

**KLASTERISASI SISWA BIMBINGAN BELAJAR TESLA SIDOARJO  
MENGGUNAKAN ALGORITMA K-MEANS**

**SKRIPSI**

**Oleh :**  
**ADYATMA ABIDIN**  
**NIM. 210605110109**



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

**KLASTERISASI SISWA BIMBINGAN BELAJAR TESLA SIDOARJO  
MENGGUNAKAN ALGORITMA K-MEANS**

**SKRIPSI**

Diajukan kepada:  
Universitas Islam Negeri Maulana Malik Ibrahim Malang  
Untuk memenuhi Salah Satu Persyaratan dalam  
Memperoleh Gelar Sarjana Komputer (S.Kom)

**Oleh :**  
**ADYATMA ABIDIN**  
**NIM. 210605110109**

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

## HALAMAN PERSETUJUAN

### KLASTERISASI SISWA BIMBINGAN BELAJAR TESLA SIDOARJO MENGGUNAKAN ALGORITMA K-MEANS

#### SKRIPSI

Oleh :  
**ADYATMA ABIDIN**  
**NIM. 210605110109**

Telah Diperiksa dan Disetujui untuk Diuji:  
Tanggal: 20 November 2025

Pembimbing I,



Ashri Shabrina Afrah, M.T  
NIP. 19900430 202012 2 003

Pembimbing II,



Ajib Hanani, M.T  
NIP. 19840731 202321 1 013

Mengetahui,  
Ketua Program Studi Teknik Informatika  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Maulana Malik Ibrahim Malang



## HALAMAN PENGESAHAN

### KLASTERISASI SISWA BIMBINGAN BELAJAR TESLA SIDOARJO MENGGUNAKAN ALGORITMA K-MEANS

#### SKRIPSI

Oleh :  
**ADYATMA ABIDIN**  
**NIM. 210605110109**

Telah Dipertahankan di Depan Dewan Penguji Skripsi  
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan  
Untuk Memperoleh Gelar Sarjana Komputer ( S.Kom )  
Tanggal: 26 Desember 2025

#### Susunan Dewan Penguji

Ketua Penguji : Dr. Agung Teguh Wibowo Almais, M.T  
NIP. 19860301 202321 1 016

(  )

Anggota Penguji I : Fajar Rohman Hariri, M.Kom  
NIP. 19890515 201801 1 001

(  )

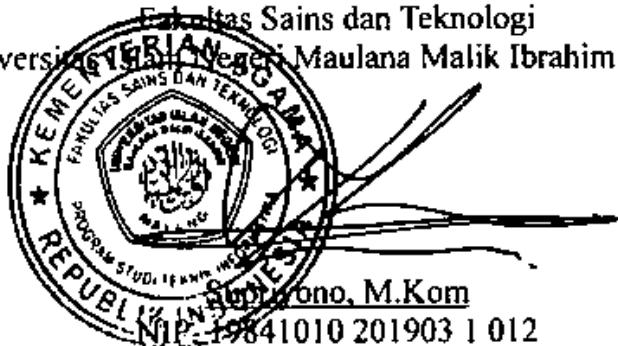
Anggota Penguji II : Ashri Shabrina Afrah, M.T  
NIP. 19900430 202012 2 003

(  )

Anggota Penguji III : Ajib Hanani, M.T  
NIP. 19840731 202321 1 013

(  )

Mengetahui dan Mengesahkan,  
Ketua Program Studi Teknik Informatika  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Maulana Malik Ibrahim Malang



## PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Adyatma Abidin  
NIM : 210605110109  
Fakultas / Program Studi : Sains dan Teknologi / Teknik Informatika  
Judul Skripsi : Klasterisasi Siswa Bimbingan Belajar Tesla  
Sidoarjo Menggunakan Algoritma K-Means

Menyatakan dengan sebenarnya bahwa Skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambil alihan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini merupakan hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 26 Desember 2025  
Yang membuat pernyataan,



Adyatma Abidin  
NIM.210605110109

## **MOTTO**

Leave no regret

## **HALAMAN PERSEMBAHAN**

Skripsi ini saya persembahkan kepada orang tua, adik, dosen, teman – teman serta seluruh pihak yang telah menemani dan membantu dalam proses penyelesaian penelitian ini baik secara mental maupun finansial. Semoga mereka semua senantiasa berada dalam lindungan Allah SWT.

## KATA PENGANTAR

Puji syukur kehadirat Allah SWT atas segala rahmat dan karunia yang telah ia limpahkan. Shalawat serta salam selalu tercurahkan kepada Nabi Muhammad Shallallah ‘alaihi wasallam. Atas limpahan berkat dan rahmat-Nya, penulis dapat menuntaskan penelitian skripsi dengan judul “Klasterisasi Siswa Bimbingan Belajar Tesla Sidoarjo Menggunakan Algoritma K-Means”, sebagai salah satu persyaratan untuk menyelesaikan pendidikan sarjana pada Program Studi Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang.

Dengan penuh rasa Syukur, penulis ingin mengucapkan terima kasih yang tak terhingga kepada seluruh pihak yang terlibat pada proses penyusunan skripsi ini. Ucapan terima kasih penulis sampaikan kepada:

1. Prof. Dr. Hj. Ilfi Nur Diana, M.Si., selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Dr. Agus Mulyono, M.Kes., selaku dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Bapak Supriyono, M.Kom., selaku Ketua Program Studi Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Ibu Ashri Shabrina Afrah, M.T., selaku dosen pembimbing I yang telah memberikan arahan, saran, serta nasihat kepada penulis selama proses penggerjaan skripsi.

5. Bapak Ajib Hanani, M.T., selaku dosen pembimbing II yang telah memberikan arahan, saran, serta nasihat kepada penulis selama proses penggerjaan skripsi.
6. Bapak Dr. Agung Teguh Wibowo Almais, M.T., selaku ketua penguji serta Bapak Fajar Rohman Hariri, M.Kom., selaku anggota penguji yang telah memberikan kritik serta saran kepada penulis sehingga dapat menuntaskan penelitian ini.
7. Bimbingan Belajar Tesla Sidoarjo, yang telah berkenan untuk membantu penulis dalam proses penggerjaan skripsi.
8. Teman teman yang telah menemani dan memberikan saran serta dukungan dalam proses penggerjaan skripsi.

Akhir kata, penulis menyadari masih terdapat banyak kekurangan pada penelitian ini. Maka, penulis berharap adanya kritik serta saran demi perbaikan di masa yang akan datang.

Malang, 13 November 2025

Penulis

## DAFTAR ISI

<b>HALAMAN PERSETUJUAN.....</b>	<b>iii</b>
<b>HALAMAN PENGESAHAN.....</b>	<b>iv</b>
<b>PERNYATAAN KEASLIAN TULISAN .....</b>	<b>v</b>
<b>MOTTO .....</b>	<b>vi</b>
<b>HALAMAN PERSEMBAHAN.....</b>	<b>vii</b>
<b>KATA PENGANTAR.....</b>	<b>viii</b>
<b>DAFTAR ISI.....</b>	<b>x</b>
<b>DAFTAR GAMBAR .....</b>	<b>xii</b>
<b>DAFTAR TABEL.....</b>	<b>xiv</b>
<b>ABSTRAK .....</b>	<b>xv</b>
<b>ABSTRACT .....</b>	<b>xvi</b>
<b>مستخلص البحث.....</b>	<b>xvii</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1    Latar Belakang .....	1
1.2    Rumusan Masalah .....	6
1.3    Batasan Masalah.....	6
1.4    Tujuan Penelitian.....	6
1.5    Manfaat Penelitian .....	7
<b>BAB II STUDI PUSTAKA .....</b>	<b>8</b>
2.1    Penelitian Terkait.....	8
2.2    Bimbingan Belajar .....	14
2.3    Gaya Belajar.....	15
2.4 <i>Index of Learning Styles</i> .....	17
2.5 <i>Clustering</i> .....	19
2.6    Algoritma K-Means <i>Clustering</i> .....	20
2.7 <i>Sihouette Score</i> .....	22
2.8 <i>Elbow Method</i> .....	23
<b>BAB III METODOLOGI PENELITIAN .....</b>	<b>25</b>
3.1    Desain Penelitian.....	25
3.1.1    Desain Sistem.....	26

3.1.2	Sumber Data.....	27
3.1.3	Praproses Data.....	28
3.1.4	<i>K-Means Clustering</i> .....	29
3.1.5	Skenario Uji Coba.....	30
3.1.6	Evaluasi <i>Silhouette Score</i> .....	36
<b>BAB IV HASIL DAN PEMBAHASAN.....</b>		<b>40</b>
4.1	Hasil Penelitian .....	40
4.1.1	Data Penelitian .....	40
4.1.2	<i>K-Means Clustering</i> .....	43
4.1.3	Skenario Uji Coba.....	47
4.1.4	Evaluasi <i>Silhouette Score</i> .....	79
4.1.5	Evaluasi <i>Elbow Method</i> .....	81
4.2	Pembahasan.....	82
4.3	Integrasi Islam.....	88
<b>BAB V KESIMPULAN DAN SARAN .....</b>		<b>90</b>
5.1.	Kesimpulan .....	90
5.2.	Saran.....	91

