - 4.72)^2 + (3.2 - 2.97)^2 + (1.3 - 1.37)^2 + (0.2 - 0.2)^2 = 0.058$$

$\bar{x}_4 = (4.6, 3.1, 1.5, 0.2)$ dengan *centroid* $\bar{\mu}_2 = \{4.72, 2.97, 1.37, 0.2\}$,

diperoleh:

$$SE_4: (4.6 - 4.72)^2 + (3.1 - 2.97)^2 + (1.5 - 1.37)^2 + (0.2 - 0.2)^2 = 0.048$$

- Untuk *cluster* 3 bernilai 0 karena tidak ada data yang masuk ke dalam

C_3

Dengan demikian,

$$\begin{aligned} total\ SSE &= SE_1 + SE_2 + SE_3 + SE_4 + SE_5 \\ &= 0.078 + 0.034 + 0.058 + 0.048 + 0.036 = 0.254 \end{aligned}$$

c. Hitung SSE untuk Partikel 3

Centroid awal (partikel 3) diambil dari tabel 4.8.

$$\bar{\mu}_1 = (4.97, 3.68, 1.23, 0.2)$$

$$\bar{\mu}_2 = (4.88, 3.13, 1.53, 0.2)$$

$$\bar{\mu}_3 = (4.82, 3.81, 1.07, 0.2)$$

Misalkan menggunakan 5 data pertama dari dataset Iris sesuai pada data tabel 4.1, kemudian menghitung jarak *Euclidean* dimulai antara data pertama ($\bar{x}_1$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_1, \bar{\mu}_1) = \|\bar{x}_1 - \bar{\mu}_1\| = 0.279$$

$$d(\bar{x}_1, \bar{\mu}_2) = \|\bar{x}_1 - \bar{\mu}_2\| = 0.449$$

$$d(\bar{x}_1, \bar{\mu}_3) = \|\bar{x}_1 - \bar{\mu}_3\| = 0.532$$

Kemudian dilanjut dengan perhitungan jarak *Euclidean* antara data kedua ($\bar{x}_2$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_2, \bar{\mu}_1) = \|\bar{x}_2 - \bar{\mu}_1\| = 0.704$$

$$d(\bar{x}_2, \bar{\mu}_2) = \|\bar{x}_2 - \bar{\mu}_2\| = 0.184$$

$$d(\bar{x}_2, \bar{\mu}_3) = \|\bar{x}_2 - \bar{\mu}_3\| = 0.878$$

Kemudian dilanjut dengan perhitungan jarak *Euclidean* antara data ketiga ($\bar{x}_3$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_3, \bar{\mu}_1) = \|\bar{x}_3 - \bar{\mu}_1\| = 0.555$$

$$d(\bar{x}_3, \bar{\mu}_2) = \|\bar{x}_3 - \bar{\mu}_2\| = 0.300$$

$$d(\bar{x}_3, \bar{\mu}_3) = \|\bar{x}_3 - \bar{\mu}_3\| = 0.662$$

Kemudian dilanjut dengan perhitungan jarak *Euclidean* antara data keempat ($\bar{x}_4$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_4, \bar{\mu}_1) = \|\bar{x}_4 - \bar{\mu}_1\| = 0.738$$

$$d(\bar{x}_4, \bar{\mu}_2) = \|\bar{x}_4 - \bar{\mu}_2\| = 0.283$$

$$d(\bar{x}_4, \bar{\mu}_3) = \|\bar{x}_4 - \bar{\mu}_3\| = 0.858$$

Terakhir dilanjut dengan perhitungan jarak *Euclidean* antara data kelima ($\bar{x}_5$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_5, \bar{\mu}_1) = \|\bar{x}_5 - \bar{\mu}_1\| = 0.190$$

$$d(\bar{x}_5, \bar{\mu}_2) = \|\bar{x}_5 - \bar{\mu}_2\| = 0.502$$

$$d(\bar{x}_5, \bar{\mu}_3) = \|\bar{x}_5 - \bar{\mu}_3\| = 0.430$$

Setelah dihitung jarak *Euclidean*, kemudian dikelompokkan ke dalam *cluster* berdasarkan jarak terkecil.

Tabel 4.15 Pengelompokan *Cluster* Partikel 3 pada Iterasi 2

$\bar{x}_i$	$d(\bar{\mu}_1)$	$d(\bar{\mu}_2)$	$d(\bar{\mu}_3)$	<i>Cluster</i>
$\bar{x}_1$	0.279	0.449	0.532	1
$\bar{x}_2$	0.704	0.184	0.878	2
$\bar{x}_3$	0.555	0.300	0.662	2
$\bar{x}_4$	0.738	0.283	0.858	2
$\bar{x}_5$	0.190	0.502	0.430	1

Setelah semua data masuk ke *cluster* masing-masing, kemudian menghitung SSE dengan persamaan (3.4).

- Untuk *cluster* 1 yang terdiri dari data $C_1 = \{\bar{x}_1, \bar{x}_5\}$,

$\bar{x}_1 = (5.1, 3.5, 1.4, 0.2)$ dengan *centroid* $\bar{\mu}_1 = \{4.97, 3.68, 1.23, 0.2\}$,

diperoleh:

$$SE_1: (5.1 - 4.97)^2 + (3.5 - 3.68)^2 + (1.4 - 1.23)^2 + (0.2 - 0.2)^2 = 0.078$$

$\bar{x}_5 = (5, 3.6, 1.4, 0.2)$ dengan *centroid* $\bar{\mu}_1 = \{4.97, 3.68, 1.23, 0.2\}$,

diperoleh:

$$SE_5: (5 - 4.97)^2 + (3.6 - 3.68)^2 + (1.4 - 1.23)^2 + (0.2 - 0.2)^2 = 0.036$$

- Untuk *cluster* 2 yang terdiri dari data $C_2 = \{\bar{x}_2, \bar{x}_3, \bar{x}_4\}$,

$\bar{x}_2 = (4.9, 3.0, 1.4, 0.2)$ dengan *centroid* $\bar{\mu}_2 = \{4.88, 3.13, 1.53, 0.2\}$,

diperoleh:

$$SE_2: (4.9 - 4.88)^2 + (3.0 - 3.13)^2 + (1.4 - 1.53)^2 + (0.2 - 0.2)^2 = 0.034$$

$\bar{x}_3 = (4.7, 3.2, 1.3, 0.2)$ dengan *centroid* $\bar{\mu}_2 = \{4.88, 3.13, 1.53, 0.2\}$,

diperoleh:

$$SE_3: (4.7 - 4.88)^2 + (3.2 - 3.13)^2 + (1.3 - 1.53)^2 + (0.2 - 0.2)^2 = 0.090$$

$\bar{x}_4 = (4.6, 3.1, 1.5, 0.2)$ dengan *centroid* $\bar{\mu}_2 = \{4.88, 3.13, 1.53, 0.2\}$,

diperoleh:

$$SE_4: (4.6 - 4.88)^2 + (3.1 - 3.13)^2 + (1.5 - 1.53)^2 + (0.2 - 0.2)^2 = 0.080$$

- Untuk *cluster* 3 bernilai 0 karena tidak ada data yang masuk ke dalam

C_3

Dengan demikian,

$$\begin{aligned} total\ SSE &= SE_1 + SE_2 + SE_3 + SE_4 + SE_5 \\ &= 0.078 + 0.034 + 0.090 + 0.080 + 0.036 = 0.318 \end{aligned}$$

2. *Update* Kecepatan dan Posisi Iterasi kedua

Karena nilai SSE pada iterasi kedua lebih besar dibandingkan iterasi pertama, maka:

- $\bar{p}_{best}$ tetap sama sesuai iterasi pertama
- $\bar{g}_{best}$ adalah $\bar{p}_{best}$ terbaik di antara semua partikel, yaitu partikel 1 ($\bar{p}_1$)

Proses *update* partikel 1 ($\bar{p}_1$) adalah sebagai berikut:

Cluster 1:

Posisi *centroid* saat ini: (5.05, 3.6, 1.15, 0.2)

$\bar{p}_{best}$: (5.1, 3.5, 1.4, 0.2)

$\bar{g}_{best}$: (5.1, 3.5, 1.4, 0.2) diambil dari $\bar{p}_1$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 2, i = 1, l = 1$, maka:

$$\bar{v}_{11}^3 = \omega \bar{v}_{11}^2 + a_1 r_1 (\bar{p}_{best_{11}} - \bar{p}_{11}^2) + a_2 r_2 (\bar{g}_{best_{11}} - \bar{p}_{11}^2)$$

$$\bar{p}_{11}^3 = \bar{p}_{11}^2 + \bar{v}_{11}^3$$

Dengan substitusi nilai:

$$\bar{v}_{11}^3 = 0.5 \cdot (-0.1) + 2.0 \cdot 0.7(5.1 - 5.05) + 2.0 \cdot 0.9(5.1 - 5.05)$$

$$\bar{p}_{11}^3 = 5.05 + \bar{v}_{11}^3$$

$$\bar{v}_{11}^3 = 0.11$$

$$\bar{p}_{11}^3 = 5.05 + 0.11 = 5.16$$

Untuk $t = 2, i = 1, l = 2$, maka:

$$\bar{v}_{12}^3 = \omega \bar{v}_{12}^2 + a_1 r_1 (\bar{p}_{best_{12}} - \bar{p}_{12}^2) + a_2 r_2 (\bar{g}_{best_{12}} - \bar{p}_{12}^2)$$

$$\bar{p}_{12}^3 = \bar{p}_{12}^2 + \bar{v}_{12}^3$$

Dengan substitusi nilai:

$$\bar{v}_{12}^3 = 0.5 \cdot (0.2) + 2.0 \cdot 0.7(3.5 - 3.6) + 2.0 \cdot 0.9(3.5 - 3.6)$$

$$\bar{p}_{12}^3 = 3.6 + \bar{v}_{12}^3$$

$$\bar{v}_{12}^3 = -0.22$$

$$\bar{p}_{12}^3 = 3.6 + (-0.22) = 3.38$$

Untuk $t = 2, i = 1, l = 3$, maka:

$$\bar{v}_{13}^3 = \omega \bar{v}_{13}^2 + a_1 r_1 (\bar{p}_{best_{13}} - \bar{p}_{13}^2) + a_2 r_2 (\bar{g}_{best_{13}} - \bar{p}_{13}^2)$$

$$\bar{p}_{13}^3 = \bar{p}_{13}^2 + \bar{v}_{13}^3$$

Dengan substitusi nilai:

$$\bar{v}_{13}^3 = 0.5 \cdot (-0.5) + 2.0 \cdot 0.7(1.4 - 1.15) + 2.0 \cdot 0.9(1.4 - 1.15)$$

$$\bar{p}_{13}^3 = 1.15 + \bar{v}_{13}^3$$

$$\bar{v}_{13}^3 = 0.55$$

$$\bar{p}_{13}^3 = 1.15 + 0.55 = 1.70$$

Untuk $t = 2, i = 1, l = 4$, maka:

$$\bar{v}_{14}^3 = \omega \bar{v}_{14}^2 + a_1 r_1 (\bar{p}_{best_{14}} - \bar{p}_{14}^2) + a_2 r_2 (\bar{g}_{best_{14}} - \bar{p}_{14}^2)$$

$$\bar{p}_{14}^3 = \bar{p}_{14}^2 + \bar{v}_{14}^3$$

Dengan substitusi nilai:

$$\bar{v}_{14}^3 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(0.2 - 0.2) + 2.0 \cdot 0.9(0.2 - 0.2)$$

$$\bar{p}_{14}^3 = 0.2 + \bar{v}_{14}^3$$

$$\bar{v}_{14}^3 = 0$$

$$\bar{p}_{14}^3 = 0.2 + 0 = 0.2$$

Maka, posisi *centroid* baru untuk *cluster* 1 adalah

$$(5.16, 3.38, 1.70, 0.2)$$

Cluster 2:

Posisi *centroid* saat ini: (4.8, 3.05, 1.45, 0.2)

$$\bar{p}_{best}: (4.7, 3.1, 1.4, 0.2)$$

$$\bar{g}_{best}: (4.7, 3.1, 1.4, 0.2) \text{ diambil dari } \bar{p}_1$$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 2, i = 1, l = 1$, maka:

$$\bar{v}_{11}^3 = \omega \bar{v}_{11}^2 + a_1 r_1 (\bar{p}_{best_{11}} - \bar{p}_{11}^2) + a_2 r_2 (\bar{g}_{best_{11}} - \bar{p}_{11}^2)$$

$$\bar{p}_{11}^3 = \bar{p}_{11}^2 + \bar{v}_{11}^3$$

Dengan substitusi nilai:

$$\bar{v}_{11}^3 = 0.5 \cdot (0.2) + 2.0 \cdot 0.7(4.7 - 4.8) + 2.0 \cdot 0.9(4.7 - 4.8)$$

$$\bar{p}_{11}^3 = 4.8 + \bar{v}_{11}^3$$

$$\bar{v}_{11}^3 = -0.22$$

$$\bar{p}_{11}^3 = 4.8 + (-0.22) = 4.58$$

Untuk $t = 2, i = 1, l = 2$, maka:

$$\bar{v}_{12}^3 = \omega \bar{v}_{12}^2 + a_1 r_1 (\bar{p}_{best_{12}} - \bar{p}_{12}^2) + a_2 r_2 (\bar{g}_{best_{12}} - \bar{p}_{12}^2)$$

$$\bar{p}_{12}^3 = \bar{p}_{12}^2 + \bar{v}_{12}^3$$

Dengan substitusi nilai:

$$\bar{v}_{12}^3 = 0.5 \cdot (-0.1) + 2.0 \cdot 0.7(3.1 - 3.05) + 2.0 \cdot 0.9(3.1 - 3.05)$$

$$\bar{p}_{12}^3 = 3.05 + \bar{v}_{12}^3$$

$$\bar{v}_{12}^3 = 0.11$$

$$\bar{p}_{12}^3 = 3.05 + 0.11 = 3.16$$

Untuk $t = 2, i = 1, l = 3$, maka:

$$\bar{v}_{13}^3 = \omega \bar{v}_{13}^2 + a_1 r_1 (\bar{p}_{best_{13}} - \bar{p}_{13}^2) + a_2 r_2 (\bar{g}_{best_{13}} - \bar{p}_{13}^2)$$

$$\bar{p}_{13}^3 = \bar{p}_{13}^2 + \bar{v}_{13}^3$$

Dengan substitusi nilai:

$$\bar{v}_{13}^3 = 0.5 \cdot (0.1) + 2.0 \cdot 0.7(1.4 - 1.45) + 2.0 \cdot 0.9(1.4 - 1.45)$$

$$\bar{p}_{13}^3 = 1.45 + \bar{v}_{13}^3$$

$$\bar{v}_{13}^3 = -0.11$$

$$\bar{p}_{13}^3 = 1.45 + (-0.11) = 1.34$$

Untuk $t = 2, i = 1, l = 4$, maka:

$$\bar{v}_{14}^3 = \omega \bar{v}_{14}^2 + a_1 r_1 (\bar{p}_{best_{14}} - \bar{p}_{14}^2) + a_2 r_2 (\bar{g}_{best_{14}} - \bar{p}_{14}^2)$$

$$\bar{p}_{14}^3 = \bar{p}_{14}^2 + \bar{v}_{14}^3$$

Dengan substitusi nilai:

$$\bar{v}_{14}^3 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(0.2 - 0.2) + 2.0 \cdot 0.9(0.2 - 0.2)$$

$$\bar{p}_{14}^3 = 0.2 + \bar{v}_{14}^3$$

$$\bar{v}_{14}^3 = 0$$

$$\bar{p}_{14}^3 = 0.2 + 0 = 0.2$$

Maka, posisi *centroid* baru untuk *cluster* 2 adalah

$$(4.58, 3.16, 1.34, 0.2)$$

Cluster 3:

Posisi *centroid* saat ini: (4.9, 3.65, 1.15, 0.2)

$$\bar{p}_{best}: (5.0, 3.6, 1.4, 0.2)$$

$$\bar{g}_{best}: (5.0, 3.6, 1.4, 0.2) \text{ diambil dari } \bar{p}_1$$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 2, i = 1, l = 1$, maka:

$$\bar{v}_{11}^3 = \omega \bar{v}_{11}^2 + a_1 r_1 (\bar{p}_{best_{11}} - \bar{p}_{11}^2) + a_2 r_2 (\bar{g}_{best_{11}} - \bar{p}_{11}^2)$$

$$\bar{p}_{11}^3 = \bar{p}_{11}^2 + \bar{v}_{11}^3$$

Dengan substitusi nilai:

$$\bar{v}_{11}^3 = 0.5 \cdot (-0.2) + 2.0 \cdot 0.7(5.0 - 4.9) + 2.0 \cdot 0.9(5.0 - 4.9)$$

$$\bar{p}_{11}^3 = 4.9 + \bar{v}_{11}^3$$

$$\bar{v}_{11}^3 = 0.22$$

$$\bar{p}_{11}^3 = 4.9 + 0.22 = 5.12$$

Untuk $t = 2, i = 1, l = 2$, maka:

$$\bar{v}_{12}^3 = \omega \bar{v}_{12}^2 + a_1 r_1 (\bar{p}_{best_{12}} - \bar{p}_{12}^2) + a_2 r_2 (\bar{g}_{best_{12}} - \bar{p}_{12}^2)$$

$$\bar{p}_{12}^3 = \bar{p}_{12}^2 + \bar{v}_{12}^3$$

Dengan substitusi nilai:

$$\bar{v}_{12}^3 = 0.5 \cdot (0.1) + 2.0 \cdot 0.7(3.6 - 3.65) + 2.0 \cdot 0.9(3.6 - 3.65)$$

$$\bar{p}_{12}^3 = 3.65 + \bar{v}_{12}^3$$

$$\bar{v}_{12}^3 = -0.11$$

$$\bar{p}_{12}^3 = 3.65 + (-0.11) = 3.54$$

Untuk $t = 2, i = 1, l = 3$, maka:

$$\bar{v}_{13}^3 = \omega \bar{v}_{13}^2 + a_1 r_1 (\bar{p}_{best_{13}} - \bar{p}_{13}^2) + a_2 r_2 (\bar{g}_{best_{13}} - \bar{p}_{13}^2)$$

$$\bar{p}_{13}^3 = \bar{p}_{13}^2 + \bar{v}_{13}^3$$

Dengan substitusi nilai:

$$\bar{v}_{13}^3 = 0.5 \cdot (-0.5) + 2.0 \cdot 0.7(1.4 - 1.15) + 2.0 \cdot 0.9(1.4 - 1.15)$$

$$\bar{p}_{13}^3 = 1.15 + \bar{v}_{13}^3$$

$$\bar{v}_{13}^3 = 0.55$$

$$\bar{p}_{13}^3 = 1.15 + 0.55 = 1.70$$

Untuk $t = 2, i = 1, l = 4$, maka:

$$\bar{v}_{14}^3 = \omega \bar{v}_{14}^2 + a_1 r_1 (\bar{p}_{best_{14}} - \bar{p}_{14}^2) + a_2 r_2 (\bar{g}_{best_{14}} - \bar{p}_{14}^2)$$

$$\bar{p}_{14}^3 = \bar{p}_{14}^2 + \bar{v}_{14}^3$$

Dengan substitusi nilai:

$$\bar{v}_{14}^3 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(0.2 - 0.2) + 2.0 \cdot 0.9(0.2 - 0.2)$$

$$\bar{p}_{14}^3 = 0.2 + \bar{v}_{14}^3$$

$$\bar{v}_{14}^3 = 0$$

$$\bar{p}_{14}^3 = 0.2 + 0 = 0.2$$

Maka, posisi *centroid* baru untuk *cluster* 3 adalah

$$(5.12, 3.54, 1.70, 0.2)$$

Proses *update* untuk partikel 2 ($\bar{p}_2$) adalah sebagai berikut:

Cluster 1:

Posisi *centroid* saat ini: (4.97, 3.68, 1.23, 0.2)

$$\bar{p}_{best}: (5.2, 3.4, 1.3, 0.2)$$

$$\bar{g}_{best}: (5.1, 3.5, 1.4, 0.2) \text{ diambil dari } \bar{p}_1$$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 2, i = 2, l = 1$, maka:

$$\bar{v}_{21}^3 = \omega \bar{v}_{21}^2 + a_1 r_1 (\bar{p}_{best_{21}} - \bar{p}_{21}^2) + a_2 r_2 (\bar{g}_{best_{21}} - \bar{p}_{21}^2)$$

$$\bar{p}_{21}^3 = \bar{p}_{21}^2 + \bar{v}_{21}^3$$

Dengan substitusi nilai:

$$\bar{v}_{21}^3 = 0.5 \cdot (-0.1) + 2.0 \cdot 0.7(5.2 - 4.97) + 2.0 \cdot 0.9(5.1 - 4.97)$$

$$\bar{p}_{21}^3 = 4.97 + \bar{v}_{21}^3$$

$$\bar{v}_{21}^3 = 0.51$$

$$\bar{p}_{21}^3 = 4.97 + 0.51 = 5.48$$

Untuk $t = 2, i = 2, l = 2$, maka:

$$\bar{v}_{22}^3 = \omega \bar{v}_{22}^2 + a_1 r_1 (\bar{p}_{best_{22}} - \bar{p}_{22}^2) + a_2 r_2 (\bar{g}_{best_{22}} - \bar{p}_{22}^2)$$

$$\bar{p}_{22}^3 = \bar{p}_{22}^2 + \bar{v}_{22}^3$$

Dengan substitusi nilai:

$$\bar{v}_{22}^3 = 0.5 \cdot (0.2) + 2.0 \cdot 0.7(3.4 - 3.68) + 2.0 \cdot 0.9(3.5 - 3.68)$$

$$\bar{p}_{22}^3 = 3.68 + \bar{v}_{22}^3$$

$$\bar{v}_{22}^3 = -0.62$$

$$\bar{p}_{22}^3 = 3.68 + (-0.62) = 3.06$$

Untuk $t = 2, i = 2, l = 3$, maka:

$$\bar{v}_{23}^3 = \omega \bar{v}_{23}^2 + a_1 r_1 (\bar{p}_{best_{23}} - \bar{p}_{23}^2) + a_2 r_2 (\bar{g}_{best_{23}} - \bar{p}_{23}^2)$$

$$\bar{p}_{23}^3 = \bar{p}_{23}^2 + \bar{v}_{23}^3$$

Dengan substitusi nilai:

$$\bar{v}_{23}^3 = 0.5 \cdot (-0.5) + 2.0 \cdot 0.7(1.3 - 1.23) + 2.0 \cdot 0.9(1.4 - 1.23)$$

$$\bar{p}_{23}^3 = 1.23 + \bar{v}_{23}^3$$

$$\bar{v}_{23}^3 = 0.15$$

$$\bar{p}_{23}^3 = 1.23 + 0.15 = 1.38$$

Untuk $t = 2, i = 2, l = 4$, maka:

$$\bar{v}_{24}^3 = \omega \bar{v}_{24}^2 + a_1 r_1 (\bar{p}_{best_{24}} - \bar{p}_{24}^2) + a_2 r_2 (\bar{g}_{best_{24}} - \bar{p}_{24}^2)$$

$$\bar{p}_{24}^3 = \bar{p}_{24}^2 + \bar{v}_{24}^3$$

Dengan substitusi nilai:

$$\bar{v}_{24}^3 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(0.2 - 0.2) + 2.0 \cdot 0.9(0.2 - 0.2)$$

$$\bar{p}_{24}^3 = 0.2 + \bar{v}_{24}^3$$

$$\bar{v}_{24}^3 = 0$$

$$\bar{p}_{24}^3 = 0.2 + 0 = 0.2$$

Maka, posisi *centroid* baru untuk *cluster* 1 adalah

$$(5.48, 3.06, 1.38, 0.2)$$

Cluster 2:

Posisi *centroid* saat ini: (4.72, 2.97, 1.37, 0.2)

$$\bar{p}_{best}: (4.8, 3.2, 1.5, 0.2)$$

$$\bar{g}_{best}: (4.7, 3.1, 1.4, 0.2) \text{ diambil dari } \bar{p}_1$$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 2, i = 2, l = 1$, maka:

$$\bar{v}_{21}^3 = \omega \bar{v}_{21}^2 + a_1 r_1 (\bar{p}_{best_{21}} - \bar{p}_{21}^2) + a_2 r_2 (\bar{g}_{best_{21}} - \bar{p}_{21}^2)$$

$$\bar{p}_{21}^3 = \bar{p}_{21}^2 + \bar{v}_{21}^3$$

Dengan substitusi nilai:

$$\bar{v}_{21}^3 = 0.5 \cdot (0.2) + 2.0 \cdot 0.7(4.8 - 4.72) + 2.0 \cdot 0.9(4.7 - 4.72)$$

$$\bar{p}_{21}^3 = 4.72 + \bar{v}_{21}^3$$

$$\bar{v}_{21}^3 = 0.18$$

$$\bar{p}_{21}^3 = 4.72 + 0.18 = 4.90$$

Untuk $t = 2, i = 2, l = 2$, maka:

$$\bar{v}_{22}^3 = \omega \bar{v}_{22}^2 + a_1 r_1 (\bar{p}_{best_{22}} - \bar{p}_{22}^2) + a_2 r_2 (\bar{g}_{best_{22}} - \bar{p}_{22}^2)$$

$$\bar{p}_{22}^3 = \bar{p}_{22}^2 + \bar{v}_{22}^3$$

Dengan substitusi nilai:

$$\bar{v}_{22}^3 = 0.5 \cdot (-0.1) + 2.0 \cdot 0.7(3.2 - 2.97) + 2.0 \cdot 0.9(3.1 - 2.97)$$

$$\bar{p}_{22}^3 = 2.97 + \bar{v}_{22}^3$$

$$\bar{v}_{22}^3 = 0.51$$

$$\bar{p}_{22}^3 = 2.97 + 0.51 = 3.48$$

Untuk $t = 2, i = 2, l = 3$, maka:

$$\bar{v}_{23}^3 = \omega \bar{v}_{23}^2 + a_1 r_1 (\bar{p}_{best_{23}} - \bar{p}_{23}^2) + a_2 r_2 (\bar{g}_{best_{23}} - \bar{p}_{23}^2)$$

$$\bar{p}_{23}^3 = \bar{p}_{23}^2 + \bar{v}_{23}^3$$

Dengan substitusi nilai:

$$\bar{v}_{23}^3 = 0.5 \cdot (0.1) + 2.0 \cdot 0.7(1.5 - 1.37) + 2.0 \cdot 0.9(1.4 - 1.37)$$

$$\bar{p}_{23}^3 = 1.37 + \bar{v}_{23}^3$$

$$\bar{v}_{23}^3 = 0.29$$

$$\bar{p}_{23}^3 = 1.37 + 0.29 = 1.66$$

Untuk $t = 2, i = 2, l = 4$, maka:

$$\bar{v}_{24}^3 = \omega \bar{v}_{24}^2 + a_1 r_1 (\bar{p}_{best_{24}} - \bar{p}_{24}^2) + a_2 r_2 (\bar{g}_{best_{24}} - \bar{p}_{24}^2)$$

$$\bar{p}_{24}^3 = \bar{p}_{24}^2 + \bar{v}_{24}^3$$

Dengan substitusi nilai:

$$\bar{v}_{24}^3 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(0.2 - 0.2) + 2.0 \cdot 0.9(0.2 - 0.2)$$

$$\bar{p}_{24}^3 = 0.2 + \bar{v}_{24}^3$$

$$\bar{v}_{24}^3 = 0$$

$$\bar{p}_{24}^3 = 0.2 + 0 = 0.2$$

Maka, posisi *centroid* baru untuk *cluster* 2 adalah

$$(4.90, 3.48, 1.66, 0.2)$$

Cluster 3:

Posisi *centroid* saat ini: (4.98, 3.73, 1.23, 0.2)

$$\bar{p}_{best}: (4.9, 3.5, 1.3, 0.2)$$

$\bar{g}_{best}$: (5.0, 3.6, 1.4, 0.2) diambil dari $\bar{p}_1$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 2, i = 2, l = 1$, maka:

$$\bar{v}_{21}^3 = \omega \bar{v}_{21}^2 + a_1 r_1 (\bar{p}_{best_{21}} - \bar{p}_{21}^2) + a_2 r_2 (\bar{g}_{best_{21}} - \bar{p}_{21}^2)$$

$$\bar{p}_{21}^3 = \bar{p}_{21}^2 + \bar{v}_{21}^3$$

Dengan substitusi nilai:

$$\bar{v}_{21}^3 = 0.5 \cdot (-0.2) + 2.0 \cdot 0.7(4.9 - 4.98) + 2.0 \cdot 0.9(5.0 - 4.98)$$

$$\bar{p}_{21}^3 = 4.98 + \bar{v}_{21}^3$$

$$\bar{v}_{21}^3 = -0.18$$

$$\bar{p}_{21}^3 = 4.98 + (-0.18) = 4.80$$

Untuk $t = 2, i = 2, l = 2$, maka:

$$\bar{v}_{22}^3 = \omega \bar{v}_{22}^2 + a_1 r_1 (\bar{p}_{best_{22}} - \bar{p}_{22}^2) + a_2 r_2 (\bar{g}_{best_{22}} - \bar{p}_{22}^2)$$

$$\bar{p}_{22}^3 = \bar{p}_{22}^2 + \bar{v}_{22}^3$$

Dengan substitusi nilai:

$$\bar{v}_{22}^3 = 0.5 \cdot (0.1) + 2.0 \cdot 0.7(3.5 - 3.73) + 2.0 \cdot 0.9(3.6 - 3.73)$$

$$\bar{p}_{22}^3 = 3.73 + \bar{v}_{22}^3$$

$$\bar{v}_{22}^3 = -0.51$$

$$\bar{p}_{22}^3 = 3.73 + (-0.51) = 3.22$$

Untuk $t = 2, i = 2, l = 3$, maka:

$$\bar{v}_{23}^3 = \omega \bar{v}_{23}^2 + a_1 r_1 (\bar{p}_{best_{23}} - \bar{p}_{23}^2) + a_2 r_2 (\bar{g}_{best_{23}} - \bar{p}_{23}^2)$$

$$\bar{p}_{23}^3 = \bar{p}_{23}^2 + \bar{v}_{23}^3$$

Dengan substitusi nilai:

$$\bar{v}_{23}^3 = 0.5 \cdot (-0.5) + 2.0 \cdot 0.7(1.3 - 1.23) + 2.0 \cdot 0.9(1.4 - 1.23)$$

$$\bar{p}_{23}^3 = 1.23 + \bar{v}_{23}^3$$

$$\bar{v}_{23}^3 = 0.15$$

$$\bar{p}_{23}^3 = 1.23 + 0.15 = 1.38$$

Untuk $t = 2, i = 2, l = 4$, maka:

$$\bar{v}_{24}^3 = \omega \bar{v}_{24}^2 + a_1 r_1 (\bar{p}_{best_{24}} - \bar{p}_{24}^2) + a_2 r_2 (\bar{g}_{best_{24}} - \bar{p}_{24}^2)$$

$$\bar{p}_{24}^3 = \bar{p}_{24}^2 + \bar{v}_{24}^3$$

Dengan substitusi nilai:

$$\bar{v}_{24}^3 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(0.2 - 0.2) + 2.0 \cdot 0.9(0.2 - 0.2)$$

$$\bar{p}_{24}^3 = 0.2 + \bar{v}_{24}^3$$

$$\bar{v}_{24}^3 = 0$$

$$\bar{p}_{24}^3 = 0.2 + 0 = 0.2$$

Maka, posisi *centroid* baru untuk *cluster* 3 adalah

$$(4.80, 3.22, 1.38, 0.2)$$

Proses *update* untuk partikel 3 ($\bar{p}_3$) adalah sebagai berikut:

Cluster 1:

Posisi *centroid* saat ini: (4.97, 3.68, 1.23, 0.2)

$\bar{p}_{best}$: (5.2, 3.4, 1.3, 0.2)

$\bar{g}_{best}$: (5.1, 3.5, 1.4, 0.2) diambil dari $\bar{p}_1$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 2, i = 3, l = 1$, maka:

$$\bar{v}_{31}^3 = \omega \bar{v}_{31}^2 + a_1 r_1 (\bar{p}_{best_{31}} - \bar{p}_{31}^2) + a_2 r_2 (\bar{g}_{best_{31}} - \bar{p}_{31}^2)$$

$$\bar{p}_{31}^3 = \bar{p}_{31}^2 + \bar{v}_{31}^3$$

Dengan substitusi nilai:

$$\bar{v}_{31}^3 = 0.5 \cdot (-0.1) + 2.0 \cdot 0.7(5.2 - 4.97) + 2.0 \cdot 0.9(5.1 - 4.97)$$

$$\bar{p}_{31}^3 = 4.97 + \bar{v}_{31}^3$$

$$\bar{v}_{31}^3 = 0.51$$

$$\bar{p}_{31}^3 = 4.97 + 0.51 = 5.48$$

Untuk $t = 2, i = 3, l = 2$, maka:

$$\bar{v}_{32}^3 = \omega \bar{v}_{32}^2 + a_1 r_1 (\bar{p}_{best_{32}} - \bar{p}_{32}^2) + a_2 r_2 (\bar{g}_{best_{32}} - \bar{p}_{32}^2)$$

$$\bar{p}_{32}^3 = \bar{p}_{32}^2 + \bar{v}_{32}^3$$

Dengan substitusi nilai:

$$\bar{v}_{32}^3 = 0.5 \cdot (0.2) + 2.0 \cdot 0.7(3.4 - 3.68) + 2.0 \cdot 0.9(3.5 - 3.68)$$

$$\bar{p}_{32}^3 = 3.68 + \bar{v}_{32}^3$$

$$\bar{v}_{32}^3 = -0.62$$

$$\bar{p}_{32}^3 = 3.68 + (-0.62) = 3.06$$

Untuk $t = 2, i = 3, l = 3$, maka:

$$\bar{v}_{33}^3 = \omega \bar{v}_{33}^2 + a_1 r_1 (\bar{p}_{best_{33}} - \bar{p}_{33}^2) + a_2 r_2 (\bar{g}_{best_{33}} - \bar{p}_{33}^2)$$

$$\bar{p}_{33}^3 = \bar{p}_{33}^2 + \bar{v}_{33}^3$$

Dengan substitusi nilai:

$$\bar{v}_{33}^3 = 0.5 \cdot (-0.5) + 2.0 \cdot 0.7(1.3 - 1.23) + 2.0 \cdot 0.9(1.4 - 1.23)$$

$$\bar{p}_{33}^3 = 1.23 + \bar{v}_{33}^3$$

$$\bar{v}_{33}^3 = 0.15$$

$$\bar{p}_{33}^3 = 1.23 + 0.15 = 1.38$$

Untuk $t = 2, i = 3, l = 4$, maka:

$$\bar{v}_{34}^3 = \omega \bar{v}_{34}^2 + a_1 r_1 (\bar{p}_{best_{34}} - \bar{p}_{34}^2) + a_2 r_2 (\bar{g}_{best_{34}} - \bar{p}_{34}^2)$$

$$\bar{p}_{34}^3 = \bar{p}_{34}^2 + \bar{v}_{34}^2$$

Dengan substitusi nilai:

$$\bar{v}_{34}^3 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(0.2 - 0.2) + 2.0 \cdot 0.9(0.2 - 0.2)$$

$$\bar{p}_{34}^3 = 0.2 + \bar{v}_{34}^3$$

$$\bar{v}_{34}^3 = 0$$

$$\bar{p}_{34}^3 = 0.2 + 0 = 0.2$$

Maka, posisi *centroid* baru untuk *cluster* 1 adalah

$$(5.48, 3.06, 1.38, 0.2)$$

Cluster 2:

Posisi *centroid* saat ini: (4.88, 3.13, 1.53, 0.2)

$$\bar{p}_{best}: (4.6, 3.0, 1.3, 0.2)$$

$$\bar{g}_{best}: (4.7, 3.1, 1.4, 0.2) \text{ diambil dari } \bar{p}_1$$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 2, i = 3, l = 1$, maka:

$$\bar{v}_{31}^3 = \omega \bar{v}_{31}^2 + a_1 r_1 (\bar{p}_{best_{31}} - \bar{p}_{31}^2) + a_2 r_2 (\bar{g}_{best_{31}} - \bar{p}_{31}^2)$$

$$\bar{p}_{31}^3 = \bar{p}_{31}^2 + \bar{v}_{31}^3$$

Dengan substitusi nilai:

$$\bar{v}_{31}^3 = 0.5 \cdot (0.2) + 2.0 \cdot 0.7(4.6 - 4.88) + 2.0 \cdot 0.9(4.7 - 4.88)$$

$$\bar{p}_{31}^3 = 4.88 + \bar{v}_{31}^3$$

$$\bar{v}_{31}^3 = -0.62$$

$$\bar{p}_{31}^3 = 4.88 + (-0.62) = 4.26$$

Untuk $t = 2, i = 3, l = 2$, maka:

$$\bar{v}_{32}^3 = \omega \bar{v}_{32}^2 + a_1 r_1 (\bar{p}_{best_{32}} - \bar{p}_{32}^2) + a_2 r_2 (\bar{g}_{best_{32}} - \bar{p}_{32}^2)$$

$$\bar{p}_{32}^3 = \bar{p}_{32}^2 + \bar{v}_{32}^3$$

Dengan substitusi nilai:

$$\bar{v}_{32}^3 = 0.5 \cdot (-0.1) + 2.0 \cdot 0.7(3.0 - 3.13) + 2.0 \cdot 0.9(3.1 - 3.13)$$

$$\bar{p}_{32}^3 = 3.13 + \bar{v}_{32}^3$$

$$\bar{v}_{32}^3 = -0.29$$

$$\bar{p}_{32}^3 = 3.13 + (-0.29) = 2.84$$

Untuk $t = 2, i = 3, l = 3$, maka:

$$\bar{v}_{33}^3 = \omega \bar{v}_{33}^2 + a_1 r_1 (\bar{p}_{best_{33}} - \bar{p}_{33}^2) + a_2 r_2 (\bar{g}_{best_{33}} - \bar{p}_{33}^2)$$

$$\bar{p}_{33}^3 = \bar{p}_{33}^2 + \bar{v}_{33}^3$$

Dengan substitusi nilai:

$$\bar{v}_{33}^3 = 0.5 \cdot (0.1) + 2.0 \cdot 0.7(1.3 - 1.53) + 2.0 \cdot 0.9(1.4 - 1.53)$$

$$\bar{p}_{33}^3 = 1.53 + \bar{v}_{33}^3$$

$$\bar{v}_{33}^3 = -0.51$$

$$\bar{p}_{33}^3 = 1.53 + (-0.51) = 1.02$$

Untuk $t = 2, i = 3, l = 4$, maka:

$$\bar{v}_{34}^3 = \omega \bar{v}_{34}^2 + a_1 r_1 (\bar{p}_{best_{34}} - \bar{p}_{34}^2) + a_2 r_2 (\bar{g}_{best_{34}} - \bar{p}_{34}^2)$$

$$\bar{p}_{34}^3 = \bar{p}_{34}^2 + \bar{v}_{34}^3$$

Dengan substitusi nilai:

$$\bar{v}_{34}^3 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(0.2 - 0.2) + 2.0 \cdot 0.9(0.2 - 0.2)$$

$$\bar{p}_{34}^3 = 0.2 + \bar{v}_{34}^3$$

$$\bar{v}_{34}^3 = 0$$

$$\bar{p}_{34}^3 = 0.2 + 0 = 0.2$$

Maka, posisi *centroid* baru untuk *cluster* 2 adalah

$$(4.26, 2.84, 1.02, 0.2)$$

Cluster 3:

Posisi *centroid* saat ini: = (4.82, 3.81, 1.07, 0.2)

$$\bar{p}_{best} = (5.1, 3.4, 1.5, 0.2)$$

$$\bar{g}_{best} = (5.0, 3.6, 1.4, 0.2) \text{ diambil dari } \bar{p}_1$$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 2, i = 3, l = 1$, maka:

$$\bar{v}_{31}^3 = \omega \bar{v}_{31}^2 + a_1 r_1 (\bar{p}_{best_{31}} - \bar{p}_{31}^2) + a_2 r_2 (\bar{g}_{best_{31}} - \bar{p}_{31}^2)$$

$$\bar{p}_{31}^3 = \bar{p}_{31}^2 + \bar{v}_{31}^3$$

Dengan substitusi nilai:

$$\bar{v}_{31}^3 = 0.5 \cdot (-0.2) + 2.0 \cdot 0.7(5.1 - 4.82) + 2.0 \cdot 0.9(5.0 - 4.82)$$

$$\bar{p}_{31}^3 = 4.82 + \bar{v}_{31}^3$$

$$\bar{v}_{31}^3 = 0.62$$

$$\bar{p}_{31}^3 = 4.82 + 0.62 = 5.44$$

Untuk $t = 2, i = 3, l = 2$, maka:

$$\bar{v}_{32}^3 = \omega \bar{v}_{32}^2 + a_1 r_1 (\bar{p}_{best_{32}} - \bar{p}_{32}^2) + a_2 r_2 (\bar{g}_{best_{32}} - \bar{p}_{32}^2)$$

$$\bar{p}_{32}^3 = \bar{p}_{32}^2 + \bar{v}_{32}^3$$

Dengan substitusi nilai:

$$\bar{v}_{32}^3 = 0.5 \cdot (0.1) + 2.0 \cdot 0.7(3.4 - 3.81) + 2.0 \cdot 0.9(3.6 - 3.81)$$

$$\bar{p}_{32}^3 = 3.81 + \bar{v}_{32}^3$$

$$\bar{v}_{32}^3 = -0.90$$

$$\bar{p}_{32}^3 = 3.81 + (-0.90) = 2.91$$

Untuk $t = 2, i = 3, l = 3$, maka:

$$\bar{v}_{33}^3 = \omega \bar{v}_{33}^2 + a_1 r_1 (\bar{p}_{best_{33}} - \bar{p}_{33}^2) + a_2 r_2 (\bar{g}_{best_{33}} - \bar{p}_{33}^2)$$

$$\bar{p}_{33}^3 = \bar{p}_{33}^2 + \bar{v}_{33}^3$$

Dengan substitusi nilai:

$$\bar{v}_{33}^3 = 0.5 \cdot (-0.5) + 2.0 \cdot 0.7(1.5 - 1.07) + 2.0 \cdot 0.9(1.4 - 1.07)$$

$$\bar{p}_{33}^3 = 1.07 + \bar{v}_{33}^3$$

$$\bar{v}_{33}^3 = 0.95$$

$$\bar{p}_{33}^3 = 1.07 + 0.95 = 2.02$$

Untuk $t = 2, i = 3, l = 4$, maka:

$$\bar{v}_{34}^3 = \omega \bar{v}_{34}^2 + a_1 r_1 (\bar{p}_{best_{34}} - \bar{p}_{34}^2) + a_2 r_2 (\bar{g}_{best_{34}} - \bar{p}_{34}^2)$$

$$\bar{p}_{34}^3 = \bar{p}_{34}^2 + \bar{v}_{34}^3$$

Dengan substitusi nilai:

$$\bar{v}_{34}^3 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(0.2 - 0.2) + 2.0 \cdot 0.9(0.2 - 0.2)$$

$$\bar{p}_{34}^3 = 0.2 + \bar{v}_{34}^3$$

$$\bar{v}_{34}^3 = 0$$

$$\bar{p}_{34}^3 = 0.2 + 0 = 0.2$$

Maka, posisi *centroid* baru untuk *cluster* 3 adalah

$$(5.44, 2.91, 2.02, 0.2)$$

Berdasarkan perhitungan manual PSO di atas hanya dilakukan 2 iterasi saja.

Sedangkan implementasi menggunakan pemrograman dilakukan hingga 30 iterasi dan menghasilkan *centroid* terbaik (*Final $\bar{g}_{best}$*) sebagai berikut:

$$\bar{\mu}_1 = (5.2, 3.4, 1.3, 0.2)$$

$$\bar{\mu}_2 = (4.8, 3.2, 1.5, 0.2)$$

$$\bar{\mu}_3 = (4.9, 3.5, 1.3, 0.2)$$

dengan nilai nilai $\bar{g}_{best}$ pada iterasi ke-30 sebesar 0.16.

4.3.2 Dataset Iris (Menggunakan Seluruh Data)

PSO diterapkan untuk mengoptimalkan pemilihan *centroid* dalam algoritma *K-Means Clustering* pada seluruh dataset Iris. Proses optimasi dilakukan hingga 100 iterasi, dengan tujuan meningkatkan kualitas pengelompokan dibandingkan dengan *K-Means* standar. Pada setiap iterasi, PSO memperbarui posisi dan kecepatan partikel berdasarkan nilai terbaik individu ($\bar{p}_{best}$) dan nilai terbaik global ($\bar{g}_{best}$). Setelah 100 iterasi, algoritma mencapai *Final* $\bar{g}_{best}$ (*centroid* terbaik) sebagai berikut:

$$\bar{\mu}_1 = (5.01, 3.44, 1.45, 0.26)$$

$$\bar{\mu}_2 = (7.53, 3.82, 6.48, 1.44)$$

$$\bar{\mu}_3 = (6.17, 2.95, 4.65, 1.66)$$

dengan nilai nilai $\bar{g}_{best}$ pada iterasi ke-100 sebesar 117.94

4.3.3 Dataset Wine (5 Data Pertama)

1. Inisialisasi Partikel PSO

a. Inisialisasi Posisi Partikel

$$\bar{p}_i = (p_{i1}, p_{i2}, p_{i3}, \dots, p_{il})$$

Tabel 4.16 Inisialisasi Posisi Partikel pada PSO

$\bar{p}_i$	$\bar{\mu}_1$	$\bar{\mu}_2$	$\bar{\mu}_3$
$\bar{p}_1$	(13.93, 1.81, 2.36, 14.53)	(13.16, 2.36, 2.67, 18.6)	(13.24, 2.59, 2.87, 21)
$\bar{p}_2$	(14.37, 1.95, 2.5, 16.8)	(13.2, 1.78, 2.14, 11.2)	(14.23, 1.71, 2.43, 15.6)
$\bar{p}_3$	(14.2, 1.7, 2.4, 15.5)	(13.1, 2.3, 2.6, 18.5)	(13.3, 2.5, 2.8, 20.9)

Tabel 4.16 menunjukkan hasil inisialisasi posisi awal partikel pada algoritma PSO. Setiap partikel terdiri dari tiga *centroid* dengan nilai fitur masing-masing.

b. Inisialisasi Kecepatan Partikel

$$\bar{v}_i = (v_{i1}, v_{i2}, \dots, v_{il})$$

Pada tahap awal, kecepatan partikel diinisialisasi secara acak untuk setiap dimensi. Nilai kecepatan awal untuk masing-masing partikel sebagai berikut:

- $\bar{v}_1 = (-0.1, 0.2, -0.5, 0.0)$
- $\bar{v}_2 = (0.2, -0.1, 0.1, 0.0)$
- $\bar{v}_3 = (-0.2, 0.1, -0.5, 0.0)$

c. Parameter PSO

Parameter pada PSO juga ditentukan dengan nilai sebagai berikut:

- $\omega = 0.5$
- $a_1 = 2.0$
- $a_2 = 2.0$
- $r_1 = 0.7$
- $r_2 = 0.9$

2. Menghitung SSE (Fungsi Objektif) Menggunakan Persamaan (3.4)

a. Hitung SSE untuk Partikel 1

Centroid awal (partikel 1) diambil dari tabel 4.16.

$$\bar{\mu}_1 = (13.93, 1.81, 2.36, 14.53)$$

$$\bar{\mu}_2 = (13.16, 2.36, 2.67, 18.6)$$

$$\bar{\mu}_3 = (13.24, 2.59, 2.87, 21)$$

Misalkan menggunakan 5 data pertama dari dataset Wine sesuai pada data tabel 4.2, kemudian menghitung jarak *Euclidean* dimulai antara data pertama ($\bar{x}_1$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_1, \bar{\mu}_1) = \|\bar{x}_1 - \bar{\mu}_1\| = 1.11$$

$$d(\bar{x}_1, \bar{\mu}_2) = \|\bar{x}_1 - \bar{\mu}_2\| = 3.26$$

$$d(\bar{x}_1, \bar{\mu}_3) = \|\bar{x}_1 - \bar{\mu}_3\| = 5.57$$

Kemudian dilanjutkan dengan perhitungan jarak *Euclidean* antara data kedua ($\bar{x}_2$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_2, \bar{\mu}_1) = \|\bar{x}_2 - \bar{\mu}_1\| = 3.41$$

$$d(\bar{x}_2, \bar{\mu}_2) = \|\bar{x}_2 - \bar{\mu}_2\| = 7.43$$

$$d(\bar{x}_2, \bar{\mu}_3) = \|\bar{x}_2 - \bar{\mu}_3\| = 9.85$$

Kemudian dilanjutkan dengan perhitungan jarak *Euclidean* antara data ketiga ($\bar{x}_3$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_3, \bar{\mu}_1) = \|\bar{x}_3 - \bar{\mu}_1\| = 4.19$$

$$d(\bar{x}_3, \bar{\mu}_2) = \|\bar{x}_3 - \bar{\mu}_2\| = 0$$

$$d(\bar{x}_3, \bar{\mu}_3) = \|\bar{x}_3 - \bar{\mu}_3\| = 2.42$$

Kemudian dilanjutkan dengan perhitungan jarak *Euclidean* antara data keempat ($\bar{x}_4$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_4, \bar{\mu}_1) = \|\bar{x}_4 - \bar{\mu}_1\| = 2.32$$

$$d(\bar{x}_4, \bar{\mu}_2) = \|\bar{x}_4 - \bar{\mu}_2\| = 2.21$$

$$d(\bar{x}_4, \bar{\mu}_3) = \|\bar{x}_4 - \bar{\mu}_3\| = 4.41$$

Terakhir dilanjutkan dengan perhitungan jarak *Euclidean* antara data kelima ($\bar{x}_5$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_5, \bar{\mu}_1) = \|\bar{x}_5 - \bar{\mu}_1\| = 6.51$$

$$d(\bar{x}_5, \bar{\mu}_2) = \|\bar{x}_5 - \bar{\mu}_2\| = 2.41$$

$$d(\bar{x}_5, \bar{\mu}_3) = \|\bar{x}_5 - \bar{\mu}_3\| = 0$$

Setelah dihitung jarak *Euclidean*, kemudian dikelompokkan ke dalam cluster berdasarkan jarak terkecil.

Tabel 4.17 Pengelompokan *Cluster* Partikel 1 pada Iterasi 1

$\bar{x}_i$	$d(\bar{\mu}_1)$	$d(\bar{\mu}_2)$	$d(\bar{\mu}_3)$	<i>Cluster</i>
$\bar{x}_1$	1.11	3.26	5.57	1
$\bar{x}_2$	3.41	7.43	9.85	1
$\bar{x}_3$	4.19	0	2.42	2
$\bar{x}_4$	2.32	2.21	4.41	2
$\bar{x}_5$	6.51	2.41	0	3

Setelah semua data masuk ke *cluster* masing-masing, kemudian menghitung SSE dengan persamaan (3.4).