## **DAFTAR PUSTAKA**

## DAFTAR GAMBAR

Gambar 3.1. Desain Penelitian.....	25
Gambar 3.2. Desain Sistem.....	27
Gambar 3.3. <i>Flowchart K-Means Clustering</i> .....	30
Gambar 4.1. Tampilan Awal Sistem .....	44
Gambar 4.2. Tampilan Hasil <i>Clustering</i> .....	45
Gambar 4.3. Analisa <i>Cluster</i> .....	46
Gambar 4.4. Visualisasi <i>Radar Chart</i> Hasil <i>Clustering</i> .....	47
Gambar 4.5. Hasil <i>Clustering</i> Skenario 1 .....	48
Gambar 4.6. Analisa <i>Cluster</i> Skenario 1 .....	49
Gambar 4.7. Visualisasi Skenario 1 .....	50
Gambar 4.8. Hasil <i>Clustering</i> Skenario 2 .....	51
Gambar 4.9. Analisa <i>Cluster</i> Skenario 2 .....	52
Gambar 4.10. Visualisasi Skenario 2 .....	53
Gambar 4.11. Hasil <i>Clustering</i> Skenario 3 .....	54
Gambar 4.12. Analisa <i>Cluster</i> Skenario 3 .....	55
Gambar 4.13. Visualisasi Skenario 3 .....	56
Gambar 4.14. Hasil <i>Clustering</i> Skenario 4 .....	57
Gambar 4.15. Analisa <i>Cluster</i> Skenario 4 .....	58
Gambar 4.16. Visualisasi Skenario 4 .....	59
Gambar 4.17. Hasil <i>Clustering</i> Skenario 5 .....	60
Gambar 4.18. Analisa <i>Cluster</i> Skenario 5.....	61
Gambar 4.19. Visualisasi Skenario 5 .....	62
Gambar 4.20. Hasil <i>Clustering</i> Skenario 6 .....	64
Gambar 4.21. Analisa <i>Cluster</i> Skenario 6.....	65
Gambar 4.22. Visualisasi Skenario 6 .....	66
Gambar 4.23. Hasil <i>Clustering</i> Skenario 7 .....	67
Gambar 4.24. Analisa <i>Cluster</i> Skenario 7 .....	68
Gambar 4.25. Visualisasi Skenario 7 .....	69
Gambar 4.26. Hasil <i>Clustering</i> Skenario 8 .....	70
Gambar 4.27. Analisa <i>Cluster</i> Skenario 8.....	71
Gambar 4.28. Visualisasi Skenario 8 .....	72
Gambar 4.29. Hasil <i>Clustering</i> Skenario 9 .....	73

Gambar 4.30. Analisa <i>Cluster</i> Skenario 9.....	74
Gambar 4.31. Visualisasi Skenario 9 .....	75
Gambar 4.32. Hasil <i>Clustering</i> Skenario 10 .....	76
Gambar 4.33. Analisa <i>Cluster</i> Skenario 10.....	77
Gambar 4.34. Visualisasi Skenario 10 .....	78
Gambar 4.35. Grafik <i>Silhouette Score</i> .....	80
Gambar 4.36. Visualisasi Grafik SSE .....	82
Gambar 4.37. Visualisasi Seluruh Skenario .....	83

## DAFTAR TABEL

Tabel 2.1. Tabel Penelitian Terkait .....	12
Tabel 2.2. Interpretasi Nilai <i>Silhouette Score</i> .....	23
Tabel 3.1. Tabel Data Skor ILS Siswa .....	28
Tabel 3.2. Tabel Data Normalisasi .....	29
Tabel 3.3. Skenario Uji Coba <i>Cluster</i> .....	31
Tabel 3.4. Tabel <i>Centroid</i> pada Skenario 1 .....	31
Tabel 3.5. Jarak Titik Data dengan <i>Centroid</i> Iterasi 1 pada Skenario 1 .....	33
Tabel 3.6. Persebaran Data Iterasi 1 pada Skenario 1 .....	33
Tabel 3.7. Jarak Titik Data dengan <i>Centroid</i> Baru Iterasi 2 pada Skenario 1 .....	34
Tabel 3.8. Hasil Klasterisasi Siswa pada Skenario 1 .....	35
Tabel 3.9. Hasil Klasterisasi Siswa pada Skenario 2 .....	35
Tabel 3.10. Hasil Klasterisasi Siswa pada Skenario 3 .....	36
Tabel 3.11. Hasil Perhitungan <i>Silhouette Score</i> pada Skenario 1 .....	38
Tabel 3.12. Skenario Uji Coba .....	39
Tabel 4.1. Tabel Data .....	41
Tabel 4.3. Implementasi <i>Source Code</i> Normalisasi .....	41
Tabel 4.4. Implementasi <i>Source Code</i> Normalisasi .....	42
Tabel 4.5. Implementasi <i>Source Code</i> Normalisasi pada Data .....	42
Tabel 4.6. Tabel Normalisasi Data .....	43
Tabel 4.7. Implementasi <i>Source Code Silhouette Score</i> .....	79
Tabel 4.8. Hasil <i>Silhouette Score</i> pada Skenario Percobaan .....	80
Tabel 4.9. Nilai SSE pada Skenario Percobaan .....	81
Tabel 4.10. Jumlah Iterasi pada Skenario .....	86

## ABSTRAK

Abidin, Adyatma. 2025. **Klasterisasi Siswa Bimbingan Belajar Tesla Sidoarjo Menggunakan Algoritma K-Means**. Skripsi. Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Ashri Shabrina Afrah, M.T (II) Ajib Hanani, M.T.

Kata kunci: K-Means Clustering, Gaya Belajar, *Index of Learning Styles*(ILS).

Pendidikan merupakan salah satu aspek penting yang terus berkembang seiring dengan perkembangan teknologi. Salah satu media yang kerap digunakan untuk mendapatkan akses pendidikan adalah bimbingan belajar. Masing – masing siswa memiliki perbedaan kemampuan serta gaya belajar yang diminati. Hal ini tentu berpengaruh pada hasil akhir berupa perbedaan pemahaman siswa pada materi yang dijelaskan. Fenomena ini dapat diatasi dengan mengelompokkan siswa menjadi beberapa kelompok. Penelitian ini bertujuan untuk mengelompokkan siswa yang memiliki ciri serta karakteristik yang serupa menggunakan Algoritma K-Means Clustering berdasarkan hasil kuesioner *Index of Learning Styles*(ILS). Proses pengelompokan akan dilakukan sesuai dengan skenario yang telah ditentukan. Hasil dari skenario tersebut kemudian akan dievaluasi menggunakan *Silhouette Score* untuk menentukan skenario pengelompokan yang paling optimal. Pada penelitian ini, skenario 3 keluar sebagai skenario terbaik dengan nilai *Silhouette Score* tertinggi diantara skenario yang lain dengan nilai 0.24523330344105687. Hasil yang diperoleh dapat dimanfaatkan oleh para tenaga pendidik untuk menyesuaikan strategi serta model pembelajaran yang kiranya cocok untuk digunakan pada kelompok siswa dengan gaya belajar tertentu.

## ABSTRACT

Abidin, Adyatma. 2025. **Clustering of Tesla Sidoarjo Tutoring Students Using the K-Means Algorithm.** Thesis. Departement of Informatics Engineering, Faculty on Science and Technology, State Islamic University Maulana Malik Ibrahim Malang. Advisor: (I) Ashri Shabrina Afrah, M.T (II) Ajib Hanani, M.T.

Keywords: K-Means Clustering, Learning Styles, *Index of Learning Styles*(ILS).

Education is an important aspect that continues to evolve alongside technological developments. One medium that is often used to gain access to education is tutoring. Each student has different abilities and preferred learning styles. This certainly affects the final results in the form of differences in students' understanding of the material being explained. This phenomenon can be overcome by grouping students into several groups. This study aims to group students who have similar traits and characteristics using the K-Means Clustering Algorithm based on the results of the Index of Learning Styles(ILS) questionnaire. The grouping process will be carried out according to a predetermined scenario. The results of the scenario will then be evaluated using the Silhouette Score to determine the most optimal grouping scenario. In this study, scenario 3 emerged as the best scenario with the highest Silhouette Score value among the other scenarios with a value of 0.24523330344105687. The results obtained can be used by educators to adjust strategies and learning models that are suitable for use in groups of students with specific learning styles.

## مستخلص البحث

أبيدين، أدياتا. 2025. تقسيم طلاب دروس تسلا سيدوارجو باستخدام خوارزمية K-Means. أطروحة. قسم هندسة المعلوماتية، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية الحكومية، مالانج. المشرفون (I): أشري شابربينا أفراح، M.T (II): أجيبي حناني، M.T

الكلمات المفتاحية: تجميع K-Means ، أنماط التعلم، مؤشر أنماط التعلم. (ILS)

التعليم هو جانب مهم يستمر في التطور جنباً إلى جنب مع التطورات التكنولوجية. أحد الوسائل التي غالباً ما تستخدم للوصول إلى التعليم هو التدريس الخصوصي. لكل طالب قدرات مختلفة وأساليب تعلم مختلفة. وهذا يؤثر بالتأكيد على النتائج النهائية في شكل اختلافات في فهم الطلاب للمواد التي يتم شرحها. يمكن التغلب على هذه الظاهرة من خلال تجميع الطلاب في عدة مجموعات. تهدف هذه الدراسة إلى تجميع الطلاب الذين لديهم خصائص مشابهة باستخدام خوارزمية التجميع K-Means استناداً إلى نتائج استبيان مؤشر أنماط التعلم. (ILS) سيتم تنفيذ عملية التجميع وفقاً لسيناريو محدد مسبقاً. سيتم بعد ذلك تقييم نتائج السيناريو باستخدام معامل الصورة الظلية لتحديد السيناريو الأمثل للتجميع. في هذه الدراسة، ظهر السيناريو 3 كأفضل سيناريو مع أعلى قيمة لمعامل الصورة الظلية بين السيناريوهات الأخرى. يمكن للمعلمين استخدام النتائج التي تم الحصول عليها لتعديل الاستراتيجيات وفما ذر التعلم المناسب للاستخدام في مجموعات الطلاب ذات أنماط التعلم المحددة.

## **BAB I**

### **PENDAHULUAN**

#### **1.1 Latar Belakang**

Pendidikan merupakan salah satu sektor penting yang terus berkembang seiring dengan berjalannya waktu. Salah satu media yang biasa digunakan untuk memperoleh pendidikan adalah sekolah. Dengan sekolah, siswa bisa mengakses pendidikan dengan baik, di mana terdapat tenaga pendidik atau guru yang akan mengajarkan serta menjelaskan materi pelajaran untuk meningkatkan pemahaman para siswa dengan berbagai macam inovasi yang dikembangkan (Najm Al Inu et al., 2023).