- Untuk *cluster* 1 yang terdiri dari data $C_1 = \{\bar{x}_1, \bar{x}_2\}$,

$\bar{x}_1 = (14.23, 1.71, 2.43, 15.6)$ dengan *centroid* $\bar{\mu}_1 = (13.93, 1.81, 2.36, 14.53)$, diperoleh:

$$SE_1: (14.23 - 13.93)^2 + (1.71 - 1.81)^2 + (2.43 - 2.36)^2 + (15.6 - 14.53)^2 = 1.249$$

$\bar{x}_2 = (13.2, 1.78, 2.14, 11.2)$ dengan *centroid* $\bar{\mu}_1 = (13.93, 1.81, 2.36, 14.53)$, diperoleh:

$$SE_2: (13.2 - 13.93)^2 + (1.78 - 1.81)^2 + (2.14 - 2.36)^2 + (11.2 - 14.53)^2 = 11.671$$

- Untuk *cluster* 2 yang terdiri dari data $C_2 = \{\bar{x}_3, \bar{x}_4\}$,

$\bar{x}_3 = (13.16, 2.36, 2.67, 18.6)$ dengan *centroid* $\bar{\mu}_2 = (13.16, 2.36, 2.67, 18.6)$, diperoleh:

$$SE_3: (13.16 - 13.16)^2 + (2.36 - 2.36)^2 + (2.67 - 2.67)^2 + (18.6 - 18.6)^2 = 0$$

$\bar{x}_4 = (14.37, 1.95, 2.5, 16.8)$ dengan *centroid* $\bar{\mu}_2 = (13.16, 2.36, 2.67, 18.6)$, diperoleh:

$$SE_4: (14.37 - 13.16)^2 + (1.95 - 2.36)^2 + (2.5 - 2.67)^2 + (16.8 - 18.6)^2 = 4.901$$

- Untuk *cluster* 3 yang terdiri dari data $C_3 = \{\bar{x}_5\}$,

$\bar{x}_5 = (13.24, 2.59, 2.87, 21)$ dengan *centroid* $\bar{\mu}_3 = (13.24, 2.59, 2.87, 21)$, diperoleh:

$$SE_5: (13.24 - 13.24)^2 + (2.59 - 2.59)^2 + (2.87 - 2.87)^2 + (21 - 21)^2 = 0$$

Dengan demikian,

$$\begin{aligned} total\ SSE &= SE_1 + SE_2 + SE_3 + SE_4 + SE_5 \\ &= 1.249 + 11.671 + 0 + 4.901 + 0 = 17.821 \end{aligned}$$

b. Hitung SSE untuk Partikel 2

Centroid awal (partikel 2) diambil dari tabel 4.16.

$$\bar{\mu}_1 = (14.37, 1.95, 2.5, 16.8)$$

$$\bar{\mu}_2 = (13.2, 1.78, 2.14, 11.2)$$

$$\bar{\mu}_3 = (14.23, 1.71, 2.43, 15.6)$$

Misalkan menggunakan 5 data pertama dari dataset Iris sesuai pada data tabel 4.2, kemudian menghitung jarak *Euclidean* antara data pertama ($\bar{x}_1$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_1, \bar{\mu}_1) = \|\bar{x}_1 - \bar{\mu}_1\| = 1.23$$

$$d(\bar{x}_1, \bar{\mu}_2) = \|\bar{x}_1 - \bar{\mu}_2\| = 4.52$$

$$d(\bar{x}_1, \bar{\mu}_3) = \|\bar{x}_1 - \bar{\mu}_3\| = 0$$

Kemudian dilanjut perhitungan jarak *Euclidean* dimulai antara data kedua ($\bar{x}_2$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_2, \bar{\mu}_1) = \|\bar{x}_2 - \bar{\mu}_1\| = 5.73$$

$$d(\bar{x}_2, \bar{\mu}_2) = \|\bar{x}_2 - \bar{\mu}_2\| = 0$$

$$d(\bar{x}_2, \bar{\mu}_3) = \|\bar{x}_2 - \bar{\mu}_3\| = 4.52$$

Kemudian dilanjut perhitungan jarak *Euclidean* dimulai antara data ketiga ($\bar{x}_3$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_3, \bar{\mu}_1) = \|\bar{x}_3 - \bar{\mu}_1\| = 2.21$$

$$d(\bar{x}_3, \bar{\mu}_2) = \|\bar{x}_3 - \bar{\mu}_2\| = 7.43$$

$$d(\bar{x}_3, \bar{\mu}_3) = \|\bar{x}_3 - \bar{\mu}_3\| = 3.25$$

Kemudian dilanjut perhitungan jarak *Euclidean* dimulai antara data keempat ($\bar{x}_4$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_4, \bar{\mu}_1) = \|\bar{x}_4 - \bar{\mu}_1\| = 0$$

$$d(\bar{x}_4, \bar{\mu}_2) = \|\bar{x}_4 - \bar{\mu}_2\| = 5.73$$

$$d(\bar{x}_4, \bar{\mu}_3) = \|\bar{x}_4 - \bar{\mu}_3\| = 1.23$$

Terakhir dilanjut dengan perhitungan jarak *Euclidean* antara data kelima ($\bar{x}_5$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_5, \bar{\mu}_1) = \|\bar{x}_5 - \bar{\mu}_1\| = 4.40$$

$$d(\bar{x}_5, \bar{\mu}_2) = \|\bar{x}_5 - \bar{\mu}_2\| = 9.85$$

$$d(\bar{x}_5, \bar{\mu}_3) = \|\bar{x}_5 - \bar{\mu}_3\| = 5.57$$

Setelah dihitung jarak *Euclidean*, kemudian dikelompokkan ke dalam cluster berdasarkan jarak terkecil.

Tabel 4.18 Pengelompokan *Cluster* Partikel 2 pada Iterasi 1

$\bar{x}_i$	$d(\bar{\mu}_1)$	$d(\bar{\mu}_2)$	$d(\bar{\mu}_3)$	Cluster
$\bar{x}_1$	1.23	4.52	0	3
$\bar{x}_2$	5.73	0	4.52	2
$\bar{x}_3$	2.21	7.43	3.25	1
$\bar{x}_4$	0	5.73	1.23	1
$\bar{x}_5$	4.40	9.85	5.57	1

Setelah semua data masuk ke *cluster* masing-masing, kemudian menghitung SSE dengan persamaan (3.4).

- Untuk *cluster* 1 yang terdiri dari data $C_1 = \{\bar{x}_3, \bar{x}_4, \bar{x}_5\}$,

$\bar{x}_3 = (13.16, 2.36, 2.67, 18.6)$ dengan *centroid* $\bar{\mu}_1 = (14.37, 1.95, 2.5, 16.8)$, diperoleh:

$$SE_3: (13.16 - 14.37)^2 + (2.36 - 1.95)^2 + (2.67 - 2.5)^2 + (18.6 - 16.8)^2 = 4.901$$

$\bar{x}_4 = (14.37, 1.95, 2.5, 16.8)$ dengan *centroid* $\bar{\mu}_1 = (14.37, 1.95, 2.5, 16.8)$, diperoleh:

$$SE_4: (14.37 - 14.37)^2 + (1.95 - 1.95)^2 + (2.5 - 2.5)^2 + (16.8 - 16.8)^2 = 0$$

$\bar{x}_5 = (13.24, 2.59, 2.87, 21)$ dengan *centroid* $\bar{\mu}_1 = (14.37, 1.95, 2.5, 16.8)$, diperoleh:

$$SE_5: (13.24 - 14.37)^2 + (2.59 - 1.95)^2 + (2.87 - 2.5)^2 + (21 - 16.8)^2 = 19.463$$

- Untuk *cluster* 2 yang terdiri dari data $C_2 = \{\bar{x}_2\}$,

$\bar{x}_2 = (13.2, 1.78, 2.14, 11.2)$ dengan *centroid* $\bar{\mu}_2 = (13.2, 1.78, 2.14, 11.2)$, diperoleh:

$$SE_2: (13.2 - 13.2)^2 + (1.78 - 1.78)^2 + (2.14 - 2.14)^2 + (11.2 - 11.2)^2 = 0$$

- Untuk *cluster* 3 yang terdiri dari data $C_3 = \{\bar{x}_1\}$,

$\bar{x}_1 = (14.23, 1.71, 2.43, 15.6)$ dengan *centroid* $\bar{\mu}_3 = (14.23, 1.71, 2.43, 15.6)$, diperoleh:

$$SE_1: (14.23 - 14.23)^2 + (1.71 - 1.71)^2 + (2.43 - 2.43)^2 + (15.6 - 15.6)^2 = 0$$

Dengan demikian,

$$\begin{aligned} total\ SSE &= SE_1 + SE_2 + SE_3 + SE_4 + SE_5 \\ &= 0 + 0 + 4.901 + 0 + 19.463 = 24.364 \end{aligned}$$

c. Hitung SSE untuk Partikel 3

Centroid awal (partikel 3):

$$\bar{\mu}_1 = (14.2, 1.7, 2.4, 15.5)$$

$$\bar{\mu}_2 = (13.1, 2.3, 2.6, 18.5)$$

$$\bar{\mu}_3 = (13.3, 2.5, 2.8, 20.9)$$

Misalkan menggunakan 5 data pertama dari dataset Iris sesuai pada data tabel 4.2, kemudian menghitung jarak *Euclidean* antara data pertama ($\bar{x}_1$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_1, \bar{\mu}_1) = \|\bar{x}_1 - \bar{\mu}_1\| = 0.10$$

$$d(\bar{x}_1, \bar{\mu}_2) = \|\bar{x}_1 - \bar{\mu}_2\| = 3.17$$

$$d(\bar{x}_1, \bar{\mu}_3) = \|\bar{x}_1 - \bar{\mu}_3\| = 5.37$$

Kemudian dilanjut perhitungan jarak *Euclidean* dimulai antara data kedua ($\bar{x}_2$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_2, \bar{\mu}_1) = \|\bar{x}_2 - \bar{\mu}_1\| = 4.42$$

$$d(\bar{x}_2, \bar{\mu}_2) = \|\bar{x}_2 - \bar{\mu}_2\| = 7.33$$

$$d(\bar{x}_2, \bar{\mu}_3) = \|\bar{x}_2 - \bar{\mu}_3\| = 9.75$$

Kemudian dilanjut perhitungan jarak *Euclidean* dimulai antara data ketiga ($\bar{x}_3$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_3, \bar{\mu}_1) = \|\bar{x}_3 - \bar{\mu}_1\| = 3.34$$

$$d(\bar{x}_3, \bar{\mu}_2) = \|\bar{x}_3 - \bar{\mu}_2\| = 0.14$$

$$d(\bar{x}_3, \bar{\mu}_3) = \|\bar{x}_3 - \bar{\mu}_3\| = 2.31$$

Kemudian dilanjut perhitungan jarak *Euclidean* dimulai antara data keempat ($\bar{x}_4$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_4, \bar{\mu}_1) = \|\bar{x}_4 - \bar{\mu}_1\| = 1.33$$

$$d(\bar{x}_4, \bar{\mu}_2) = \|\bar{x}_4 - \bar{\mu}_2\| = 2.15$$

$$d(\bar{x}_4, \bar{\mu}_3) = \|\bar{x}_4 - \bar{\mu}_3\| = 4.28$$

Terakhir dilanjut dengan perhitungan jarak *Euclidean* antara data kelima ($\bar{x}_5$) dan setiap *centroid* $\bar{\mu}_1, \bar{\mu}_2$, serta $\bar{\mu}_3$.

$$d(\bar{x}_5, \bar{\mu}_1) = \|\bar{x}_5 - \bar{\mu}_1\| = 5.67$$

$$d(\bar{x}_5, \bar{\mu}_2) = \|\bar{x}_5 - \bar{\mu}_2\| = 2.53$$

$$d(\bar{x}_5, \bar{\mu}_3) = \|\bar{x}_5 - \bar{\mu}_3\| = 0.16$$

Setelah dihitung jarak *Euclidean*, kemudian dikelompokkan ke dalam cluster berdasarkan jarak terkecil.

Tabel 4.19 Pengelompokan *Cluster* Partikel 3 pada Iterasi 1

$\bar{x}_i$	$d(\bar{\mu}_1)$	$d(\bar{\mu}_2)$	$d(\bar{\mu}_3)$	<i>Cluster</i>
$\bar{x}_1$	0.10	3.17	5.37	1
$\bar{x}_2$	4.42	7.33	9.75	1
$\bar{x}_3$	3.34	0.14	2.31	2
$\bar{x}_4$	1.33	2.15	4.28	1
$\bar{x}_5$	5.67	2.53	0.16	3

Setelah semua data masuk ke *cluster* masing-masing, kemudian menghitung SSE dengan persamaan (3.4).

- Untuk *cluster* 1 yang terdiri dari data $C_1 = \{\bar{x}_1, \bar{x}_2, \bar{x}_4\}$,

$\bar{x}_1 = (14.23, 1.71, 2.43, 15.6)$ dengan *centroid* $\bar{\mu}_1 = (14.2, 1.7, 2.4, 15.5)$, diperoleh:

$$SE_1: (14.23 - 14.2)^2 + (1.71 - 1.7)^2 + (2.43 - 2.4)^2 + (15.6 - 15.5)^2 = 0.011$$

$\bar{x}_2 = (13.2, 1.78, 2.14, 11.2)$ dengan *centroid* $\bar{\mu}_1 = (14.2, 1.7, 2.4, 15.5)$, diperoleh:

$$SE_2: (13.2 - 14.2)^2 + (1.78 - 1.7)^2 + (2.14 - 2.4)^2 + (11.2 - 15.5)^2 = 15.564$$

$\bar{x}_4 = (14.37, 1.95, 2.5, 16.8)$ dengan *centroid* $\bar{\mu}_1 = (14.2, 1.7, 2.4, 15.5)$, diperoleh:

$$SE_4: (14.37 - 14.2)^2 + (1.95 - 1.7)^2 + (2.5 - 2.4)^2 + (16.8 - 15.5)^2 = 1.791$$

- Untuk *cluster* 2 yang terdiri dari data $C_2 = \{\bar{x}_2\}$,

$\bar{x}_2 = (13.2, 1.78, 2.14, 11.2)$ dengan *centroid* $\bar{\mu}_2 = (13.1, 2.3, 2.6, 18.5)$, diperoleh:

$$SE_3: (13.16 - 13.1)^2 + (2.36 - 2.3)^2 + (2.67 - 2.6)^2 + (18.6 - 18.5)^2 = 0.022$$

- Untuk *cluster* 3 yang terdiri dari data $C_3 = \{\bar{x}_5\}$,

$\bar{x}_5 = (13.24, 2.59, 2.87, 21)$ dengan *centroid* $\bar{\mu}_3 = (13.3, 2.5, 2.8, 20.9)$, diperoleh:

$$SE_5: (13.24 - 13.3)^2 + (2.59 - 2.5)^2 + (2.87 - 2.8)^2 + (21 - 20.9)^2 = 0.026$$

Dengan demikian,

$$\begin{aligned} total\ SSE &= SE_1 + SE_2 + SE_3 + SE_4 + SE_5 \\ &= 0.011 + 15.564 + 0.022 + 1.791 + 0.026 = 17.414 \end{aligned}$$

3. *Update* Kecepatan dan Posisi

Karena ini adalah iterasi pertama, maka:

- $\bar{p}_{best}$ untuk setiap partikel adalah posisi awal partikel tersebut.
- $\bar{g}_{best}$ adalah posisi dari partikel dengan nilai SSE terkecil, yaitu partikel 3 ($\bar{p}_3$)

Proses *update* partikel 1 ($\bar{p}_1$) adalah sebagai berikut:

Cluster 1:

Posisi *centroid* saat ini: (13.93, 1.81, 2.36, 14.53)

$\bar{p}_{best}$: (13.93, 1.81, 2.36, 14.53)

$\bar{g}_{best}$: (14.2, 1.7, 2.4, 15.5) diambil dari $\bar{p}_3$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6)

Untuk $t = 1, i = 1, l = 1$, maka:

$$\bar{v}_{11}^2 = 0.5 \cdot (-0.1) + 2.0 \cdot 0.7(13.93 - 13.93) + 2.0 \cdot 0.9(14.2 - 13.93)$$

$$\bar{p}_{11}^2 = 13.93 + \bar{v}_{11}^2$$

$$\bar{v}_{11}^2 = 0.43$$

$$\bar{p}_{11}^2 = 13.93 + 0.43 = 14.36$$

Untuk $t = 1, i = 1, l = 2$, maka:

$$\bar{v}_{12}^2 = 0.5 \cdot (0.2) + 2.0 \cdot 0.7(1.81 - 1.81) + 2.0 \cdot 0.9(1.7 - 1.81)$$

$$\bar{p}_{12}^2 = 1.81 + \bar{v}_{12}^2$$

$$\bar{v}_{12}^2 = -0.10$$

$$\bar{p}_{12}^2 = 1.81 + (-0.10) = 1.71$$

Untuk $t = 1, i = 1, l = 3$, maka:

$$\bar{v}_{13}^2 = 0.5 \cdot (-0.5) + 2.0 \cdot 0.7(2.36 - 2.36) + 2.0 \cdot 0.9(2.4 - 2.36)$$

$$\bar{p}_{13}^2 = 2.36 + \bar{v}_{13}^2$$

$$\bar{v}_{13}^2 = (-0.18)$$

$$\bar{p}_{13}^2 = 2.36 + (-0.18) = 2.18$$

Untuk $t = 1, i = 1, l = 4$, maka:

$$\bar{v}_{14}^2 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(14.53 - 14.53) + 2.0 \cdot 0.9(15.5 - 14.53)$$

$$\bar{p}_{14}^2 = 14.53 + \bar{v}_{14}^2$$

$$\bar{v}_{14}^2 = 1.75$$

$$\bar{p}_{14}^2 = 14.53 + 1.75 = 16.28$$

Maka, posisi *centroid* baru untuk *cluster* 1 adalah

$$(14.36, 1.71, 2.18, 16.28)$$

Cluster 2:

Posisi *centroid* saat ini: (13.16, 2.36, 2.67, 18.6)

$$\bar{p}_{best}: (13.16, 2.36, 2.67, 18.6)$$

$$\bar{g}_{best}: (13.1, 2.3, 2.6, 18.5) \text{ diambil dari } \bar{p}_3$$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 1, i = 1, l = 1$, maka:

$$\bar{v}_{11}^2 = 0.5 \cdot (0.2) + 2.0 \cdot 0.7(13.16 - 13.16) + 2.0 \cdot 0.9(13.1 - 13.16)$$

$$\bar{p}_{11}^2 = 13.16 + \bar{v}_{11}^2$$

$$\bar{v}_{11}^2 = -0.01$$

$$\bar{p}_{11}^2 = 13.16 + (-0.01) = 13.15$$

Untuk $t = 1, i = 1, l = 2$, maka:

$$\bar{v}_{12}^2 = 0.5 \cdot (-0.1) + 2.0 \cdot 0.7(2.36 - 2.36) + 2.0 \cdot 0.9(2.3 - 2.36)$$

$$\bar{p}_{12}^2 = 2.36 + \bar{v}_{12}^2$$

$$\bar{v}_{12}^2 = -0.16$$

$$\bar{p}_{12}^2 = 2.36 + (-0.16) = 2.20$$

Untuk $t = 1, i = 1, l = 3$, maka:

$$\bar{v}_{13}^2 = 0.5 \cdot (0.1) + 2.0 \cdot 0.7(2.67 - 2.67) + 2.0 \cdot 0.9(2.6 - 2.67)$$

$$\bar{p}_{13}^2 = 2.67 + \bar{v}_{13}^2$$

$$\bar{v}_{13}^2 = -0.08$$

$$\bar{p}_{13}^2 = 2.67 + (-0.08) = 2.59$$

Untuk $t = 1, i = 1, l = 4$, maka:

$$\bar{v}_{14}^2 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(18.6 - 18.6) + 2.0 \cdot 0.9(18.5 - 18.6)$$

$$\bar{p}_{14}^2 = 18.6 + \bar{v}_{14}^2$$

$$\bar{v}_{14}^2 = -0.18$$

$$\bar{p}_{14}^2 = 18.6 + (-0.18) = 18.42$$

Maka, posisi *centroid* baru untuk *cluster* 2 adalah

$$(13.15, 2.20, 2.59, 18.42)$$

Cluster 3:

Posisi *centroid* saat ini: (13.24, 2.59, 2.87, 21)

$$\bar{p}_{best}: (13.24, 2.59, 2.87, 21)$$

$$\bar{g}_{best}: (13.3, 2.5, 2.8, 20.9) \text{ dari } \bar{p}_3$$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 1, i = 1, l = 1$, maka:

$$\bar{v}_{11}^2 = 0.5 \cdot (-0.2) + 2.0 \cdot 0.7(13.24 - 13.24) + 2.0 \cdot 0.9(13.3 - 13.24)$$

$$\bar{p}_{11}^2 = 13.24 + \bar{v}_{11}^2$$

$$\bar{v}_{11}^2 = 0.01$$

$$\bar{p}_{11}^2 = 13.24 + 0.01 = 13.25$$

Untuk $t = 1, i = 1, l = 2$, maka:

$$\bar{v}_{12}^2 = 0.5 \cdot (0.1) + 2.0 \cdot 0.7(2.59 - 2.59) + 2.0 \cdot 0.9(2.5 - 2.59)$$

$$\bar{p}_{12}^2 = 2.59 + \bar{v}_{12}^2$$

$$\bar{v}_{12}^2 = -0.11$$

$$\bar{p}_{12}^2 = 2.59 + (-0.11) = 2.48$$

Untuk $t = 1, i = 1, l = 3$, maka:

$$\bar{v}_{13}^2 = 0.5 \cdot (-0.5) + 2.0 \cdot 0.7(2.87 - 2.87) + 2.0 \cdot 0.9(2.8 - 2.87)$$

$$\bar{p}_{13}^2 = 2.87 + \bar{v}_{13}^2$$

$$\bar{v}_{13}^2 = -0.38$$

$$\bar{p}_{13}^2 = 2.87 + (-0.38) = 2.49$$

Untuk $t = 1, i = 1, l = 4$, maka:

$$\bar{v}_{14}^2 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(21 - 21) + 2.0 \cdot 0.9(20.9 - 21)$$

$$\bar{p}_{14}^2 = 21 + \bar{v}_{14}^2$$

$$\bar{v}_{14}^2 = -0.18$$

$$\bar{p}_{14}^2 = 21 + (-0.18) = 20.82$$

Maka, posisi *centroid* baru untuk *cluster* 3 adalah

$$(13.25, 2.48, 2.49, 20.82)$$

Proses *update* untuk partikel 2 ($\bar{p}_2$) adalah sebagai berikut:

Cluster 1:

Posisi *centroid* saat ini: (14.37, 1.95, 2.5, 16.8)

$$\bar{p}_{best}: (14.37, 1.95, 2.5, 16.8)$$

$$\bar{g}_{best}: (14.2, 1.7, 2.4, 15.5) \text{ diambil dari } \bar{p}_3$$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 1, i = 2, l = 1$, maka:

$$\bar{v}_{21}^2 = 0.5 \cdot (-0.1) + 2.0 \cdot 0.7(14.37 - 14.37) + 2.0 \cdot 0.9(14.2 - 14.37)$$

$$\bar{p}_{21}^2 = 14.37 + \bar{v}_{21}^2$$

$$\bar{v}_{21}^2 = -0.36$$

$$\bar{p}_{21}^2 = 14.37 + (-0.36) = 14.01$$

Untuk $t = 1, i = 2, l = 2$, maka:

$$\bar{v}_{22}^2 = 0.5 \cdot (0.2) + 2.0 \cdot 0.7(1.95 - 1.95) + 2.0 \cdot 0.9(1.7 - 1.95)$$

$$\bar{p}_{22}^2 = 1.95 + \bar{v}_{22}^2$$

$$\bar{v}_{22}^2 = -0.35$$

$$\bar{p}_{22}^2 = 1.95 + (-0.35) = 1.60$$

Untuk $t = 1, i = 2, l = 3$, maka:

$$\bar{v}_{23}^2 = 0.5 \cdot (-0.5) + 2.0 \cdot 0.7(2.5 - 2.5) + 2.0 \cdot 0.9(2.4 - 2.5)$$

$$\bar{p}_{23}^2 = 2.5 + \bar{v}_{23}^2$$

$$\bar{v}_{23}^2 = -0.43$$

$$\bar{p}_{23}^2 = 2.5 + (-0.43) = 2.07$$

Untuk $t = 1, i = 2, l = 4$, maka:

$$\bar{v}_{24}^2 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(16.8 - 16.8) + 2.0 \cdot 0.9(15.5 - 16.8)$$

$$\bar{p}_{24}^2 = 16.8 + \bar{v}_{24}^2$$

$$\bar{v}_{24}^2 = -2.34$$

$$\bar{p}_{24}^2 = 16.8 + (-2.34) = 14.46$$

Maka, posisi *centroid* baru untuk *cluster* 1 adalah

$$(14.01, 1.60, 2.07, 14.46)$$

Cluster 2:

Posisi *centroid* saat ini: (13.2, 1.78, 2.14, 11.2)

$\bar{p}_{best}$: (13.2, 1.78, 2.14, 11.2)

$\bar{g}_{best}$: (13.1, 2.3, 2.6, 18.5) diambil dari $\bar{p}_3$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 1, i = 2, l = 1$, maka:

$$\bar{v}_{21}^2 = 0.5 \cdot (0.2) + 2.0 \cdot 0.7(13.2 - 13.2) + 2.0 \cdot 0.9(13.1 - 13.2)$$

$$\bar{p}_{21}^2 = 13.2 + \bar{v}_{21}^2$$

$$\bar{v}_{21}^2 = -0.08$$

$$\bar{p}_{21}^2 = 13.2 + (-0.08) = 13.12$$

Untuk $t = 1, i = 2, l = 2$, maka:

$$\bar{v}_{22}^2 = 0.5 \cdot (-0.1) + 2.0 \cdot 0.7(1.78 - 1.78) + 2.0 \cdot 0.9(2.3 - 1.78)$$

$$\bar{p}_{22}^2 = 1.78 + \bar{v}_{22}^2$$

$$\bar{v}_{22}^2 = 0.89$$

$$\bar{p}_{22}^2 = 1.78 + 0.89 = 2.67$$

Untuk $t = 1, i = 2, l = 3$, maka:

$$\bar{v}_{23}^2 = 0.5 \cdot (0.1) + 2.0 \cdot 0.7(2.14 - 2.14) + 2.0 \cdot 0.9(2.6 - 2.14)$$

$$\bar{p}_{23}^2 = 2.14 + \bar{v}_{23}^2$$

$$\bar{v}_{23}^2 = 0.88$$

$$\bar{p}_{23}^2 = 2.14 + 0.88 = 3.02$$

Untuk $t = 1, i = 2, l = 4$, maka:

$$\bar{v}_{24}^2 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(11.2 - 11.2) + 2.0 \cdot 0.9(18.5 - 11.2)$$

$$\bar{p}_{24}^2 = 11.2 + \bar{v}_{24}^2$$

$$\bar{v}_{24}^2 = 13.14$$

$$\bar{p}_{24}^2 = 11.2 + 13.14 = 24.34$$

Maka, posisi *centroid* baru untuk *cluster* 2 adalah

$$(13.12, 2.67, 3.02, 24.34)$$

Cluster 3:

Posisi *centroid* saat ini: (14.23, 1.71, 2.43, 15.6)

$$\bar{p}_{best}: (14.23, 1.71, 2.43, 15.6)$$

$$\bar{g}_{best}: (13.3, 2.5, 2.8, 20.9) \text{ diambil dari } \bar{p}_3$$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 1, i = 2, l = 1$, maka:

$$\bar{v}_{21}^2 = 0.5 \cdot (-0.2) + 2.0 \cdot 0.7(14.23 - 14.23) + 2.0 \cdot 0.9(13.3 - 14.23)$$

$$\bar{p}_{21}^2 = 14.23 + \bar{v}_{21}^2$$

$$\bar{v}_{21}^2 = -1.77$$

$$\bar{p}_{21}^2 = 14.23 + (-1.77) = 12.46$$

Untuk $t = 1, i = 2, l = 2$, maka:

$$\bar{v}_{22}^2 = 0.5 \cdot (0.1) + 2.0 \cdot 0.7(1.71 - 1.71) + 2.0 \cdot 0.9(2.5 - 1.71)$$

$$\bar{p}_{22}^2 = 1.71 + \bar{v}_{22}^2$$

$$\bar{v}_{22}^2 = 1.47$$

$$\bar{p}_{22}^2 = 1.71 + 1.47 = 3.18$$

Untuk $t = 1, i = 2, l = 3$, maka:

$$\bar{v}_{23}^2 = 0.5 \cdot (-0.5) + 2.0 \cdot 0.7(2.43 - 2.43) + 2.0 \cdot 0.9(2.8 - 2.43)$$

$$\bar{p}_{23}^2 = 2.43 + \bar{v}_{23}^2$$

$$\bar{v}_{23}^2 = 0.42$$

$$\bar{p}_{23}^2 = 2.43 + 0.42 = 2.85$$

Untuk $t = 1, i = 2, l = 4$, maka:

$$\bar{v}_{24}^2 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(15.6 - 15.6) + 2.0 \cdot 0.9(20.9 - 15.6)$$

$$\bar{p}_{24}^2 = 15.6 + \bar{v}_{24}^2$$

$$\bar{v}_{24}^2 = 9.54$$

$$\bar{p}_{24}^2 = 15.6 + 9.54 = 25.14$$

Maka, posisi *centroid* baru untuk *cluster* 3 adalah

$$(12.46, 3.18, 2.85, 25.14)$$

Proses *update* untuk partikel 3 ($\bar{p}_3$) adalah sebagai berikut:

Cluster 1:

Posisi *centroid* saat ini: (14.2, 1.7, 2.4, 15.5)

$$\bar{p}_{best}: (14.2, 1.7, 2.4, 15.5)$$

$$\bar{g}_{best}: (14.2, 1.7, 2.4, 15.5) \text{ diambil dari } \bar{p}_3$$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 1, i = 3, l = 1$, maka:

$$\bar{v}_{31}^2 = 0.5 \cdot (-0.1) + 2.0 \cdot 0.7(14.2 - 14.2) + 2.0 \cdot 0.9(14.2 - 14.2)$$

$$\bar{p}_{31}^2 = 14.2 + \bar{v}_{31}^2$$

$$\bar{v}_{31}^2 = -0.05$$

$$\bar{p}_{31}^2 = 14.2 + (-0.05) = 14.15$$

Untuk $t = 1, i = 3, l = 2$, maka:

$$\bar{v}_{32}^2 = 0.5 \cdot (0.2) + 2.0 \cdot 0.7(1.7 - 1.7) + 2.0 \cdot 0.9(1.7 - 1.7)$$

$$\bar{p}_{32}^2 = 1.7 + \bar{v}_{32}^2$$

$$\bar{v}_{32}^2 = 0.1$$

$$\bar{p}_{32}^2 = 1.7 + 0.1 = 1.8$$

Untuk $t = 1, i = 3, l = 3$, maka:

$$\bar{v}_{33}^2 = 0.5 \cdot (-0.5) + 2.0 \cdot 0.7(2.4 - 2.4) + 2.0 \cdot 0.9(2.4 - 2.4)$$

$$\bar{p}_{33}^2 = 2.4 + \bar{v}_{33}^2$$

$$\bar{v}_{33}^2 = -0.25$$

$$\bar{p}_{33}^2 = 2.4 + (-0.25) = 2.15$$

Untuk $t = 1, i = 3, l = 4$, maka:

$$\bar{v}_{34}^2 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(15.5 - 15.5) + 2.0 \cdot 0.9(15.5 - 15.5)$$

$$\bar{p}_{24}^2 = 15.5 + \bar{v}_{34}^2$$

$$\bar{v}_{34}^2 = 0$$

$$\bar{p}_{24}^2 = 15.5 + 0 = 15.5$$

Maka, posisi *centroid* baru untuk *cluster* 1 adalah

$$(14.15, 1.8, 2.15, 15.5)$$

Cluster 2:

Posisi *centroid* saat ini: (13.1, 2.3, 2.6, 18.5)

$$\bar{p}_{best}: (13.1, 2.3, 2.6, 18.5)$$

$$\bar{g}_{best}: (13.1, 2.3, 2.6, 18.5) \text{ dari } \bar{p}_3$$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 1, i = 3, l = 1$, maka:

$$\bar{v}_{31}^2 = 0.5 \cdot (0.2) + 2.0 \cdot 0.7(13.1 - 13.1) + 2.0 \cdot 0.9(13.1 - 13.1)$$

$$\bar{p}_{31}^2 = 13.1 + \bar{v}_{31}^2$$

$$\bar{v}_{31}^2 = 0.1$$

$$\bar{p}_{31}^2 = 13.1 + 0.1 = 13.2$$

Untuk $t = 1, i = 3, l = 2$, maka:

$$\bar{v}_{32}^2 = 0.5 \cdot (-0.1) + 2.0 \cdot 0.7(2.3 - 2.3) + 2.0 \cdot 0.9(2.3 - 2.3)$$

$$\bar{p}_{32}^2 = 2.3 + \bar{v}_{32}^2$$

$$\bar{v}_{32}^2 = -0.05$$

$$\bar{p}_{32}^2 = 2.3 + (-0.05) = 2.25$$

Untuk $t = 1, i = 3, l = 3$, maka:

$$\bar{v}_{33}^2 = 0.5 \cdot (0.1) + 2.0 \cdot 0.7(2.6 - 2.6) + 2.0 \cdot 0.9(2.6 - 2.6)$$

$$\bar{p}_{33}^2 = 2.6 + \bar{v}_{33}^2$$

$$\bar{v}_{33}^2 = 0.05$$

$$\bar{p}_{33}^2 = 2.6 + 0.05 = 2.65$$

Untuk $t = 1, i = 3, l = 4$, maka:

$$\bar{v}_{34}^2 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(18.5 - 18.5) + 2.0 \cdot 0.9(18.5 - 18.5)$$

$$\bar{p}_{24}^2 = 18.5 + \bar{v}_{34}^2$$

$$\bar{v}_{34}^2 = 0$$

$$\bar{p}_{24}^2 = 18.5 + 0 = 18.5$$

Maka, posisi *centroid* baru untuk *cluster* 2 adalah

$$(13.2, 2.25, 2.65, 18.5)$$

Cluster 3:

Posisi *centroid* saat ini: $= (13.3, 2.5, 2.8, 20.9)$

$$\bar{p}_{best}: = (13.3, 2.5, 2.8, 20.9)$$

$$\bar{g}_{best}: (13.3, 2.5, 2.8, 20.9) \text{ dari } \bar{p}_1$$

Menghitung kecepatan dan posisi baru menggunakan persamaan (3.5) dan persamaan (3.6):

Untuk $t = 1, i = 3, l = 1$, maka:

$$\bar{v}_{31}^2 = 0.5 \cdot (-0.2) + 2.0 \cdot 0.7(13.3 - 13.3) + 2.0 \cdot 0.9(13.3 - 13.3)$$

$$\bar{p}_{31}^2 = 13.3 + \bar{v}_{31}^2$$

$$\bar{v}_{31}^2 = -0.1$$

$$\bar{p}_{31}^2 = 13.3 + (-0.1) = 13.2$$

Untuk $t = 1, i = 3, l = 2$, maka:

$$\bar{v}_{32}^2 = 0.5 \cdot (0.1) + 2.0 \cdot 0.7(2.5 - 2.5) + 2.0 \cdot 0.9(2.5 - 2.5)$$

$$\bar{p}_{32}^2 = 2.5 + \bar{v}_{32}^2$$

$$\bar{v}_{32}^2 = 0.05$$

$$\bar{p}_{32}^2 = 2.5 + 0.05 = 2.55$$

Untuk $t = 1, i = 3, l = 3$, maka:

$$\bar{v}_{33}^2 = 0.5 \cdot (-0.5) + 2.0 \cdot 0.7(2.8 - 2.8) + 2.0 \cdot 0.9(2.8 - 2.8)$$

$$\bar{p}_{33}^2 = 2.8 + \bar{v}_{33}^2$$

$$\bar{v}_{33}^2 = -0.25$$

$$\bar{p}_{33}^2 = 2.8 + (-0.25) = 2.55$$

Untuk $t = 1, i = 3, l = 4$, maka:

$$\bar{v}_{34}^2 = 0.5 \cdot (0.0) + 2.0 \cdot 0.7(20.9 - 20.9) + 2.0 \cdot 0.9(20.9 - 20.9)$$

$$\bar{p}_{24}^2 = 0.2 + \bar{v}_{34}^2$$

$$\bar{v}_{34}^2 = 0$$

$$\bar{p}_{24}^2 = 20.9 + 0 = 20.9$$

Maka, posisi *centroid* baru *cluster* 3: (13.2, 2.55, 2.55, 20.9)

Implementasi menggunakan pemrograman dilakukan hingga 30 iterasi dan menghasilkan *centroid* terbaik (*Final $\bar{g}_{best}$*) sebagai berikut:

$$\bar{\mu}_1 = (14.40, 1.17, 2.91, 14.74)$$

$$\bar{\mu}_2 = (12.58, 2.02, 4.04, 14.24)$$

$$\bar{\mu}_3 = (12.21, 2.85, 4.77, 14.85)$$

dengan nilai nilai $\bar{g}_{best}$ pada iterasi ke-30 sebesar 6.12.