Media lain yang juga kerap digunakan untuk mengakses pendidikan adalah bimbingan belajar. Bimbingan belajar adalah salah satu upaya yang kerap dilakukan siswa untuk membantu menyelesaikan masalah belajar yang mereka alami. Pada sebuah sesi bimbingan belajar, siswa akan ditemani oleh seorang pendamping yang akan membantu siswa menyelesaikan masalah belajar mereka. Bimbingan belajar biasanya dilakukan diluar jam pelajaran di sekolah. Para siswa tentunya tidak belajar sendiri, mereka akan ditempatkan di sebuah ruang kelas seperti halnya belajar di sekolah. Biasanya, bimbingan belajar akan menempatkan siswa ke ruang kelas dengan jumlah siswa yang cenderung sedikit. Hal ini dilakukan supaya pendamping dan siswa mampu menciptakan suasana belajar yang lebih kondusif (Wahyuningsih et al., 2020).

Dalam proses belajar mengajar ini, guru harus memiliki kemampuan untuk memilih metode pembelajaran atau strategi pembelajaran yang sesuai untuk siswa. Hal ini diperlukan karena tingkat pemahaman antara satu siswa dengan siswa yang lain tidaklah sama. Dalam materi tertentu, beberapa siswa bisa jadi kurang memahami apa yang dijelaskan oleh guru di kelas (Azzahrah Putri et al., 2021). Kejadian ini bukanlah sesuatu yang aneh, biasanya siswa akan bertanya tentang materi yang tidak mereka pahami, dan tentunya guru akan dengan senang hati menjawab pertanyaan tersebut. Namun, jika guru terlalu fokus untuk menjelaskan materi kepada salah satu siswa, siswa dengan tingkat pemahaman yang berbeda terpaksa harus mengikuti tempo belajar siswa tersebut.

Fenomena ini bisa terjadi karena banyak faktor, salah satunya adalah gaya belajar siswa yang berbeda. Menurut Richard Felder, tiap siswa memiliki gaya belajar serta pendekatan yang berbeda satu sama lain. Perbedaan ini berpengaruh kuat pada cara mereka memproses informasi yang diperoleh. Perbedaan ini juga berdampak pada hasil yang didapatkan oleh siswa tersebut, beberapa siswa mungkin saja mendapatkan hasil akhir yang berbeda dengan siswa lainnya (Felder & Brent, 2005).

Masalah ini dapat diselesaikan dengan cara membagi siswa menjadi beberapa kelompok. Sebuah kelompok akan terdiri dari siswa yang memiliki gaya belajar yang serupa. Dengan adanya kelompok yang berisi siswa dengan gaya belajar yang serupa, guru bisa mengambil langkah yang tepat untuk membantu proses pemahaman materi pada siswa yang diajar.

Hal ini juga sesuai dengan firman Allah SWT pada Al-Qur'an surat *Az-Zumar* ayat 9 yang berbunyi,

أَمْنٌ هُوَ فَيَنْتَهِ إِلَيْهِ الْمُنْتَهَىٰ وَقَائِمًا يَخْذُلُ الْآخِرَةَ وَيَرْجُوا رَحْمَةَ رَبِّهِ ۖ قُلْ هُنَّ يَسْتَوِي الْأَذْيَانُ يَعْلَمُونَ وَالْأَذْيَانُ لَا يَعْلَمُونَ ۖ إِنَّمَا يَتَذَكَّرُ أُولُو الْأَلْبَابِ ۙ

*"(Apakah orang musyrik yang lebih beruntung) ataukah orang yang beribadah pada waktu malam dalam keadaan bersujud, berdiri, takut pada (azab) akhirat, dan mengharapkan rahmat Tuhaninya? Katakanlah, "Apakah sama orang yang mengetahui dengan yang tidak mengetahui?" Sesungguhnya hanya orang yang berakal yang dapat menerima pelajaran". (Q.S. Az Zumar:9).*

Ayat tersebut menjelaskan adanya perbedaan di antara beberapa kaum. Menurut Ibnu Katsir, terdapat perbedaan antara orang – orang musyrik dengan orang yang beribadah di waktu malam, bersujud dan tunduk kepada Allah SWT. Orang – orang demikian tidak mempunyai derajat yang sama di mata Allah Swt. Hanya orang – orang yang mempunyai akal yang dapat mengetahui perbedaan antara kedua golongan tersebut (Ghaffar & al-Atsari, 2004).

Begitu halnya pada diri seorang siswa, setiap dari diri mereka memiliki perbedaan antara satu sama lain. Setiap siswa memiliki perbedaan karakter, kemampuan diri, akademis dan berbagai faktor lainnya. Karena adanya perbedaan pada siswa inilah pengelompokan siswa dilakukan. Dengan membagi siswa yang memiliki karakteristik serupa ke dalam beberapa kelompok yang berbeda, diharapkan proses pembelajaran akan berjalan secara lebih baik serta efektif.

Salah satu metode yang seringkali digunakan untuk melakukan pembagian kelompok adalah *clustering*. *Clustering* merupakan sebuah teknik *data mining* yang kerap digunakan untuk mengelompokkan atau melakukan segmentasi terhadap sebuah data yang memiliki pola serta karakteristik yang serupa (Ayu Syahfitri et al.,

2024). Dengan menggunakan *clustering*, kita bisa membagi siswa yang ada menjadi beberapa kelompok klaster yang terdiri dari siswa dengan gaya belajar yang serupa.

Pada penelitian yang dilakukan oleh Yotaman pada tahun 2020, peneliti mengelompokkan siswa untuk menyesuaikan strategi pengajaran yang efektif. Peneliti menggunakan kuesioner *Index of Learning Styles* (ILS) pada 48 siswa yang diolah menggunakan *Hierarchical Clustering*. Peneliti juga menggunakan menggunakan *Euclidean Distance* serta *Manhattan Distance*, untuk menentukan jumlah klaster yang sesuai. Dari penelitian ini, siswa terbagi menjadi 7 klaster yang memiliki gaya belajar yang serupa (Yotaman et al., 2020).

Pada penelitian yang dilakukan oleh Maryono pada tahun 2022, peneliti mengelompokkan siswa berdasarkan gaya belajar yang serupa. Peneliti menggunakan kuesioner *Index of Learning Styles* (ILS) pada 58 responden yang diolah menggunakan *k-means clustering* untuk menentukan jumlah klaster yang sesuai. Dari evaluasi tersebut, responden terbagi menjadi 2 kelompok yang menunjukkan karakteristik serta kecenderungan yang serupa (Maryono et al., 2022).

Pada penelitian yang dilakukan oleh Chi pada tahun 2022, peneliti mengelompokkan siswa berdasarkan kemampuan belajar untuk mengevaluasi kinerja siswa. Peneliti menggunakan nilai ujian akhir siswa pada 7 mata Pelajaran yang diolah menggunakan *k-means clustering*. Dari penelitian tersebut Jumlah k yang paling optimal jatuh pada nilai  $k = 5$  yang memiliki *squared error* terkecil dengan nilai 2.44 dengan  $C1 = 30$  siswa,  $C2 = 30$  siswa,  $C3 = 60$  siswa,  $C4 = 10$  siswa, dan  $C5 = 30$  siswa (Chi, 2021).

Dari beberapa penelitian yang telah dilakukan sebelumnya, Algoritma K-*Means Clustering* serta Algoritma *Hierarchical Clustering* merupakan metode yang paling sering digunakan untuk mengolah data yang digunakan. Perbedaan yang ada diantara keduanya terletak pada pendekatan yang digunakan, dimana *hierarchical clustering* akan membagi data dalam bentuk *dendogram*, sedangkan *k-means clustering* akan membagi data dalam bentuk partisi. Dari hasil yang didapatkan, penelitian dengan Algoritma K-*Means Clustering* terbukti berhasil mengelompokkan seluruh data pada sebuah klaster (Maryono et al., 2022). Sedangkan Algoritma *Hierarchical Clustering* hanya dapat mengelompokkan sebagian data saja (Pamungkas et al., 2021). Data yang tidak memiliki karakteristik yang serupa dengan klaster yang telah terbentuk keluar sebagai *outlier*.

Algoritma K-*Means Clustering* juga merupakan algoritma yang lebih sederhana serta memiliki tingkat efisiensi yang tinggi. Meskipun digunakan pada data dengan jumlah sampel yang kecil, algoritma ini dapat mencapai akurasi yang tinggi serta dapat menurunkan tingkat kompleksitas waktu yang digunakan (Chi, 2021). Dari hasil tersebut, penelitian ini akan menggunakan Algoritma K-*Means Clustering* untuk mengolah serta mengelompokkan siswa berdasarkan gaya belajar yang serupa.

Pada penelitian kali ini, parameter yang digunakan berupa hasil kuesioner *Index of Learning Styles* (ILS). Penelitian ini dilakukan dengan tujuan untuk membantu tenaga pendidik untuk menentukan strategi pembelajaran yang sesuai.

## 1.2 Rumusan Masalah

Pernyataan masalah pada penelitian ini adalah bagaimana *K-Means Clustering* dapat digunakan untuk mengelompokkan siswa sesuai dengan gaya belajar berdasarkan hasil kuesioner *Index of Learning Styles* (ILS) yang dibagikan?

## 1.3 Batasan Masalah

Untuk membatasi pembahasan pada penelitian ini, peneliti menetapkan beberapa batasan, antara lain,

1. Sumber data yang digunakan pada penelitian ini berasal dari siswa Bimbingan Belajar Tesla Sidoarjo pada Tahun Ajaran 2024/2025 sebanyak 58 siswa SMA.
2. Parameter yang digunakan untuk mengelompokkan siswa meliputi hasil kuesioner *Index of Learning Styles* (ILS).
3. Penelitian ini akan menggunakan Algoritma *K-Means Clustering* untuk mengelompokkan siswa sesuai dengan gaya belajar yang serupa.
4. Penelitian ini akan menggunakan *Silhouette Score* untuk menguji kualitas klaster yang keluar sebagai hasil *clustering*.

## 1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk mengelompokkan siswa sesuai dengan gaya belajar yang serupa menggunakan Algoritma *K-Means Clustering* berdasarkan hasil kuesioner ILS supaya dapat membantu tenaga pendidik atau guru untuk melakukan analisa serta mengambil langkah yang sesuai pada proses pembelajaran.

## 1.5 Manfaat Penelitian

Adapun manfaat dari penelitian ini antara lain,

1. Penelitian ini diharapkan dapat menambah khasanah keilmuan mengenai *clustering* menggunakan Algoritma *K-Means Clustering* untuk mengelompokkan siswa berdasarkan gaya belajar yang serupa.
2. Hasil dari penelitian diharapkan dapat digunakan untuk membantu guru atau pengajar untuk melakukan analisa dan menentukan sikap apa yang harus diambil pada siswa dengan gaya belajar tertentu.

## **BAB II**

### **STUDI PUSTAKA**

Bagian ini menjelaskan mengenai penelitian yang telah dilakukan dengan pembahasan seputar *clustering* siswa berdasarkan gaya belajar siswa. Bagian ini juga menjelaskan mengenai landasan teori yang mendukung atribut – atribut yang akan digunakan pada penelitian ini.

#### **2.1 Penelitian Terkait**

Penelitian pertama dilakukan oleh Dyulicheva et al., pada tahun 2020. Penelitian ini membahas mengenai pengelompokan siswa, di mana siswa akan mengisi kuesioner honey dan mumford yang dirancang oleh Peter Honey dan Alan Mumford serta kuesioner *Index of Learning Styles* yang dirancang oleh Felder dan Soloman. Peneliti menggunakan algoritma k-means serta *hierarchical clustering* untuk mengelompokkan data siswa. Dari penelitian yang dilakukan, terbentuk beberapa kelompok yang terdiri dari siswa dengan preferensi serta orientasi belajar yang serupa. Namun, penelitian ini hanya berhasil mengelompokkan 35 siswa dari total 61 siswa yang mengisi kuesioner (Dyulicheva & Kosova, 2020).

Penelitian selanjutnya dilakukan oleh Yotaman et al., pada tahun 2020. Peneliti menggunakan Algoritma *Hierarchical Clustering* untuk mengelompokkan 48 siswa responden kuesioner. Peneliti juga menggunakan *Linkage Criterion* yang diukur menggunakan *Euclidean Distance* serta *Manhattan Distance*, untuk menentukan jumlah klaster yang sesuai dengan nilai *error* (SSE) terkecil. Dari penelitian ini, siswa terbagi menjadi 7 klaster berdasarkan gaya belajar siswa

dengan *error* terkecil yaitu 569.46 pada *Euclidean Distance* yang dikombinasikan dengan *Ward Linkage Criterion* (Yotaman et al., 2020).

Penelitian selanjutnya dilakukan oleh Maryono et al., pada tahun 2022. Penelitian ini mengukur gaya belajar 58 siswa menggunakan kuesioner *Index of Learning Styles* yang disusun oleh Felder dan Soloman untuk menilai preferensi belajar siswa. Peneliti menggunakan algoritma *k-means* untuk pengelompokan siswa dengan jumlah *k* bernilai 2, 3, 4, dan 8. Untuk menentukan jumlah *k* yang optimal, peneliti menggunakan evaluasi *Sum of Square Error*. Dari evaluasi tersebut, jumlah *k* dengan nilai 2 keluar sebagai hasil pengelompokan terbaik. Dimana klaster 1 terdiri dari 26 siswa dengan dimensi *active* pada gaya belajar mereka, siswa dengan dimensi ini cenderung belajar dengan cara eksperimen serta melakukan percobaan. Klaster 2 terdiri dari 32 siswa dengan dimensi *reflective* pada gaya belajar mereka, siswa dengan dimensi ini cenderung belajar dengan cara memikirkan secara matang tentang apa yang dipelajari (Maryono et al., 2022).

Penelitian selanjutnya dilakukan oleh Chi et al., pada tahun 2021. Peneliti menggunakan Algoritma *K-Means* untuk membagi siswa menjadi beberapa klaster berdasarkan kinerja atau nilai akhir siswa pada beberapa mata pelajaran. Kemudian peneliti mengelompokkan siswa dengan jumlah *k* dengan nilai diantara 2 sampai dengan 6. Peneliti menggunakan evaluasi *squared error* untuk menentukan nilai *k* paling optimal. Jumlah *k* yang paling optimal jatuh pada nilai *k* = 5 yang memiliki *squared error* terkecil dengan nilai 2.44 dengan  $C1 = 30$  siswa,  $C2 = 30$  siswa,  $C3 = 60$  siswa,  $C4 = 10$  siswa, dan  $C5 = 30$  siswa (Chi, 2021).

Penelitian selanjutnya dilakukan oleh Syahputra et al., pada tahun 2022. Penelitian ini menggunakan nilai 80 siswa yang telah melewati proses normalisasi yang kemudian akan dikelompokkan menjadi 3 *cluster* “High”, “Medium”, dan “Low” menggunakan algoritma *k-means*. Pengelompokan ini digunakan sebagai sistem rekomendasi perankingan kelas. Dari hasil pengelompokan, didapatkan 3 buah *cluster* dengan C1 = 43 siswa yang termasuk dalam kelas superior, C2 = 18 siswa, dan C3 = 19 siswa di mana *cluster* 2 dan 3 masuk kedalam kelas reguler (Syahputra & Hutagalung, 2022).

Penelitian selanjutnya dilakukan oleh Pamungkas et al., pada tahun 2021. Penelitian ini menggunakan algoritma *hierarchical clustering* untuk mengelompokkan siswa yang telah mengisi kuesioner *Index of Learning Styles* yang disusun oleh Felder dan Soloman. Setelah pra pemrosesan data, sebuah dendogram akan digambar pada data yang memiliki karakteristik yang serupa menjadi sebuah kelompok. Dari dendogram tersebut, 2 buah klaster keluar sebagai hasil pengelompokan, di mana C1 terdiri dari 29 data yang memiliki preferensi *reflective* dalam proses pembelajaran dengan rata rata -3.76, serta C2 yang terdiri dari 25 data yang memiliki preferensi *active* dalam proses pembelajaran dengan rata rata 2.52. Selain kedua klaster tersebut, terdapat sebuah *outlier* berisi 4 data yang tidak memiliki karakteristik yang serupa dengan klaster yang telah terbentuk (Pamungkas et al., 2021).

Penelitian selanjutnya dilakukan oleh Kartina dan Safii pada tahun 2021. Penelitian ini membahas mengenai pemantauan kualitas pendidikan yang terjadi pada jenjang SMA. Pemantauan ini dilakukan sebagai upaya dari pengendalian

serta jaminan mutu pendidikan. Peneliti menggunakan Algoritma *K-Means Clustering* untuk mengelompokkan 28 sekolah berdasarkan nilai ujian akhir sekolah. Dari pengolahan tersebut, sekolah terbagi menjadi 2 klaster, dimana klaster pertama terdiri dari 21 sekolah dengan nilai yang lebih tinggi sedangkan klaster kedua terdiri dari 7 sekolah dengan nilai yang lebih rendah (Kartina & Safii, 2021).

Penelitian selanjutnya dilakukan oleh Asroni et al., pada tahun 2020. Peneliti menggunakan Algoritma *K-Means Clustering* untuk mengelompokkan 395 mahasiswa berdasarkan nilai pada beberapa mata kuliah serta indeks prestasi. Pengelompokan ini bertujuan untuk membagi mahasiswa berdasarkan nilai mahasiswa untuk membantu pengembangan kualitas pendidikan yang terjadi. Dari penelitian ini, mahasiswa terbagi menjadi 3 klaster. Klaster 1 terdiri dari 72 mahasiswa dengan nilai rendah dan indeks prestasi rendah dengan rata – rata 2.72. Klaster 2 terdiri dari 190 mahasiswa dengan nilai tinggi dan indeks prestasi tinggi dengan rata – rata 3.60. Klaster 3 terdiri dari 133 mahasiswa dengan nilai tinggi dan indeks prestasi yang cenderung tinggi, namun lebih rendah dibandingkan klaster 2 dengan rata – rata 3.13. (Asroni et al., 2020).

Penelitian saat ini akan menggunakan skor kuesioner *Index of Learning Styles* (ILS) dari siswa pada Bimbingan Belajar Tesla Sidoarjo. Data tersebut akan diolah menggunakan Algoritma *K-Means Clustering*, untuk mengelompokkan siswa berdasarkan gaya belajar yang serupa. Penelitian ini akan menggunakan 10 skenario, dengan perbedaan jumlah  $k$  atau klaster pada masing – masing skenario dengan nilai  $k$  diantara 2 sampai dengan 11. Penelitian ini akan menggunakan

*Silhouette Score* untuk menentukan skenario terbaik yang memiliki tingkat klasterisasi lebih baik dengan skor *silhouette* yang lebih tinggi.