4.3.4 Dataset Wine (Menggunakan Seluruh Data)

Hasil PSO pada dataset Wine secara keseluruhan dilakukan hingga 50 iterasi untuk memperoleh *centroid* terbaik. Setelah iterasi ke-50, diperoleh nilai *Final $\bar{g}_{best}$* (*centroid* terbaik) sebagai berikut:

$$\bar{\mu}_1 = (-0.22, 0.64, -0.01, 0.35, -0.33, -0.53, -0.03, 0.39, 0.23, -0.09, -0.23, -0.49, -0.54)$$

$$\bar{\mu}_2 = (0.60, 0.34, 0.30, -1.11, 0.14, 0.46, 0.98, -0.66, 0.63, 0.51, 0.50, 0.23, 1.41)$$

$$\bar{\mu}_3 = (-1.26, 1.31, -0.18, 2.10, 2.83, 0.62, 0.90, -0.10, 0.88, 1.21, 0.89, 0.36, 1.46)$$

dengan nilai nilai $\bar{g}_{best}$ pada iterasi ke-50 sebesar 1809.71

4.4 Implementasi *K-Means* - PSO

4.4.1 *K-Means* - PSO Pada Dataset Iris (5 Data Pertama)

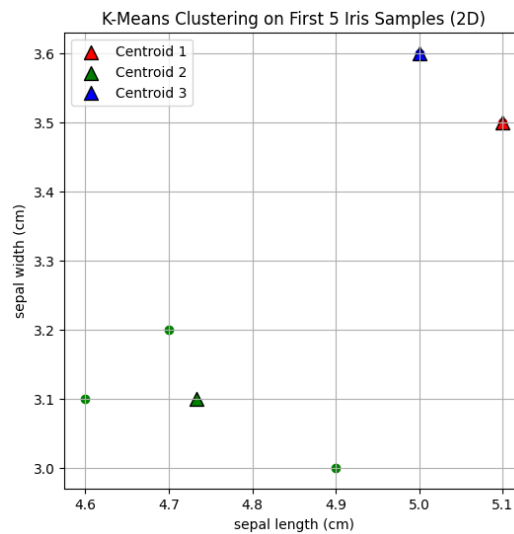
Pada metode *K-Means* yang dioptimalkan dengan PSO, inisialisasi *centroid* awal diperoleh melalui optimasi menggunakan PSO sebelum proses *clustering* dimulai. PSO membantu menentukan titik pusat awal yang lebih optimal. Pada perhitungan ini tidak dilakukan perhitungan manual, karena langkah-langkah nya sama seperti *K-Means* standar yang sudah dilakukan sebelumnya, hanya saja pada langkah inisialisasi *centroid* tidak diambil secara acak lagi melainkan menggunakan *centroid* yang didapat dari perhitungan PSO sebelumnya.

Perhitungan ini mencapai konvergensi hingga 2 iterasi dan diperoleh nilai *final centroid* yaitu:

$$\bar{\mu}_1 = (5.1, 3.5, 1.4, 0.2)$$

$$\bar{\mu}_2 = (4.7, 3.1, 1.4, 0.2)$$

$$\bar{\mu}_3 = (5.0, 3.6, 1.4, 0.2)$$



Gambar 4.5 Hasil *K-Means* – PSO Dataset Iris (5 Data Pertama)

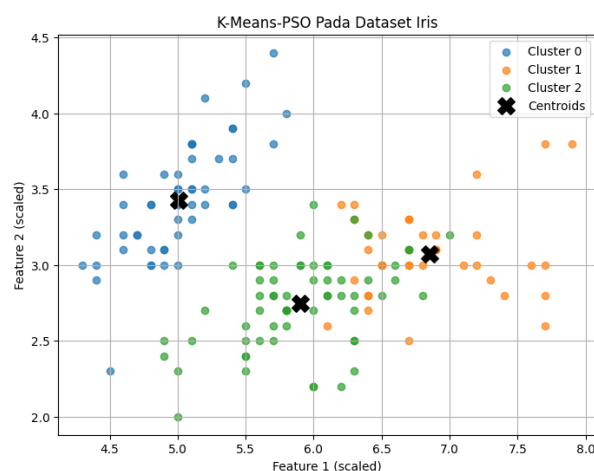
4.4.2 *K-Means* - PSO Pada Dataset Iris (Seluruh Data)

Pada dataset Iris untuk keseluruhan data dilakukan perhitungan menggunakan pemrograman *Python* yang bisa dilihat melalui tautan berikut: <https://tinyurl.com/Dataset-Iris>.

Berdasarkan lima kali percobaan yang dilakukan oleh penulis, diperoleh hasil *clustering* yang konsisten, baik dari sisi jumlah iterasi, nilai SSE, nilai *Silhouette Score*, maupun akurasi. Jumlah iterasi yang dihasilkan untuk mencapai konvergen pada setiap percobaan adalah sebanyak 11 iterasi. Nilai SSE yang dihasilkan sebesar 78.85, menunjukkan bahwa nilai tersebut sedikit lebih baik dibandingkan dengan *K-Means* standar, yang menunjukkan bahwa penggunaan PSO dalam menentukan *centroid* awal membantu meningkatkan kepadatan atau kekompakan *cluster*.

Selain itu, nilai *Silhouette Score* yang diperoleh adalah 0.553 menunjukkan bahwa nilai ini sedikit lebih tinggi dari *K-Means* standar, dan akurasi nya sebesar

89.3%, yang berarti sebagian besar data berhasil dikelompokkan ke kluster yang sesuai dengan kelas aslinya. Artinya metode *K-Means* + PSO menunjukkan performa yang lebih baik secara keseluruhan dibandingkan *K-Means* standar. Peningkatan terlihat pada nilai SSE, *Silhouette Score*, dan akurasi. Hal ini menunjukkan bahwa PSO efektif dalam membantu *K-Means* menemukan inisialisasi *centroid* yang lebih optimal, sehingga menghasilkan hasil pengelompokan yang lebih akurat dan terstruktur.



Gambar 4.6 Hasil *K-Means* – PSO Dataset Iris

4.4.3 *K-Means* - PSO Pada Dataset Wine (5 Data Pertama)

Pada dataset Wine, penggunaan PSO untuk inisialisasi *centroid* pada *K-Means* menunjukkan peningkatan efisiensi dalam proses *clustering*. Algoritma *K-Means* PSO mencapai konvergensi hanya dalam 2 iterasi. Sebaliknya, *K-Means* standar yang menggunakan inisialisasi *centroid* acak memerlukan 4 iterasi sebelum mencapai kondisi stabil.

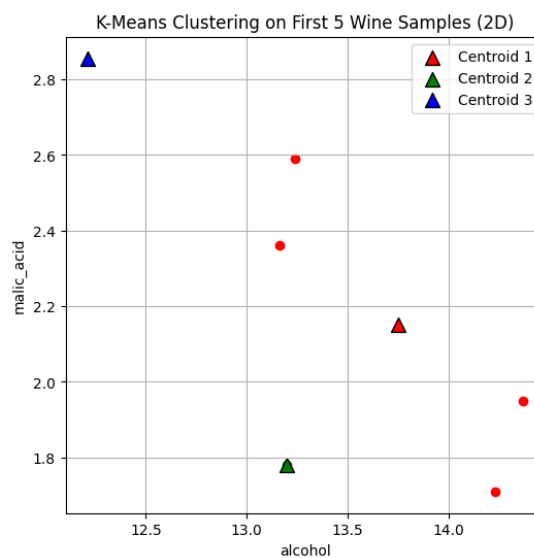
Hasil ini menunjukkan bahwa inisialisasi *centroid* yang lebih baik dapat mempercepat proses konvergensi *K-Means*, mengurangi jumlah iterasi yang diperlukan, dan meningkatkan efisiensi komputasi. Dengan iterasi yang lebih

sedikit, penggunaan PSO dapat menghemat waktu pemrosesan, yang sangat bermanfaat terutama untuk dataset dengan dimensi yang lebih tinggi atau jumlah data yang besar. Selain itu, *centroid* akhir yang diperoleh juga menunjukkan stabilitas yang lebih baik dalam membentuk *cluster* yang optimal.

$$\bar{\mu}_1 = (13.75, 2.15, 2.61, 18)$$

$$\bar{\mu}_2 = (13.2, 1.78, 2.14, 11.2)$$

$$\bar{\mu}_3 = (12.21, 2.85, 4.77, 14.83)$$



Gambar 4.7 Hasil *K-Means* – PSO Dataset Wine (5 Data Pertama)

4.4.4 *K-Means* - PSO Pada Dataset Wine (Seluruh Data)

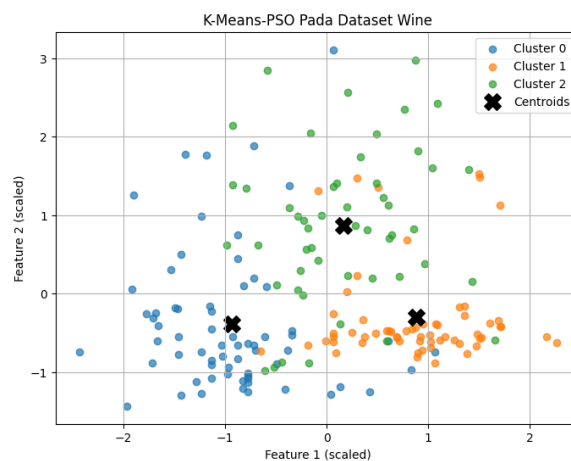
Pada dataset Wine untuk keseluruhan data dilakukan perhitungan menggunakan pemrograman *Python* yang bisa dilihat melalui tautan berikut: <https://tinyurl.com/Dataset-Wine>.

Berdasarkan lima kali percobaan yang dilakukan oleh penulis, diperoleh hasil *clustering* yang konsisten, baik dari sisi jumlah iterasi, nilai SSE, nilai *Silhouette Score*, maupun akurasi. Jumlah iterasi yang dihasilkan untuk mencapai konvergen pada setiap percobaan adalah sebanyak 3 iterasi. Nilai SSE yang

dihasilkan sebesar 1278.76, menunjukkan bahwa data dalam tiap klaster relatif dekat terhadap *centroid* masing-masing.

Selain itu, nilai *Silhouette Score* yang diperoleh adalah 0.286, dan akurasi nya sebesar 97.2%, yang berarti sebagian besar data berhasil dikelompokkan ke klaster yang sesuai dengan kelas aslinya.

Berdasarkan penjelasan di atas, hasil pada *K-Means* yang dioptimasi dengan PSO terdapat perubahan yakni pada nilai SSE, *Silhouette Score*, dan akurasi. Dimana terlihat ada kenaikan pada akurasi nya, dan *Silhouette Score*.



Gambar 4.8 Hasil *K-Means* – PSO Dataset Wine

4.5 Evaluasi dan Validasi Hasil

4.5.1 Evaluasi Hasil *Clustering* pada Dataset Iris

Tabel 4.20 Rata-rata Evaluasi *K-Means* pada Dataset Iris

Metode	SSE (rata-rata)	Silhouette Score (rata-rata)	Akurasi (rata-rata)
<i>K-Means</i>	78.86	0.551	88.7%
<i>K-Means</i> + PSO	78.85	0.553	89.3%

Berdasarkan tabel 4.20, dapat disimpulkan bahwa metode *K-Means* + PSO mampu menghasilkan performa *clustering* yang sedikit lebih baik. Hal ini ditunjukkan dari nilai SSE rata-rata yang lebih rendah, yaitu 78.85 pada *K-Means*

+ PSO dibandingkan 78.86 pada *K-Means* standar. Nilai SSE yang lebih rendah mengindikasikan bahwa hasil *clustering* memiliki jarak yang lebih kecil terhadap *centroid*-nya, sehingga lebih kompak.

Selain itu, nilai *Silhouette Score* rata-rata pada metode *K-Means* + PSO juga lebih tinggi yaitu 0.553, dibandingkan 0.551 pada *K-Means* standar. Peningkatan ini menunjukkan bahwa pengelompokan dihasilkan oleh *K-Means* + PSO memiliki batas antar *cluster* yang lebih jelas dan terdefinisi lebih baik.

Kemudian dari segi akurasi, metode *K-Means* + PSO memperoleh nilai rata-rata 89.3%, lebih tinggi dibandingkan *K-Means* standar yang hanya mencapai 88.7%. Ini berarti metode yang dioptimasi dengan PSO mampu menghasilkan pengelompokan yang lebih sesuai dengan label sebenarnya

Secara keseluruhan, penerapan algoritma PSO dalam menentukan inisialisasi *centroid* pada metode *K-Means* memberikan hasil *clustering* yang lebih baik. Hal ini terlihat dari kelompok data yang lebih rapat, pemisahan antar kelompok yang lebih jelas, dan tingkat akurasi yang lebih tinggi. Oleh karena itu, menggabungkan PSO dengan *K-Means* bisa menjadi cara yang efektif untuk meningkatkan kualitas pengelompokan data.

4.5.2 Evaluasi Hasil *Clustering* pada Dataset Wine

Tabel 4.21 Rata-rata Evaluasi *K-Means* pada Dataset Wine

Metode	SSE (rata-rata)	<i>Silhouette Score</i> (rata-rata)	Akurasi (rata-rata)
<i>K-Means</i>	1277.93	0.285	96.6%
<i>K-Means</i> + PSO	1278.76	0.286	97.2%

Berdasarkan tabel 4.21, menunjukkan hasil evaluasi rata-rata metode *K-Means* standar dan *K-Means* + PSO pada dataset Wine. Berdasarkan nilai SSE, metode *K-Means* standar menghasilkan nilai lebih rendah yaitu 1277.93,

dibandingkan dengan *K-Means* + PSO yang memperoleh 1278.76. Meskipun nilai SSE pada *K-Means* standar lebih kecil, selisihnya sangat kecil sehingga perbedaannya tidak terlalu signifikan. Namun, pada *Silhouette Score*, yang menunjukkan seberapa baik objek berada dalam *cluster*-nya, *K-Means* + PSO memiliki nilai lebih tinggi yaitu 0.286 dibandingkan 0.285 pada *K-Means* standar. Ini menunjukkan bahwa PSO sedikit membantu menghasilkan pengelompokan yang lebih terpisah dan terstruktur. Dari segi akurasi, *K-Means* + PSO juga unggul dengan nilai 97.2%, sedangkan *K-Means* standar mencapai 96.6%. Artinya, metode yang dioptimasi dengan PSO mampu mengelompokkan data dengan lebih sesuai terhadap label sebenarnya. Secara keseluruhan, meskipun peningkatannya tidak terlalu besar, integrasi PSO ke dalam algoritma *K-Means* mampu memberikan hasil *clustering* yang sedikit lebih baik, terutama dari segi pemisahan *cluster* dan akurasi.

4.5.3 Tabel Klasifikasi/*Clustering* Data Iris

Tabel 4.22 Perbandingan Evaluasi pada Penelitian Sebelumnya

Metode	Akurasi	SSE	<i>Silhouette Score</i>	Waktu
SVM (<i>Polynomial Kernel</i>) (Desiani et al., 2019)	96.7%	-	-	-
SVM (<i>Optimized</i> dengan GA) (Hussain et al., 2020)	98.7%	-	-	-
<i>K-Nearest Neighbor</i> (Tuntun et al., 2022)	95.33%	-	-	-
<i>K-Nearest Neighbor</i> (Rahman et al., 2023)	97%	-	-	-
<i>Random Forest</i> (Rahman et al., 2023)	100%	-	-	-
PSO -LVQ (Romadhona et al., 2018)	93.334%	-	-	0.3 detik
SVM (Muhammad et al., 2022)	95%	-	-	-
<i>K-Nearest Neighbor</i> (Mutrofin et al., 2014)	83.88%	-	-	-
<i>K-Means</i> (Vysala & Gomes, 2020)	-	-	0.55 (<i>K</i> =3)	-
<i>K-Means</i> +++ (Susanto, 2016)	89.33%	-	0.735	-
DBSCAN (Aref et al., 2020)	86.4%	-	-	-
<i>K-Means</i> + PSO	89.3%	78.85	0.553	0.006 detik

Berdasarkan tabel 4.22 di atas, terlihat bahwa sebagian besar penelitian sebelumnya menggunakan pendekatan klasifikasi, tetapi terdapat beberapa penelitian menggunakan pendekatan *clustering*. Namun demikian, dari penelitian yang telah di sebutkan terdapat perbedaan mencolok yaitu kelengkapan evaluasi. Dalam banyak penelitian sebelumnya, informasi evaluasi seperti SSE, Silhouette Score, dan waktu komputasi tidak lengkap atau bahkan tidak dicantumkan sama sekali.

Oleh karena itu, dapat disimpulkan bahwa dari segi akurasi, metode *Random Forest* menunjukkan performa terbaik dengan akurasi sebesar 100%. Hal ini dapat terjadi karena dataset Iris merupakan dataset yang secara umum lebih sesuai untuk pendekatan klasifikasi dan memang sering digunakan dalam berbagai studi klasifikasi. Oleh karena itu, pendekatan *clustering* yang digunakan oleh penulis tidak menghasilkan performa yang lebih baik dibandingkan metode klasifikasi seperti *Random Forest*. Meskipun demikian, akurasi dari pendekatan *clustering* yang dilakukan penulis, yaitu metode *K-Means* + PSO, menunjukkan performa yang lebih baik dibandingkan metode *clustering* lainnya, seperti DBSCAN.

Kemudian dari sisi SSE, yang merupakan indikator seberapa baik data dikelompokkan ke dalam klaster, nilai yang lebih kecil menunjukkan hasil *clustering* yang lebih baik. Hasil menunjukkan bahwa *K-Means* + PSO memiliki nilai SSE paling rendah, yaitu 78.85.

Selanjutnya, dilihat dari *Silhouette Score* yang lebih baik yaitu pada metode *K-Means* +++ dengan nilai sebesar 0.735, dan terakhir waktu komputasi yang lebih cepat adalah 0.0063 detik pada *K-Means* + PSO.

Dengan demikian, dapat disimpulkan bahwa masing-masing metode memiliki keunggulan tersendiri. *Random Forest* unggul dalam klasifikasi dengan akurasi sempurna, sementara untuk metode *clustering*, *K-Means* +++ unggul dalam kualitas klaster, dan *K-Means* + PSO unggul dalam efisiensi waktu serta nilai SSE yang paling optimal.