Tabel 2.1. Tabel Penelitian Terkait

No	Penulis	Sumber Data dan Metode	Hasil
1	Yu.Yu. Dyulicheva, Y. Kosova (2020)	Skor Kuesioner Honey Mumford, Skor Kuesioner <i>Index of Learning Styles</i> (ILS), Ward.D2, <i>K-Means Clustering</i>	Dari penelitian yang dilakukan, terbentuk beberapa kelompok yang terdiri dari siswa dengan preferensi serta orientasi belajar yang serupa. Namun, penelitian ini hanya berhasil mengelompokkan 35 siswa dari total 61 siswa yang mengisi kuesioner.
2	Yotaman et al. (2020)	Skor Kuesioner <i>Index of Learning Styles</i> (ILS), <i>Hierarchical Clustering</i> , <i>Euclidean Distance</i> , <i>Manhattan Distance</i> , <i>Sum of Square Error</i> (SSE)	Dari penelitian ini, siswa terbagi menjadi 7 <i>cluster</i> yang memiliki gaya belajar yang serupa.
3	Maryono et al. (2022)	Skor Kuesioner <i>Index of Learning Styles</i> (ILS), <i>K-Means Clustering</i> , <i>Sum of Square Error</i> (SSE)	Peneliti menggunakan algoritma k-means untuk pengelompokan siswa dengan jumlah k bernilai 2, 3, 4, dan 8. Dari evaluasi SSE, jumlah k dengan nilai 2 keluar sebagai hasil pengelompokan terbaik.
4	Chi et al. (2021)	Nilai Akhir Siswa, <i>K-Means Clustering</i> , <i>Sum of Squared Error</i> (SSE)	Peneliti menggunakan evaluasi squared error untuk menentukan nilai k paling optimal. Jumlah k yang paling optimal jatuh pada nilai k = 5 yang memiliki squared error terkecil dengan nilai 2.44 dengan C1 = 30 siswa, C2 = 30 siswa, C3 =

			60 siswa, C4 = 10 siswa, dan C5 = 30 siswa.
5	Syahputra et al. (2022)	Nilai Siswa, <i>K-Means Clustering</i>	Dari hasil pengelompokan, didapatkan 3 buah <i>cluster</i> dengan C1 = 43 siswa yang termasuk dalam kelas superior, C2 = 18 siswa, dan C3 = 19 siswa dimana <i>cluster</i> 2 dan 3 masuk kedalam kelas reguler.
6	Pamungkas et al. (2021)	Skor Kuesioner <i>Index of Learning Styles</i> (ILS), <i>Hierarchical Clustering</i>	Setelah pra pemrosesan data, sebuah dendogram akan digambar pada data yang memiliki karakteristik yang serupa menjadi sebuah kelompok. Dari dendogram tersebut, 2 buah <i>cluster</i> keluar sebagai hasil pengelompokan, dimana C1 terdiri dari 29 data yang memiliki preferensi reflektif dalam proses pembelajaran dengan rata rata -3.76, serta C2 yang terdiri dari 25 data yang memiliki preferensi aktif dalam proses pembelajaran dengan rata rata 2.52. Selain kedua <i>cluster</i> tersebut, terdapat sebuah <i>outlier</i> berisi 4 data yang tidak memiliki karakteristik yang serupa dengan <i>clutser</i> yang telah terbentuk.
7	Kartina & Safii (2021)	Nilai Ujian Akhir Sekolah, <i>K-Means Clustering</i>	Peneliti mengelompokkan 28 sekolah berdasarkan nilai ujian akhir sekolah. Dari pengolahan tersebut, sekolah terbagi menjadi 2 klaster, dimana klaster pertama terdiri dari 21 sekolah dengan nilai yang lebih tinggi sedangkan klaster kedua terdiri dari 7 sekolah dengan nilai yang lebih rendah

8	Asroni et al. (2020)	Nilai Mata Kuliah, Indeks Prestasi, <i>K-Means Clustering</i>	Dari penelitian ini, mahasiswa terbagi menjadi 3 klaster. Klaster 1 terdiri dari 72 mahasiswa dengan nilai rendah dan indeks prestasi rendah dengan rata – rata 2.72. Klaster 2 terdiri dari 190 mahasiswa dengan nilai tinggi dan indeks prestasi tinggi dengan rata – rata 3.60. Klaster 3 terdiri dari 133 mahasiswa dengan nilai tinggi dan indeks prestasi yang cenderung tinggi, namun lebih rendah dibandingkan klaster 2 dengan rata – rata 3.13.
9	Penelitian saat ini	Skor Kuesioner <i>Index of Learning Styles</i> (ILS), <i>K-Means Clustering, Silhouette Score</i>	

## 2.2 Bimbingan Belajar

Bimbingan belajar merupakan salah satu sarana yang kerap dilaksanakan diluar jam pelajaran sekolah untuk membantu kesulitan yang dihadapi oleh siswa. Salah satu kesulitan yang dialami siswa adalah kesulitan untuk memahami materi yang dijelaskan di sekolah. Hal ini terjadi bukan semata – mata karena kebodohan siswa. Kondisi belajar yang tidak nyaman dan kondusif, penjelasan guru yang kurang efektif, serta kemampuan individu siswa dalam memahami materi dapat menyebabkan kesulitan ketika belajar. Untuk mengatasi masalah tersebut, bimbingan belajar diadakan dengan tujuan membantu siswa mengatasi kesulitan

serta masalah yang mereka hadapi ketika belajar di sekolah (Subakti & Handayani, 2020).

### **2.3 Gaya Belajar**

Gaya belajar merupakan sebuah rangkaian serta kombinasi dari kemampuan siswa untuk menyerap, mengolah, serta menerapkan sebuah informasi atau pengetahuan (Ediyanto, 2022). Gaya belajar merepresentasikan cara bagaimana kita memproses sebuah informasi yang kemudian akan diolah menjadi ilmu yang bisa kita amalkan. Hal ini juga dapat diartikan sebagai kebiasaan atau tatacara yang sering kita lakukan ketika sedang mempelajari sesuatu yang baru. Seorang siswa sangat mungkin untuk memiliki gaya belajar yang berbeda dengan siswa yang lain.

Perbedaan yang dimaksud adalah perbedaan kemampuan masing – masing siswa dalam proses mengolah informasi. Ketika guru menjelaskan sebuah materi di kelas, tidak semua siswa bisa memahami apa yang telah dijelaskan oleh guru tersebut secara langsung. Beberapa siswa mungkin memerlukan waktu untuk memahami materi tersebut. Proses memahami yang dilakukan oleh siswa inilah yang memiliki kaitan erat dengan gaya belajar. Menurut David Kolb, gaya belajar merupakan salah satu instrumen penting yang diperlukan untuk memahami sebuah informasi secara efektif. Seorang siswa perlu menerapkan gaya belajar yang sesuai dengan kemampuan mereka supaya bisa belajar secara efisien dan mendapat hasil belajar yang baik (Telaumbanua & Harefa, 2024).

Untuk menyelesaikan permasalahan belajar pada siswa, banyak model pembelajaran yang telah dikembangkan. Salah satunya adalah model yang dikembangkan oleh David Kolb. Menurut David Kolb, pelajar atau siswa memiliki

preferensi tersendiri pada cara mereka menerima dan memproses informasi ( Felder & Brent, 2005). Pada model *Kolb's Experiental Learning Model*, David Kolb (dalam Felder & Brent, 2005) membagi beberapa skema yang cenderung digunakan pelajar, diantara lain:

1. *Divergen*, di mana pelajar atau siswa akan memahami pelajaran dengan baik jika dijelaskan dengan menghubungkan pelajaran tersebut dengan pengalaman serta minat mereka.
2. *Assimilator*, di mana pelajar atau siswa akan memahami pelajaran dengan baik jika dijelaskan dengan terorganisir, serta logis. Hal ini akan bermanfaat jika siswa diberi waktu untuk refleksi serta memahaminya perlahan.
3. *Convergen*, di mana pelajar atau siswa akan memahami pelajaran dengan baik jika mereka memiliki kesempatan untuk berkekspresi dan bekerja secara aktif pada tugas yang diberikan serta belajar secara uji coba pada lingkungan yang tidak mengharuskan tugas diselesaikan secara sempurna.
4. *Accomodator*, di mana pelajar atau siswa akan memahami pelajaran dengan baik jika mereka diberi kesempatan untuk menyelesaikan permasalahan nyata yang sedang mereka hadapi.

Menurut Richard Felder, proses pembelajaran melibatkan dua tahap penting, yaitu tahap penerimaan dan tahap pemrosesan informasi yang diperoleh. Pada kedua tahap ini, siswa akan menerima seluruh informasi yang diberikan, kemudian siswa akan menghafal, menalar serta merefleksikan informasi tersebut sebagai bentuk pemrosesan informasi. Dari proses tersebut, muncul sebuah hasil

berupa sebuah pemahaman, apakah siswa tersebut memahami materi tersebut atau tidak (Felder & Silverman, 1988). Pada model *Felder-Silverman Model*, Felder dan Silverman mendefinisikan gaya belajar sesuai dengan jawaban dari beberapa pertanyaan, diantara lain:

1. Jenis informasi yang disukai siswa, yaitu *sensory* yang cenderung berorientasi pada fakta atau prosesdur langsung, atau *intuitive* yang lebih tertarik dengan teori serta cenderung lebih baik ketika menyelesaikan masalah.
2. Jenis informasi sensorik yang paling efektif, yaitu *visual* seperti gambar, diagram serta demonstrasi atau *verbal* seperti deskripsi dan penjelasan secara lisan.
3. Bagaimana siswa memproses informasi, yaitu *active*, di mana siswa akan terlibat dalam sebuah diskusi atau *reflective*, di mana siswa cenderung melakukan introspeksi secara mandiri.
4. Bagaimana karakteristik pemahaman siswa, yaitu *sequentially*, di mana siswa akan berpikir secara berurutan dan mampu mengerjakan sesuatu hanya dengan sebagian materi atau *globally*, di mana siswa memerlukan pemahaman sepenuhnya pada sebuah materi untuk bisa menerapkannya (Felder & Brent, 2005).

#### **2.4 Index of Learning Styles**

*Index of Learning Styles* merupakan kuesioner dengan 44 butir pertanyaan yang seringkali digunakan sebagai alat ukur untuk menilai preferensi pada 4 dimensi yang diusulkan oleh Richard Felder dan Linda Silverman pada model

Felder-Silverman *Model*. Dimensi tersebut ialah *active/reflective*, *sensing/intuitive*, *visual/verbal*, *sequential/global*. Kuesioner ini pertama kali diciptakan oleh Richard Felder dan Barbara Solomon pada tahun 1991. Setiap dimensi pada model Felder-Silverman *Model* memiliki hubungan pada pertanyaan di kuesioner. Masing – masing dimensi akan direpresentasikan dengan 11 pertanyaan yang sesuai dengan salah satu kategori pada dimensi tersebut.