4.5.4 Tabel Klasifikasi/*Clustering* Data Wine

Tabel 4.23 Perbandingan Evaluasi pada Penelitian Sebelumnya

Metode	Akurasi	SSE	Silhouette Score	Waktu
<i>K-Nearest Neighbor</i> (Mutrofin et al., 2014)	84.24%	-	-	-
K-NN(Z-Score, tanpa citric acid) (Khakim et al., 2023)	75.62%	-	-	-
SVM (Decimal Scaling) (Adams et al., 2021)	57.10%	-	-	-
Decision Tree (Min-Max Scaling) (Adams et al., 2021)	46.70%	-	-	-
<i>K-Means</i> (Suraya et al., 2023)	-	-	0.284 ($K=3$)	-
<i>K-Means</i> +++ (Susanto, 2016)	70.22%%	-	0.728	-
DBSCAN (Aref et al., 2020)	58.1%	-	-	-
<i>K-Means</i> + PSO	97.2%	1278.76	0.286	0.0124 detik

Berdasarkan tabel 4.23, dapat dilihat bahwa metode dengan akurasi tertinggi adalah *K-Means* + PSO, yaitu sebesar 97.2%. Hal ini menunjukkan bahwa pada dataset yang digunakan dalam penelitian sebelumnya, pendekatan *clustering*, khususnya *K-Means* yang dioptimasi dengan PSO, dapat menghasilkan akurasi yang lebih tinggi dibandingkan dengan metode klasifikasi seperti SVM, KNN, dan lainnya. Hal ini dapat terjadi karena dataset Wine memang sering digunakan dalam studi-studi *clustering*, sehingga pendekatan tanpa supervisi dapat memberikan hasil yang kompetitif bahkan melampaui metode klasifikasi tertentu. Meskipun demikian, akurasi dari pendekatan *clustering* yang dilakukan penulis, yaitu metode *K-Means* + PSO, menunjukkan performa yang lebih baik dibandingkan metode *clustering* lainnya, seperti DBSCAN.

Kemudian dari segi nilai dengan nilai SSE nilai terbaik (terendah) dicapai oleh *K-Means* + PSO yang memiliki nilai 1278.76. Selanjutnya, *Silhouette Score* terbaik yaitu sebesar 0.728 pada metode *K-Means* + PSO. Terakhir, waktu komputasi tercepat yaitu pada metode *K-Means* + PSO memerlukan waktu 0.0124 detik.

Secara keseluruhan, dapat disimpulkan bahwa *K-Means* + PSO unggul dari nilai SSE dan segi akurasi, serta waktu komputasi. Sedangkan *K-Means* + PSO memberikan kualitas kluster terbaik berdasarkan *Silhouette Score*.

Keterbatasan dari metode yang digunakan penulis terletak pada metode *K-Means*. Salah satu kelemahan utama dari *K-Means* adalah hasil klusterisasi sangat bergantung pada pemilihan pusat kluster (*centroid*) di awal. Jika pemilihan *centroid* awal tidak tepat, maka algoritma dapat menghasilkan kluster yang kurang optimal karena mudah terjebak pada solusi lokal. Untuk mengatasi masalah tersebut, metode *K-Means* kemudian dioptimasi menggunakan PSO. Pendekatan ini bertujuan untuk mencari posisi *centroid* yang lebih baik secara global, sehingga diharapkan dapat meningkatkan kualitas hasil klusterisasi.

Namun, metode *K-Means* + PSO juga memiliki beberapa keterbatasan. Proses optimasi menggunakan PSO memerlukan waktu komputasi yang lebih lama karena melibatkan iterasi berulang untuk mencari solusi terbaik. Selain itu, hasil yang diperoleh dari PSO sangat bergantung pada pengaturan parameter seperti jumlah partikel, nilai inersia (kecepatan), dan jumlah iterasi. Jika parameter-parameter ini tidak diatur dengan tepat, maka hasil optimasi juga bisa menjadi kurang maksimal.

4.6 Kajian Penelitian dalam Perspektif Islam

Terdapat ayat Al-Qur'an QS. Al-Infithar ayat 13-14, di mana Allah SWT menjelaskan tentang pengelompokan manusia pada hari kiamat, manusia terbagi menjadi dua golongan utama.

إِنَّ الْأَبْرَارَ لَفِي نَعِيمٍ ﴿١٣﴾ وَإِنَّ الْفُجَّارَ لَفِي جَحِيمٍ ﴿١٤﴾

(Kemenag, 2024)

Artinya: “*Sesungguhnya orang-orang yang berbakti benar-benar berada dalam (surga yang penuh) kenikmatan, dan sesungguhnya orang-orang yang durhaka benar-benar berada dalam (neraka) Jahim.*”

Menurut Tafsir Ibnu Katsir (Al-Sheikh, 2005) golongan-golongan tersebut ialah:

1. *Al-Abrar*

Mereka yang beriman dengan benar, beramal saleh, dan menjaga akhlak mulia dalam interaksi sesama manusia. Mereka adalah individu yang menghiasi hidupnya dengan ketaatan kepada Allah SWT dan kepedulian terhadap sesama. Karena itu, dalam ayat Al-Qur'an (QS. Al-Infithar:13), mereka digambarkan sebagai orang-orang yang akan memperoleh kenikmatan di surga, yang merupakan sebuah balasan atas keimanan dan amal baik yang mereka tanamkan sepanjang hidup.

2. *Al-Fujjar*

Golongan manusia yang hidup dalam kedurhakaan, melakukan kemaksiatan, dan mengabaikan perintah Allah SWT. Mereka lebih mementingkan hawa nafsu, duniawi, dan cenderung melakukan kerusakan serta kezaliman. Oleh sebab itu, dalam QS. Al-Infithar ayat 14, mereka digambarkan sebagai penghuni neraka Jahim, tempat siksaan yang sangat

pedih, sebagai konsekuensi dari kelalaian dan dosa-dosa besar yang mereka lakukan selama hidup.

Jika dikaitkan dengan konsep *clustering* dalam analisis data, ayat-ayat ini secara tegas menunjukkan bahwa manusia secara alami dapat dikelompokkan berdasarkan atribut perilaku dan keimanan. Dalam hal ini, atribut seperti keimanan, amal perbuatan, dan perilaku sosial menjadi fitur utama dalam membentuk klaster. Dengan menggunakan metode *K-Means Clustering*, misalnya, manusia dapat dipetakan ke dalam dua *centroid* utama. Satu *centroid* mewakili karakteristik *Al-Abrar* (keimanan tinggi, amal saleh, perilaku baik) dan satu *centroid* lagi mewakili karakteristik *Al-Fujjar* (keimanan rendah atau tidak ada, amal buruk, perilaku tercela). Proses *clustering* ini menyerupai pemisahan manusia di hari kiamat berdasarkan kualitas iman dan amal mereka, sebagaimana dijelaskan dalam Al-Qur'an.

Misal:

Data manusia dinotasikan sebagai:

$$X = (\bar{x}_1, \bar{x}_2, \bar{x}_3, \dots, \bar{x}_m)$$

dengan masing-masing $\bar{x}_i$ adalah vektor fitur:

$$\bar{x}_i = (x_{i1}, x_{i2}, \dots, x_{iq})$$

dengan i merupakan jumlah individu, dan q mewakili fitur, fitur yang digunakan adalah tiga, yaitu:

- x_{i1} = skor keimanan
- x_{i2} = skor amal perbuatan
- x_{i3} = skor perilaku sosial

Kemudian, menentukan jumlah klaster, $K = 2$, karena terdapat 2 golongan.

Setelah itu dilakukan inisialisasi *centroid*:

- *Centroid* untuk *Al-Abrar* ($\bar{\mu}_1$):

$$\bar{\mu}_1 = (\text{tinggi}, \text{baik}, \text{mulia})$$

- *Centroid* untuk *Al-Fujjar* ($\bar{\mu}_2$):

$$\bar{\mu}_2 = (\text{rendah}, \text{buruk}, \text{tercela})$$

Contoh di misalkan dengan nilai angka:

- “Tinggi/Mulia” = 9-10
- “Baik” = 7-8
- “Buruk/Tercela” = 1-3
- “Rendah” = 0-2

Sehingga di misalkan:

$$\bar{\mu}_1 = (9,8,9) \text{ dan } \bar{\mu}_2 = (2,2,2)$$

Kemudian dilanjutkan dengan langkah-langkah pada *K-Means Clustering* yang telah dijelaskan pada bab 3 yaitu pada subbab 3.4 Tahapan Penelitian. Jika permissalan di atas terdapat 3 manusia atau individu, dan sudah dilakukan langkah-langkah tersebut, dan misal didapatkan hasil berikut:

- $\bar{x}_1 = (9,8,9)$, artinya iman tinggi, amal baik, akhlak mulia.
- $\bar{x}_2 = (2,1,3)$, artinya iman rendah, amal buruk, perilaku buruk.
- $\bar{x}_3 = (8,7,8)$, artinya iman tinggi, amal baik, perilaku sangat baik.

Maka:

- $\bar{x}_1$ dan $\bar{x}_3$ akan masuk ke *cluster Al-Abrar*.
- $\bar{x}_2$ akan masuk ke *cluster Al-Fujjar*.

Dengan demikian, QS. Al-Infithar tidak hanya mengandung makna teologis tentang balasan akhirat, tetapi juga dapat dipahami dengan pendekatan analitis

modern seperti *clustering* dalam ilmu data. Sehingga pendekatan ini mengilustrasikan bahwa nilai-nilai agama dapat direpresentasikan ke dalam model matematis yang sistematis, yang pada gilirannya membantu kita memahami bagaimana keimanan dan amal perbuatan menjadi faktor kunci dalam menentukan kedudukan manusia di sisi Allah SWT. Ini juga menekankan pentingnya upaya untuk meningkatkan kualitas diri, agar dapat termasuk ke dalam kelompok orang-orang yang beriman dan beramal saleh.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penelitian ini bertujuan untuk mengoptimalkan algoritma *K-Means* menggunakan *Particle Swarm Optimization* (PSO), serta menganalisis dan membandingkan performanya dengan *K-Means* standar. Hasil yang diperoleh menunjukkan bahwa pada dataset Iris, metode *K-Means* yang dioptimasi dengan PSO mampu menghasilkan performa *clustering* yang lebih baik dibandingkan metode clustering lainnya, ditunjukkan oleh nilai SSE terendah, *Silhouette Score* yang lebih tinggi, dan waktu komputasi yang lebih cepat. Meskipun metode klasifikasi seperti *Random Forest* memiliki akurasi yang lebih tinggi, yaitu 100%, hal ini dapat dimaklumi karena dataset Iris lebih cocok digunakan untuk klasifikasi.

Sementara pada dataset Wine, metode *K-Means* + PSO terbukti mampu meningkatkan performa *clustering* dibandingkan *K-Means* standar, terutama dari segi akurasi dan nilai SSE. Hal ini menunjukkan bahwa penggunaan PSO dapat membantu menemukan posisi centroid yang lebih optimal, sehingga meningkatkan kualitas pengelompokan data. Namun, metode ini juga memiliki keterbatasan, seperti kebutuhan waktu komputasi yang lebih tinggi dan ketergantungan terhadap pengaturan parameter PSO. Oleh karena itu, pengaturan parameter yang tepat sangat penting agar hasil optimasi dapat memberikan performa *clustering* yang maksimal.

Secara keseluruhan, penelitian ini membuktikan bahwa *K-Means* yang dioptimasi dengan PSO memberikan hasil *clustering* yang lebih baik dan efisien

dibandingkan *K-Means* standar, meskipun memerlukan waktu komputasi lebih lama dan pengaturan parameter yang tepat agar hasilnya maksimal.

5.2 Saran untuk Penelitian Lanjutan

Berdasarkan hasil penelitian ini, penulis memberikan beberapa saran untuk pengembangan dan penelitian lebih lanjut:

1. Penelitian dapat diperluas dengan menggunakan dataset yang lebih kompleks dan berukuran besar, untuk menguji performa algoritma *K-Means-PSO* dalam skala data yang lebih luas dan kompleks.
2. Integrasi PSO dengan algoritma *clustering* lainnya selain *K-Means*, seperti DBSCAN atau *Fuzzy C-Means* juga dapat menjadi penelitian menarik, untuk dikaji agar dapat dibandingkan efektivitasnya dalam pengelompokan data.

DAFTAR PUSTAKA

- Adams, F., Dwi Anggoro, R. A., Satria, M. B., Oktavia, A. W., & Chamidah, N. (2021). Perbandingan Normalisasi Data untuk Klasifikasi Wine Menggunakan Algoritma Naïve Bayes, Decision Tree, dan Support Vector Machine. *Seminar Nasional Mahasiswa Ilmu Komputer Dan Aplikasinya (SENAMIKA)*, 2(2), 260–268.
<https://conference.upnvj.ac.id/index.php/senamika/article/view/1744>
- Al-Sheikh. (2005). Ibnu Katsir. In *Islami* (p. 27).
- Anderson, E. (1936). *The species problem in Iris*. 457.
- Aref, Y., Cemal, K. S., Asef, Y., & Amir, S. (2020). Automatic fuzzy-DBSCAN algorithm for morphological and overlapping datasets. *Journal of Systems Engineering and Electronics*, 31(6), 1245–1253.
<https://doi.org/10.23919/JSEE.2020.000095>
- Benabdellah, A. C., Benghabrit, A., & Bouhaddou, I. (2019). A survey of clustering algorithms for an industrial context. *Procedia Computer Science*, 148, 291–302. <https://doi.org/10.1016/j.procs.2019.01.022>
- Desiani, A., Irmeilyana, Hanum, H., Andriani, Y., Maiyanti, S. I., Uteh, C. M., & Rayani, I. (2019). Penerapan Metode Support Vector Machine Dalam Klasifikasi Bunga Iris. *Sustainability (Switzerland)*, 11(1), 1–14.
- Dey, L., & Chakraborty, S. (2014). Canonical PSO Based K -Means Clustering Approach for Real Datasets . *International Scholarly Research Notices*, 2014, 1–11. <https://doi.org/10.1155/2014/414013>
- et. all. Hastie, T., Hastie, T., Tibshirani, R., & Friedman, J. (2009). Statistics The Elements of Statistical Learning. *The Mathematical Intelligencer*, 27(2), 83–85.
- Ha, J., Kambe, M., & Pe, J. (2011). Data Mining: Concepts and Techniques. *Data Mining: Concepts and Techniques*, 1–703. <https://doi.org/10.1016/C2009-0-61819-5>
- Hussain, Z. F., Ibraheem, H. R., Alsajri, M., Ali, A. H., Ismail, M. A., Kasim, S., & Sutikno, T. (2020). A new model for iris data set classification based on linear support vector machine parameter's optimization. *International Journal of Electrical and Computer Engineering*, 10(1), 1079–1084.
<https://doi.org/10.11591/ijece.v10i1.pp1079-1084>
- J. A. Hartigan, M. A. W. (2012). Algorithm AS 136 A K-Means Clustering Algorithm. *Journal of the Royal Statistical Society Series B Methodological*, 28(1), 100–108.
- Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition*

- Letters*, 31(8), 651–666. <https://doi.org/10.1016/j.patrec.2009.09.011>
- Jamhuri, M., & Utomo, T. (2024). *Penggunaan Particle Swarm Optimization Pada Jaringan Syaraf Tiruan Untuk Klasifikasi Sinyal Radar*.
- Kemenag. (2025). *Qur'an Kemenag*.
- Khakim, E. N. R., Hermawan, A., & Avianto, D. (2023). Implementasi Correlation Matrix Pada Klasifikasi Dataset Wine. *JIKO (Jurnal Informatika Dan Komputer)*, 7(1), 158. <https://doi.org/10.26798/jiko.v7i1.771>
- Muhammad, O. :, Pradana, R., Rio,), Hardinata, S., Syahputra, Z., Sains, F., Teknologi, D., Pembangunan, U., & Medan, P. (2022). *Analisa Algoritma Support Vector Machine Pada Data Bunga Iris*. 30(1), 477–487.
- Muslimatin, B. (2011). *Perbandingan K-Means dan Fuzzy C-Means(FCM) untuk Cluster Data*.
- Mutrofin, S., Izzah, A., Kurniawardhani, A., & Masrur, M. (2014). Optimasi Teknik Klasifikasi Modified K Nearest Neighbor Menggunakan Algoritma Genetika. *Jurnal Gamma*, January 2016, 130–134.
- Nafisa, N. A. (2023). *K-Means Untuk Analisis Pola Penyebaran Tempat K-Means Untuk Analisis Pola Penyebaran Tempat*.
- Nur Afidah, N. (2023). Penerapan Metode Clustering dengan Algoritma K-means untuk Pengelompokkan Data Migrasi Penduduk Tiap Kecamatan di Kabupaten Rembang. *PRISMA, Prosiding Seminar Nasional Matematika*, 6, 729–738. <https://journal.unnes.ac.id/sju/index.php/prisma/>
- R. A. Fisher, Sc.D., F. R. (1936). *The Use of Multiple Measurements in Taxonomic Problems*.
- Rahman, B., Fauzi, F., & Amri, S. (2023). Perbandingan Hasil Klasifikasi Data Iris menggunakan Algoritma K-Nearest Neighbor dan Random Forest. *Journal Of Data Insights*, 1(1), 19–26. <https://doi.org/10.26714/jodi.v1i1.135>
- Romadhona, I., Cholissodin, I., & Marji. (2018). Penerapan Algoritme Particle Swarm Optimization-Learning Vector Quantization(PSO-LVQ) Pada Klasifikasi Data Iris. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(12), 6418–6428. <http://j-ptiik.ub.ac.id>
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20(C), 53–65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)
- Sheikhpour, R., Sarram, M. A., & Sheikhpour, R. (2016). Particle swarm optimization for bandwidth determination and feature selection of kernel

- density estimation based classifiers in diagnosis of breast cancer. *Applied Soft Computing Journal*, 40, 113–131. <https://doi.org/10.1016/j.asoc.2015.10.005>
- Suraya, S., Sholeh, M., & Lestari, U. (2023). Evaluation of Data Clustering Accuracy using K-Means Algorithm. *International Journal of Multidisciplinary Approach Research and Science*, 2(01), 385–396. <https://doi.org/10.59653/ijmars.v2i01.504>
- Susanto, E. B. (2016). Evaluasi Hasil Klaster Pada Dataset Iris , Soybean-small , Wine Menggunakan Algoritma Fuzzy C-Means dan K-means++. *Surya Informatika*, 2(1), 6–13.
- Tuntun, R., Kusriani, K., & Kusnawi, K. (2022). Analisis Perbandingan Kinerja Algoritma Klasifikasi dengan Menggunakan Metode K-Fold Cross Validation. *Jurnal Media Informatika Budidarma*, 6(4), 2111. <https://doi.org/10.30865/mib.v6i4.4681>
- Vysala, A., & Gomes, D. J. (2020). *Evaluating and Validating Cluster Results*. 37–47. <https://doi.org/10.5121/csit.2020.100904>
- Wu, W. T., Li, Y. J., Feng, A. Z., Li, L., Huang, T., Xu, A. D., & Lyu, J. (2021). Data mining in clinical big data: the frequently used databases, steps, and methodological models. *Military Medical Research*, 8(1), 1–12. <https://doi.org/10.1186/s40779-021-00338-z>
- YENI, M. (2022). *Analisis K-means pada Pengelompokan Kabupaten-Kota Provinsi Jawa Timur Berdasarkan Kasus Kesembuhan dan Kasus Kematian Covid-19*. 1–88.
- Yudi Agusta. (2007). K-Means-Penerapan, Permasalahan dan Metode Terkait. *Jurnal Sistem Dan Informatika*, 3(Pebruari), 47–60.
- Zaki, M. J., & Meira, W. J. (2020). Data Mining and Machine Learning Fundamental Concepts and Algorithms. In *Journal of Chemical Information and Modeling* (Vol. 53, Issue 9).