Sebagai contoh, pada dimensi *active/reflective*, pengguna bisa menjawab 11 pertanyaan yang akan menunjukkan preferensi mereka. Jika pengguna menjawab a, maka pengguna memiliki preferensi *active*. Sebaliknya, jika pengguna menjawab b, maka pengguna memiliki preferensi *reflective*, begitu juga dengan dimensi yang lain. Dari pertanyaan yang telah dijawab, sisi yang lebih dominan akan menunjukkan preferensi siswa pada dimensi tersebut. Kuesioner ini menggunakan skala untuk menentukan preferensi, dimana skor 1 atau 3 menunjukkan preferensi yang ringan (*mild preference*), skor 5 atau 7 menunjukkan preferensi sedang (*moderate preference*), dan skor 9 atau 11 menunjukkan preferensi yang kuat (*strong preference*) pada dimensi tersebut (Felder & Spurlin, 2005). Dari kuesioner tersebut, siswa akan mendapatkan dimensi yang lebih disukai berdasarkan skala yang didapat. Sebagai contoh, pada dimensi *active/reflective* siswa mendapatkan skala 5 pada kategori *active*, yang menunjukkan bahwa siswa tersebut cenderung menyukai gaya belajar yang aktif seperti diskusi. Pada penelitian ini, peneliti akan menggunakan seluruh kategori dari 4 dimensi yang ada pada *Index of Learning Styles*.

## 2.5 *Clustering*

*Clustering* merupakan sebuah metode yang digunakan untuk mengelompokkan sekumpulan objek atau data dimana data pada satu kelompok atau klaster memiliki karakteristik serta pola yang serupa satu sama lain dibandingkan dengan data pada kelompok yang lain. *Clustering* seringkali digunakan untuk melakukan analisis sebuah data secara statistik. Setelah dilakukan proses analisa, hasil dari analisa tersebut seringkali digunakan untuk menentukan strategi atau langkah yang harus diambil (Chen et al., 2023). *Clustering* bisa diaplikasikan pada berbagai sektor. Sebagai contoh, pada sektor pendidikan, kita bisa melakukan segmentasi siswa di sebuah kelas berdasarkan nilai pada beberapa mata pelajaran. Dari nilai tersebut, kita bisa mengetahui kemampuan dari masing – masing siswa. Hasil tersebut bisa digunakan oleh guru untuk mengevaluasi proses pembelajaran yang telah berlangsung.

Untuk melakukan proses pengelompokan, telah dikembangkan berbagai macam metode yang diusulkan untuk mengolah data ke dalam beberapa klaster. Beberapa kategorinya diantara lain:

1. *Center-Based Clustering*, yang melakukan partisi data ke dalam beberapa klaster dengan mencari metrik kesalahan terkecil diantara sampel data pada sebuah klaster. Berdasarkan metrik kesalahan yang telah ditentukan, pusat klaster dapat berupa rata – rata dari sampel pada sebuah klaster (seperti *k-means*), atau *median* dari sampel pada sebuah klaster (seperti *k-medoids*).
2. *Hierarchical Clustering*, yang bertujuan untuk membagi data ke dalam sebuah hirarki yang direpresentasikan sebagai pohon biner. Simpul akar

merepresentasikan seluruh kumpulan data, simpul daun terdiri dari sampel tunggal dari kumpulan data, sedangkan simpul pohon yang tersisa merepresentasikan klaster yang ada.

3. *Mixture Model-Based Clustering*, yang menggunakan pendekatan pengelompokan berdasarkan probabilitas di mana titik data ditetapkan pada klaster secara lebih fleksibel. Salah satu contoh model yang menggunakan kategori ini adalah *Gaussian Mixture-Model*.
4. *Graph-Based Clustering*, yang memanfaatkan konsep dari grafik untuk mengelompokkan data. Data akan diubah menjadi simpul – simpul grafik yang akan membuat tepi atau *edge* di antara sampel data menggunakan metrik kesamaan atau ketidaksamaan.
5. *Fuzzy Clustering*, yang terdiri dari pendekatan pengelompokan yang fleksibel, di mana setiap data memiliki nilai *fuzzy* antara 0 dan 1 pada sebuah klaster. Salah satu contoh dari metode ini adalah *Fuzzy C-Means*.
6. *Combinational Search-Based Clustering*, yang memberikan solusi untuk masalah pengelompokan yang memerlukan hasil yang optimal. Metode ini mengombinasikan berbagai pendekatan untuk mendapatkan hasil yang optimal. pendekatan yang termasuk pada kategori ini adalah *Genetic K-means Algorithm* (Chhabra et al., 2021).

## 2.6 Algoritma K-Means Clustering

Secara umum, terdapat dua pendekatan yang digunakan untuk melakukan *clustering*. Pendekatan tersebut adalah *Hierarchical Clustering*, dimana pendekatan ini membagi kumpulan data menjadi beberapa cabang yang direpresentasikan

sebagai pohon biner. Pendekatan lain yang sering digunakan adalah *Partitional Clustering*, dimana pembagian kumpulan data direpresentasikan sebagai sebuah titik klaster yang memiliki fungsi objektifnya sendiri. Pendekatan ini memerlukan perhitungan perbedaan jarak antara titik data dengan titik klaster yang ada. Algoritma K-Means merupakan algoritma paling populer yang menggunakan pendekatan *Partitional Clustering* (Sinaga & Yang, 2020).

K-means membagi kumpulan data menjadi beberapa klaster yang disimbolkan dengan  $k$ . Pada umumnya, jumlah  $k$  akan diinisialisasi secara acak di awal proses perhitungan k-means. Metode ini terbagi menjadi dua tahap. Tahap pertama merupakan menentukan jumlah  $k$  awal yang akan menjadi jumlah klaster yang ada. Klaster ini akan memiliki sebuah titik yang mewakili pusat dari sebuah klaster, yang disebut sebagai *centroid*. Setelah mendapatkan *centroid*, kita perlu menghitung jarak antara titik data yang ada dengan titik pusat klaster atau *centroid* menggunakan *Euclidean Distance*. Berikut adalah rumus perhitungan *Euclidean Distance*.

$$d(x_i, c_i) = \sqrt{\sum_{i=1}^n (x_i - c_i)^2} \quad (2.1)$$

Keterangan :

$d(x_i, c_i)$  = nilai *Euclidean Distance*

$x_i$  = titik data data ke  $i$

$c_i$  = titik *centroid* ke  $i$

$n$  = jumlah data

Setelah jarak didapatkan, titik data akan dimasukkan pada klaster yang memiliki jarak terkecil antara titik data dengan centroid yang ada. Tahap kedua adalah menghitung ulang posisi *centroid*. Hal ini dilakukan supaya titik data masuk

ke dalam kelompok klaster dengan jarak terkecil dari tiap *centroid*. Berikut adalah rumus perhitungan ulang *centroid*.

$$c_k = \left(\frac{1}{n}\right) \sum x_i \quad (2.2)$$

Keterangan :

$c_k$  = nilai *centroid* baru

$n$  = jumlah data

$x_i$  = titik data ke  $i$

Setelah *centroid* baru didapatkan, kita akan terus mengulangi tahap pertama dan kedua sampai tidak ada perubahan nilai *centroid*. Hal ini berarti *centroid* telah berada tepat di tengah – tengah klaster yang ada (Andi Akram Nur Risal et al., 2024).

## 2.7 Sihouette Score

*Silhouette* merupakan sebuah sebuah metode yang digunakan untuk melakukan interpretasi serta mengukur validitas sebuah klaster. Perhitungan ini dilakukan supaya kita bisa mengetahui kualitas pembagian data pada klaster yang ada. Apakah klaster yang keluar sebagai hasil merupakan klaster yang tepat? Pertanyaan tersebut dapat dijawab dengan menghitung *silhouette score* pada sebaran data yang ada dengan menghitung rata – rata dari sampel data. Nilai *silhouette* berkisar pada -1 atau 1. Dimana nilai terbesar mengindikasikan tingkat pemisahan klaster yang lebih tinggi atau lebih baik (Pavlopoulos et al., 2025). Berikut merupakan rumus untuk menghitung *silhouette score*.

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (2.3)$$

Keterangan :

$s(i)$  = nilai *silhouette score*

$b(i)$  = jarak rata – rata antara titik ( $i$ ) dengan semua titik data pada klaster yang lain

$a(i)$  = jarak rata – rata antara titik ( $i$ ) dengan semua titik data pada klaster yang sama

Untuk mendapatkan nilai  $a(i)$ , dapat dihitung dengan rumus,

$$a(i) = \frac{1}{|C_i|-1} \sum_{j \in C_i, j \neq i} d(i, j) \quad (2.4)$$

Keterangan :

$a(i)$  = jarak rata – rata antara titik ( $i$ ) dengan semua titik data pada klaster yang sama

$|C_i|$  = jumlah data pada klaster

$d(i, j)$  = jarak titik  $i$  ke titik  $j$  pada klaster yang sama

Untuk mendapatkan nilai  $b(i)$ , dapat dihitung dengan rumus,

$$b(i) = \frac{1}{|C_j|} \sum_{j \in C_j} d(i, j) \quad (2.5)$$

Keterangan :

$b(i)$  = jarak rata – rata antara titik ( $i$ ) dengan semua titik data pada klaster yang lain

$|C_j|$  = jumlah data pada klaster selain  $i$

$d(i, j)$  = jarak titik  $i$  ke titik  $j$  pada klaster yang lain

Nilai  $s(i)$  berada pada rentang -1 sampai dengan 1. Jika  $s(i)$  bernilai mendekati 1, maka titik data  $i$  diklasterisasi dengan baik. Jika  $s(i)$  bernilai mendekati 0, maka titik data  $i$  berada diantara dua klaster. Jika  $s(i)$  bernilai mendekati -1, maka titik data  $i$  berada pada klaster yang tidak sesuai (Hidayati et al., 2021). Rentang nilai *silhouette score* disajikan pada Tabel 2.1.

Tabel 2.2. Interpretasi Nilai *Silhouette Score*

<i>Silhouette Score</i>	Interpretasi
0.71 – 1.00	Kuat
0.51 – 0.70	Baik
0.26 – 0.50	Lemah
$\leq 1$	Tidak terstruktur

## 2.8 Elbow Method

*Elbow Method* merupakan sebuah metode yang digunakan untuk menentukan jumlah klaster yang optimal pada proses *clustering*. Metode ini sering

digunakan untuk melakukan validasi setelah pengolahan data yang dilakukan menggunakan Algoritma K-Means Clustering. Jumlah k paling optimal didapatkan melalui hasil perbandingan persentase jumlah klaster yang membentuk pola seperti siku, dimana perhitungan ini dilakukan dengan menggunakan *Sum Squared Error* (SSE) (Anggreani et al., 2024). Nilai SSE dapat diperoleh melalui persamaan berikut:

$$SSE = \sum_{i=1}^k \sum_{x_i \in S_k} \|x_i - c_k\|^2 \quad (2.6)$$

Keterangan :

$k$  = jumlah klaster

$S_k$  = data pada klaster  $k$

$x_i$  = titik data

$C_k$  = *centroid* klaster  $i$

$\|x_i - c_k\|^2$  = jarak Euclidean antara titik data dan *centroid*

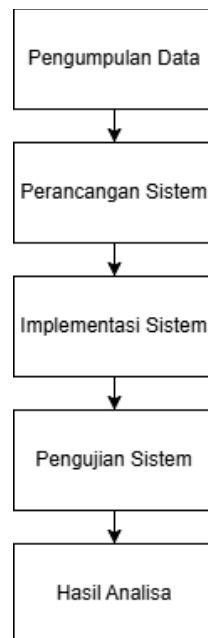
## **BAB III**

### **METODOLOGI PENELITIAN**

Bagian ini menjelaskan mengenai tahapan serta metodologi yang akan digunakan pada penelitian ini.

#### **3.1 Desain Penelitian**

Desain penelitian merupakan alur yang terdiri dari langkah – langkah atau tatacara dari pelaksanaan penelitian yang dilakukan. Gambaran desain penelitian akan disajikan pada Gambar 3.1.



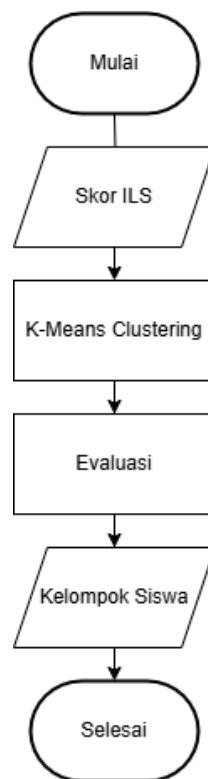
Gambar 3.1. Desain Penelitian

Penelitian ini dimulai dengan mengumpulkan data berupa kuesioner ILS yang didapatkan dari Siswa Bimbingan Belajar Tesla Sidoarjo. Kemudian, sebuah sistem yang digunakan untuk mengolah data tersebut akan dirancang. Sistem yang

dirancang pada penelitian ini merupakan sebuah sistem untuk mengelompokkan Siswa Bimbingan Belajar Tesla Sidoarjo menggunakan Algoritma *K-Means Clustering*. Sistem tersebut dirancang menggunakan Bahasa Pemrograman *Python* yang ditulis pada *Text Editor Visual Studio Code*. Sistem tersebut kemudian akan diuji menggunakan *silhouette score*, dimana nilai *silhouette score* akan menunjukkan performa dari sistem yang telah dirancang. Hasil berupa analisa serta deskripsi dari masing – masing klaster yang terdiri dari Siswa Bimbingan Belajar Tesla Sidoarjo bisa didapatkan dari implementasi dan pengujian sistem yang telah dilakukan.

### 3.1.1 Desain Sistem

Desain sistem merupakan alur atau rancangan dari proses sistem akan berjalan. Sistem ini memiliki masing – masing elemen yang berhubungan, dimana elemen ini akan dirancang supaya sistem berjalan sesuai dengan harapan. Pada penelitian ini, terdapat 3 tahapan yang akan dijalankan, yaitu tahap *input, process*, serta *output* yang akan dihasilkan. Tahapan tersebut akan disajikan pada Gambar 3.2.



Gambar 3.2. Desain Sistem

Pada tahapan *input*, data – data yang digunakan akan dimasukkan ke dalam sistem. Data tersebut akan diolah dengan algoritma yang ada untuk mendapatkan hasil yang diinginkan. Data akan diolah menggunakan Algoritma *K-Means Clustering*. Setelah *clustering* dilakukan, data yang telah terbagi menjadi klaster – klaster yang berbeda akan keluar sebagai *output* dari sistem.

### 3.1.2 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data siswa pada Bimbingan Belajar Tesla Sidoarjo. Proses pengambilan data dilakukan pada bulan Februari hingga bulan Mei Tahun 2025. Data tersebut terdiri hasil skor kuesioner ILS yang telah dibagikan. Skor kuesioner ILS berisi skor yang didapat dari

kuesioner yang terbagi menjadi 4 dimensi dengan 2 kategori pada masing – masing dimensi. Berikut merupakan contoh data siswa yang disajikan pada Tabel 3.1.

Tabel 3.1. Tabel Data Skor ILS Siswa

<b>Id</b>	<b>active</b>	<b>reflective</b>	<b>sense</b>	<b>intuitive</b>	<b>visual</b>	<b>verbal</b>	<b>sequential</b>	<b>global</b>
S1	8	3	7	4	8	3	4	7
S2	5	6	6	5	10	1	3	8
S3	3	8	9	2	5	6	5	6
S4	6	5	8	3	2	9	6	5
S5	4	7	8	3	10	1	7	4
S6	10	1	8	3	3	8	4	7

### 3.1.3 Praproses Data

Praproses data merupakan sebuah tahapan sebelum proses pengolahan data yang perlu dilakukan supaya data yang akan digunakan siap untuk diolah menggunakan algoritma yang digunakan. Data siswa akan disamakan supaya tidak ada ketimpangan antara satu atribut dengan atribut lainnya dengan cara normalisasi. Normalisasi dilakukan dengan tujuan menyamakan skala dari data yang akan diolah. Normalisasi data akan dilakukan pada data siswa yang terdiri dari hasil skor kuesioner ILS. Teknik *Min-Max Scaling* akan digunakan untuk menyamakan rentang data dengan nilai antara 0 dan 1. Normalisasi *Min-Max Scaling* dilakukan dengan Persamaan 3.1.

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (3.1)$$

Keterangan :

$x'$  = nilai data  $x$  setelah di normalisasi

$x$  = nilai data

$x_{min}$  = nilai data terkecil

$x_{max}$  = nilai data terbesar

Berikut merupakan contoh perhitungan normalisasi pada titik data S1 pada parameter *active*.

Pada titik data S1 :

$$S1 = \frac{8 - 3}{10 - 3}$$

$$= 0.714$$

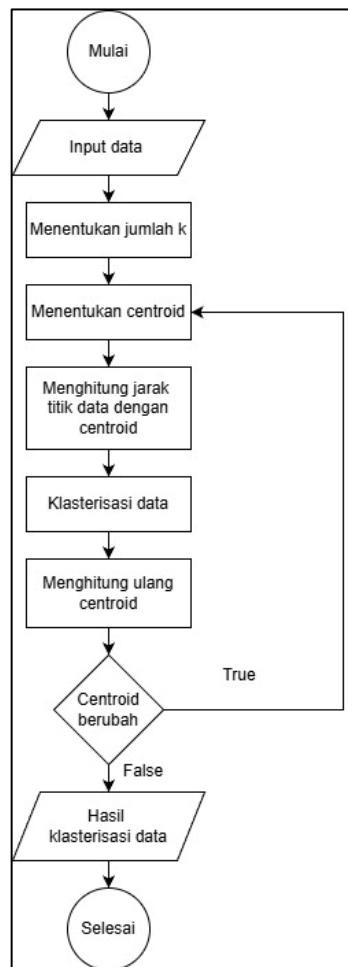
Proses normalisasi tersebut dilakukan pada setiap titik data di semua parameter yang digunakan. Berikut merupakan contoh data yang telah melalui proses normalisasi yang disajikan pada Tabel 3.2.

Tabel 3.2. Tabel Data Normalisasi

<b>Id</b>	<b>active</b>	<b>reflective</b>	<b>sense</b>	<b>intuitive</b>	<b>visual</b>	<b>verbal</b>	<b>sequential</b>	<b>global</b>
S1	0.714	0.286	0.333	0.667	0.750	0.250	0.250	0.750
S2	0.286	0.714	0.000	1.000	1.000	0.000	0.000	1.000
S3	0.000	1.000	1.000	0.000	0.375	0.625	0.500	0.500
S4	0.429	0.571	0.667	0.333	0.000	1.000	0.750	0.250
S5	0.143	0.857	0.667	0.333	1.000	0.000	1.000	0.000
S6	1.000	0.000	0.667	0.333	0.125	0.875	0.250	0.750

### 3.1.4 K-Means Clustering

Sebelum data diolah, langkah pertama yang harus dilakukan adalah menentukan jumlah  $k$ , yang menjadi jumlah klaster yang terbentuk. Setelah  $k$  inisial didapatkan, proses clustering akan dilakukan dengan cara menghitung jarak antara titik data dengan titik pusat atau *centroid* yang telah ditentukan. Data dengan jarak *centroid* terdekat akan masuk ke dalam klaster tersebut. Dari anggota klaster yang didapat, *centroid* akan dihitung ulang untuk melihat jika ada pergeseran pada titik *centroid*. Iterasi ini terus dilakukan secara berulang sampai *centroid* tidak menunjukkan perubahan nilai. Alur kerja dari Algoritma K-Means Clustering disajikan pada Gambar 3.3.