LAMPIRAN

Script python K-Means

```
# Helper functions
# -----
def init_centroids(X, k, rng):
    idx = rng.choice(len(X), k, replace=False)
    return X[idx].copy()

def pairwise_sq_dists(A, B):
    return (
        np.sum(A**2, axis=1, keepdims=True) +
        np.sum(B**2, axis=1) -
        2 * A @ B.T
    )

def assign_clusters(X, centroids):
    dists = pairwise_sq_dists(X, centroids)
    return np.argmin(dists, axis=1)

def recompute_centroids(X, labels, k, rng):
    d = X.shape[1]
    centroids = np.empty((k, d))
    for j in range(k):
        mask = labels == j
        if np.any(mask):
            centroids[j] = X[mask].mean(axis=0)
        else:
            centroids[j] = X[rng.integers(len(X))]
    return centroids

def kmeans(X, k=3, max_iter=300, tol=1e-4, random_state=None):
    """
    Run K-Means and return (labels, centroids, n_iter).

    X : (n_samples, n_features) array
    """
    rng = np.random.default_rng(random_state)
    centroids = init_centroids(X, k, rng)

    for i in range(max_iter):
        labels = assign_clusters(X, centroids)
        new_centroids = recompute_centroids(X, labels, k, rng)
        shift = np.linalg.norm(new_centroids - centroids)
        centroids = new_centroids
        if shift < tol:
            return labels, centroids, i+1
    return labels, centroids, max_iter

# Run custom K-Means
k = 3
start_time = time.time()
labels, centroids, n_iter = kmeans(X, k=k, random_state=42)
```

```

end_time = time.time()
elapsed_time = end_time - start_time
print(f"Cluster assignments: {labels[:150]}")
print(f"Converged at iteration: {n_iter}\n")
print(f"Waktu komputasi K-Means: {elapsed_time:.4f} detik\n")

#Evaluate clustering quality
# -----
-----
cm = confusion_matrix(y_true, labels)

# Best one-to-one mapping between clusters & species
row_ind, col_ind = linear_sum_assignment(-cm)
mapping = {col: row for row, col in zip(row_ind, col_ind)}
y_pred = np.array([mapping[c] for c in labels])

accuracy = (y_pred == y_true).mean()
ari = adjusted_rand_score(y_true, labels)

print("Confusion matrix (after column reordering):")
print(cm[:, col_ind])
print(f"\nAccuracy : {accuracy:.3f}")
print(f"Adjusted Rand Index: {ari:.3f}")

# Internal clustering quality measures
# -----
-----
from sklearn.metrics import silhouette_score # allowed;
uses only metrics

# --- Sum of Squared Errors (SSE) ---
def sse(X, labels, centroids):
    return np.sum((X - centroids[labels])**2)

total_sse = sse(X, labels, centroids)
print(f"Total SSE : {total_sse:.2f}")

# --- Silhouette score (needs at least 2 clusters) ---
sil_score = silhouette_score(X, labels, metric="euclidean")
print(f"Mean Silhouette Coefficient: {sil_score:.3f}")

# Plot clustering result (first two features)
plt.figure(figsize=(8, 6))
for j in range(k):
    mask = labels == j
    plt.scatter(X[mask, 0], X[mask, 1], label=f"Cluster {j}",
alpha=0.7)
# Plot centroids on same axes
plt.scatter(centroids[:, 0], centroids[:, 1], marker='X',
s=200, c='black', label='Centroids')
plt.title('K-Means Clustering Pada Dataset Iris')
plt.xlabel('Feature 1 (scaled)')
plt.ylabel('Feature 2 (scaled)')
plt.legend()
plt.grid(True)
plt.show()

```

Script python PSO

```
# --- Utility functions ---
def assign_labels(X, centroids):
    """
    Assign each sample in X to the nearest centroid.
    Returns an array of labels (shape: n_samples).
    """
    # Compute distances: shape (n_samples, n_centroids)
    dists = np.linalg.norm(X[:, np.newaxis] - centroids,
axis=2)
    # Assign to closest centroid
    return np.argmin(dists, axis=1)

def compute_sse(centroids, X):
    """
    Compute Sum of Squared Errors (SSE) for given centroids.
    """
    labels = assign_labels(X, centroids)
    # Sum squared distances of each point to its centroid
    return np.sum((X - centroids[labels]) ** 2)

# --- PSO parameters & initialization ---
num_particles = 50      # number of particles (solutions)
num_clusters = 3        # number of clusters
dim = X.shape[1]        # feature dimension
max_iter = 100          # max PSO iterations

w = 0.5      # inertia weight
a1 = 2.0     # cognitive coefficient
a2 = 2.0     # social coefficient

# Initialize particles uniformly within data range
data_min, data_max = X.min(axis=0), X.max(axis=0)
particles = np.random.uniform(low=data_min, high=data_max,
                             size=(num_particles,
num_clusters, dim))
velocities = np.zeros_like(particles)

# Initialize personal best (pbest) and global best (gbest)
pbest = particles.copy()
pbest_scores = np.full(num_particles, np.inf)
gbest = None
gbest_score = np.inf

# Evaluate initial pbest & gbest
for i in range(num_particles):
    score = compute_sse(particles[i], X)
    pbest_scores[i] = score
    if score < gbest_score:
        gbest_score = score
        gbest = particles[i].copy()

print(f"Initial gbest SSE: {gbest_score:.2f}")
```

```

# Store history of gbest for visualization
gbest_history = [gbest.copy()]

# --- PSO main loop ---
for it in range(1, max_iter + 1):
    for i in range(num_particles):
        # Update velocity and position for each centroid in
        particle i
        r1, r2 = np.random.rand(), np.random.rand()
        cognitive = a1 * r1 * (pbest[i] - particles[i])
        social = a2 * r2 * (gbest - particles[i])
        velocities[i] = w * velocities[i] + cognitive + social
        particles[i] += velocities[i]

        # Evaluate new position
        score = compute_sse(particles[i], X)
        # Update pbest
        if score < pbest_scores[i]:
            pbest_scores[i] = score
            pbest[i] = particles[i].copy()
        # Update gbest
        if score < gbest_score:
            gbest_score = score
            gbest = particles[i].copy()

    gbest_history.append(gbest.copy())
    if it % 10 == 0 or it == 1:
        print(f"Iteration {it:>2}: gbest SSE =
{gbest_score:.2f}")

# --- Final clustering & results ---
labels_final = assign_labels(X, gbest)
print(f"\nFinal gbest SSE: {gbest_score:.2f}")

# Centroids
print("Final centroids:")
print(gbest)

# --- Visualization on first two features ---
plt.figure(figsize=(8, 5))
colors = ['r', 'g', 'b']
for k in range(num_clusters):
    plt.scatter(X[labels_final == k, 0], X[labels_final == k,
1],
                s=20, alpha=0.6, label=f"Cluster {k}",
color=colors[k])
    plt.scatter(gbest[k, 0], gbest[k, 1],
                marker='x', s=100, edgecolor='k',
                label=f"Centroid {k}", color=colors[k])

plt.xlabel('Feature 1 (standardized)')
plt.ylabel('Feature 2 (standardized)')
plt.title('PSO-based Clustering on Iris (First 2 Features)')
plt.legend()
plt.tight_layout()
plt.show()

```

Script python K-Means + PSO

```

# Helper functions
# -----
def init_centroids(X, k, rng):
    """Pick k distinct observations as initial centroids."""
    idx = rng.choice(len(X), k, replace=False)
    return X[idx].copy()

def pairwise_sq_dists(A, B):
    """
    Squared Euclidean distances between each row of A (nxd)
    and each row of B (mxd) -> shape (n, m).
    """
    #  $\|a\|^2 + \|b\|^2 - 2 a \cdot b^T$  (broadcasted)
    return (
        np.sum(A**2, axis=1, keepdims=True) +
        np.sum(B**2, axis=1) -
        2 * A @ B.T
    )

def assign_clusters(X, centroids):
    """Return vector of cluster indices (0...k-1)."""
    dists = pairwise_sq_dists(X, centroids)
    return np.argmin(dists, axis=1)

def recompute_centroids(X, labels, k, rng):
    """Return new centroids after assignment (re-init empty
    clusters)."""
    d = X.shape[1]
    centroids = np.empty((k, d))
    for j in range(k):
        mask = labels == j
        if np.any(mask):
            centroids[j] = X[mask].mean(axis=0)
        else:
            # empty cluster → random
            centroids[j] = X[rng.integers(len(X))]
    return centroids

def kmeans(X, k=3, max_iter=300, tol=1e-4, random_state=None,
init_c=None):
    rng = np.random.default_rng(random_state)
    if init_c is None:
        centroids = init_centroids(X, k, rng)
    else:
        centroids = init_c.copy()

    for _ in range(max_iter):
        labels = assign_clusters(X, centroids)
        new_centroids = recompute_centroids(X, labels, k, rng)
        shift = np.linalg.norm(new_centroids - centroids)
        centroids = new_centroids
        if shift < tol:

```

```

        break

    return labels, centroids

#Run our K-Means implementation
# -----
-----
k = 3
start_time = time.time()
labels, centroids = kmeans(X, k=3, init_c=gbest)
end_time = time.time()
elapsed_time = end_time - start_time
print(f"Cluster assignments: {labels[:150]}")
print(f"Converged at iteration: {n_iter}\n")
print(f"Waktu komputasi K-Means (dengan inisialisasi PSO):
{elapsed_time:.4f} detik\n")

#Evaluate clustering quality
# -----
-----
cm = confusion_matrix(y_true, labels)

# Best one-to-one mapping between clusters & species
row_ind, col_ind = linear_sum_assignment(-cm)
mapping = {col: row for row, col in zip(row_ind, col_ind)}
y_pred = np.array([mapping[c] for c in labels])

accuracy = (y_pred == y_true).mean()
ari = adjusted_rand_score(y_true, labels)

print("Confusion matrix (after column reordering):")
print(cm[:, col_ind])
print(f"\nAccuracy : {accuracy:.3f}")
print(f"Adjusted Rand Index: {ari:.3f}")

# Internal clustering quality measures
# -----
-----
from sklearn.metrics import silhouette_score # allowed;
uses only metrics

# --- Sum of Squared Errors (SSE) ---
def sse(X, labels, centroids):
    return np.sum((X - centroids[labels])**2)

total_sse = sse(X, labels, centroids)
print(f"Total SSE : {total_sse:.2f}")

# --- Silhouette score (needs at least 2 clusters) ---
sil_score = silhouette_score(X, labels, metric="euclidean")
print(f"Mean Silhouette Coefficient: {sil_score:.3f}")

# Plot clustering result (first two features)
plt.figure(figsize=(8, 6))
for j in range(k):

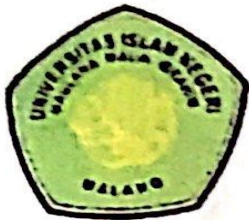
```

```
    mask = labels == j
    plt.scatter(X[mask, 0], X[mask, 1], label=f"Cluster {j}",
alpha=0.7)
# Plot centroids on same axes
plt.scatter(centroids[:, 0], centroids[:, 1], marker='x',
s=200, c='black', label='Centroids')
plt.title('K-Means-PSO Pada Dataset Iris')
plt.xlabel('Feature 1 (scaled)')
plt.ylabel('Feature 2 (scaled)')
plt.legend()
plt.grid(True)
plt.show()
```

RIWAYAT HIDUP



Muthia Afifi atau lebih dikenal sebagai Muthia, lahir di Jakarta pada 29 Desember 2002. Penulis merupakan anak ketiga dari pasangan Bapak Mursal Pili dan Ibu Agustina. Selama masa pendidikan, penulis menempuh pendidikan dasar di SDN Rawa Buaya 02 Pagi dan lulus pada tahun 2015. Kemudian penulis melanjutkan pendidikan menengah pertama di SMPN 205 Jakarta, lulus pada tahun 2018. Selanjutnya, penulis menimba ilmu agama di Pondok Pesantren Daar El Qolam 3 dan menempuh pendidikan menengah atas di SMAS Daar El Qolam 2, lulus pada tahun 2021. Setelah dinyatakan lulus dari SMA, pada tahun yang sama penulis melanjutkan pendidikan di bangku perkuliahan sebagai mahasiswa program studi Matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Selama masa perkuliahan penulis cukup aktif dalam organisasi kampus. Salah satunya sebagai anggota divisi Internal Public Relation (IPR) Himpunan Mahasiswa Program Studi Matematika (HMPS) “Integral” Matematika pada tahun 2022. Apabila terdapat pertanyaan, saran, ataupun kritik dari penelitian ini, penulis dapat dihubungi melalui email (muthiaafifi298@gmail.com) atau sosial media instagram (<https://instagram.com/muthhaf>).

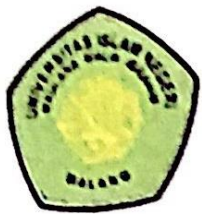


KEMENTERIAN AGAMA RI
UNIVERSITAS ISLAM NEGERI
MAULANA MALIK IBRAHIM MALANG
FAKULTAS SAINS DAN TEKNOLOGI
Jl. Gajayana No.50 Dinoyo Malang Telp. / Fax. (0341)558933

BUKTI KONSULTASI SKRIPSI

Nama : Muthia Afifi
NIM : 210601110107
Fakultas/Jurusan : Sains dan Teknologi/Matematika
Judul Skripsi : Optimasi Algoritma *K-Means* Menggunakan *Particle Swarm Optimization*
Pembimbing I : Dr. Mohammad Jamhuri, M.Si
Pembimbing II : Erna Herawati, M.Pd

No	Tanggal	Hal	Tanda Tangan
1.	03 September 2024	Konsultasi Topik	1.
2.	11 September 2024	Konsultasi Bab I, II, dan III	2.
3.	17 September 2024	Konsultasi Bab I, II, dan III	3.
4.	24 September 2024	Konsultasi Bab I, II, dan III	4.
5.	03 Oktober 2024	Konsultasi Kajian Agama Bab I dan II	5.
6.	23 Oktober 2024	Konsultasi Kajian Agama Bab I dan II	6.
7.	31 Oktober 2024	Konsultasi Kajian Agama Bab I dan II	7.
8.	04 November 2024	ACC Bab I, II, dan III	8.
9.	06 November 2024	Konsultasi Kajian Agama Bab I dan II	9.
10.	07 November 2024	ACC Kajian Agama Bab I dan II	10.
11.	11 November 2024	ACC Seminar Proposal	11.
12.	02 Desember 2024	Konsultasi Revisi Seminar Proposal	12.
13.	16 Januari 2025	Konsultasi Bab IV	13.
14.	24 Januari 2025	Konsultasi Bab IV dan V	14.
15.	07 Februari 2025	Konsultasi Bab IV dan V	15.



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

No	Tanggal	Hal	Tanda Tangan
16.	20 Maret 2025	Konsultasi Kajian Agama Bab IV	16. <i>[Signature]</i>
17.	09 April 2025	Konsultasi Kajian Agama Bab IV	17. <i>[Signature]</i>
18.	15 April 2025	Konsultasi Kajian Agama Bab IV	18. <i>[Signature]</i>
19.	23 April 2025	ACC BAB IV dan V	19. <i>[Signature]</i>
20.	23 April 2025	ACC Kajian Agama Bab IV	20. <i>[Signature]</i>
21.	24 April 2025	ACC Seminar Hasil	21. <i>[Signature]</i>
22.	27 Mei 2025	Konsultasi Revisi Seminar Hasil	22. <i>[Signature]</i>
23.	03 Juni 2025	ACC Matriks Revisi Seminar Hasil	23. <i>[Signature]</i>
24.	04 Juni 2025	ACC Sidang Skripsi	24. <i>[Signature]</i>
25.	13 Juni 2025	ACC Keseluruhan	25. <i>[Signature]</i>

Malang, 13 Juni 2025

Mengetahui,

Ketua Program Studi Matematika

[Signature]

Dr. Elly Susanti, M.Sc.

NIP. 19741129 200012 2 005