Gambar 3.3. *Flowchart K-Means Clustering*

Jika dari iterasi yang terus dilakukan tidak ada perubahan nilai dari *centroid*, maka proses *clustering* data telah selesai.

### 3.1.5 Skenario Uji Coba

Pengolahan data menggunakan Algoritma K-Means Clustering dilakukan dengan beberapa skenario. Skenario tersebut akan memiliki perbedaan pada jumlah klaster dengan nilai 2, 3, 4, 5, 6, 7, 8, 9, 10 dan 11 yang dicantumkan pada Tabel 3.3.

Tabel 3.3. Skenario Uji Coba *Cluster*

Skenario uji coba	Jumlah klaster
1	2
2	3
3	4
4	5
5	6
6	7
7	8
8	9
9	10
10	11

Berikut merupakan skenario pengujian perhitungan Algoritma *K-Means Clustering* pada percobaan skenario 1 dengan jumlah  $k$  bernilai 2.

1. Menentukan jumlah  $k$

Pada percobaan ini, jumlah klaster yang akan digunakan adalah 2.

2. Menentukan *centroid*

Setelah jumlah  $k$  ditentukan, *centroid* akan ditentukan secara acak.

Pada contoh skenario 1, *centroid* akan jatuh pada data dengan id S2 dan S6, atau  $C_0 = S_2$  dan  $C_1 = S_6$ . Titik *centroid* data disajikan pada Tabel 3.4.

Tabel 3.4. Tabel *Centroid* pada Skenario 1

Id	active	reflective	sense	intuitive	visual	verbal	sequential	global
S2	0.286	0.714	0.000	1.000	1.000	0.000	0.000	1.000
S5	1.000	0.000	0.667	0.333	0.125	0.875	0.250	0.750

### 3. Menghitung jarak titik data dengan *centroid*

Setelah *centroid* didapatkan, jarak antara setiap titik data dan *centroid* akan dihitung menggunakan *Euclidean Distance* dengan Persamaan 3.2.

$$d(x_i, c_i) = \sqrt{\sum_{i=1}^n (x_i - c_i)^2} \quad (3.2)$$

Dimana  $(x_i - c_i)$  merupakan jarak antara titik data  $i$  dengan centroid  $i$ , dan  $n$  merupakan jumlah data. Berikut merupakan contoh perhitungan jarak titik data dengan *centroid* pertama yaitu C0 pada iterasi pertama.

Pada titik data S1 dengan C0 :

$$\begin{aligned} S1 &= \sqrt{(0.714 - 0.286)^2 + (0.286 - 0.714)^2 + (0.333 - 0)^2 + (0.667 - 1)^2 + \\ &\quad (0.750 - 1)^2 + (0.250 - 0)^2 + (0.250 - 0)^2 + (0.750 - 1)^2} \\ &= 0.916 \end{aligned}$$

Pada titik data S2 dengan C0 :

$$\begin{aligned} S2 &= \sqrt{(0.286 - 0.286)^2 + (0.714 - 0.714)^2 + (0 - 0)^2 + (1 - 1)^2 + \\ &\quad (1 - 1)^2 + (0 - 0)^2 + (0 - 0)^2 + (1 - 1)^2} \\ &= 0 \end{aligned}$$

Selanjutnya jarak antara setiap titik data dengan *centroid* kedua dihitung menggunakan rumus yang sama. Berikut merupakan perhitungan titik data dengan *centroid* kedua yaitu C1.

Pada titik data S1 dengan C1 :

$$\begin{aligned} S1 &= \sqrt{(0.714 - 1)^2 + (0.286 - 0)^2 + (0.333 - 0.667)^2 + (0.667 - 0.333)^2 + \\ &\quad (0.750 - 0.125)^2 + (0.250 - 0.875)^2 + (0.250 - 0.250)^2 + (0.750 - 0.750)^2} \\ &= 1.080 \end{aligned}$$

Pada titik data S2 dengan C1 :

$$S1 = \sqrt{(0.286 - 1)^2 + (0.714 - 0)^2 + (0 - 0.667)^2 + (1 - 0.333)^2 + (1 - 0.125)^2 + (0 - 0.875)^2 + (0 - 0.250)^2 + (1 - 0.750)^2} \\ = 1.888$$

Perhitungan jarak titik data dengan *centroid* dilakukan pada seluruh titik data yang ada. Berikut merupakan hasil perhitungan *Euclidean Distance* pada iterasi pertama yang disajikan pada Tabel 3.5.

Tabel 3.5. Jarak Titik Data dengan *Centroid* Iterasi 1 pada Skenario 1

Id	Jarak C0	Jarak C1
S1	0.916	1.080
S2	0.000	1.888
S3	1.856	1.572
S4	2.014	1.088
S5	1.712	2.031
S6	1.888	0.000

Setelah jarak antara setiap titik data dengan *centroid* telah dihitung, langkah selanjutnya adalah mengelompokkan data ke dalam klaster dengan jarak terkecil pada titik data tersebut. Hasil *clustering* pada iterasi pertama disajikan pada Tabel 3.6.

Tabel 3.6. Persebaran Data Iterasi 1 pada Skenario 1

Id	Jarak C0	Jarak C1	Min	C0	C1
S1	0.916	1.080	0.916	1	0
S2	0.000	1.888	0.000	1	0
S3	1.856	1.572	1.572	0	1
S4	2.014	1.088	1.088	0	1
S5	1.712	2.031	1.712	1	0
S6	1.888	0.000	0.000	0	1

#### 4. Menghitung perubahan nilai *centroid*

Tahap selanjutnya adalah menghitung pergeseran *centroid*. Untuk menghitung pergeseran *centroid* dapat dihitung menggunakan Persamaan 3.3.

$$c_k = \left(\frac{1}{n}\right) \sum x_i \quad (3.3)$$

Dimana  $n$  merupakan jumlah data pada klaster tersebut dan  $x_i$  merupakan titik data ke  $i$ . Berikut merupakan contoh perhitungan pergeseran *centroid* pada parameter *active*.

$$\text{Pada titik } C_0 = \frac{0.714+0.286+0.143}{3}$$

$$= 0.381$$

$$\text{Pada titik } C_1 = \frac{0+0.429+1}{3}$$

$$= 0.476$$

Perhitungan tersebut dilakukan pada seluruh parameter yang digunakan. Setelah nilai *centroid* yang diperbarui telah didapatkan, langkah selanjutnya merupakan melanjutkan perhitungan pada iterasi kedua dengan menghitung jarak antara titik data dengan *centroid* yang telah diperbarui. Hasil perhitungan pada iterasi kedua dapat dilihat pada Tabel 3.7.

Tabel 3.7. Jarak Titik Data dengan *Centroid* Baru Iterasi 2 pada Skenario 1

<b>Id</b>	<b>Jarak C0</b>	<b>Jarak C1</b>	<b>Min</b>	<b>C0</b>	<b>C1</b>
S1	0.556	1.146	0.556	1	0
S2	0.688	1.781	0.688	1	0
S3	2.016	0.799	0.799	0	1
S4	2.175	0.458	0.458	0	1
S5	1.211	1.461	1.211	1	0
S6	2.534	0.838	0.838	0	1

Langkah tersebut terus menerus diulangi hingga tidak ada perubahan nilai pada *centroid*. *Centroid* merupakan pusat dari sebuah klaster. Jika *centroid* tidak berubah tempat, maka anggota pada klaster tersebut juga sama. Dikarenakan pada iterasi kedua tidak ada perubahan posisi pada anggota klaster, hal tersebut berarti tidak ada perubahan nilai *centroid*. Karena nilai *centroid* tidak berubah, maka proses perhitungan dihentikan sampai dengan iterasi ke-2 dengan hasil klaster 1 terdiri dari siswa dengan id S1, S2 dan S5 serta klaster 2 terdiri dari siswa dengan id S3, S4, S6 yang dicantumkan pada Tabel 3.8.

Tabel 3.8. Hasil Klasterisasi Siswa pada Skenario 1

<b>Id</b>	<b>Jarak C0</b>	<b>Jarak C1</b>	<b>Min</b>	<b>C0</b>	<b>C1</b>	<b>Klaster</b>
S1	0.556	1.146	0.556	1	0	0
S2	0.688	1.781	0.688	1	0	0
S3	2.016	0.799	0.799	0	1	1
S4	2.175	0.458	0.458	0	1	1
S5	1.211	1.461	1.211	1	0	0
S6	2.534	0.838	0.838	0	1	1

Pada skenario 2 dan 3, proses klasterisasi dilakukan menggunakan langkah-langkah serta persamaan yang sama yang digunakan pada percobaan skenario 1. Namun, dengan perbedaan jumlah  $k$ , proses klasterisasi juga akan mengalami beberapa perbedaan. Berikut merupakan hasil klasterisasi pada skenario 2 dengan jumlah  $k$  bernilai 3 yang disajikan pada Tabel 3.9.

Tabel 3.9. Hasil Klasterisasi Siswa pada Skenario 2

<b>Id</b>	<b>Jarak C0</b>	<b>Jarak C1</b>	<b>Jarak C2</b>	<b>Min</b>	<b>C0</b>	<b>C1</b>	<b>C2</b>	<b>Klaster</b>
S1	0.613	1.625	1.609	0.613	1	0	0	0
S2	0.613	1.755	2.062	0.613	1	0	0	0
S3	1.753	0.716	1.356	0.716	0	1	0	1
S4	1.965	1.217	0.569	0.569	0	0	1	2
S5	1.710	0.716	1.729	0.716	0	1	0	1
S6	1.696	1.723	0.569	0.569	0	0	1	2

Berikut merupakan hasil klasterisasi pada skenario 3 dengan jumlah k bernilai 4 yang disajikan pada Tabel 3.10.

Tabel 3.10. Hasil Klasterisasi Siswa pada Skenario 3

<b>Id</b>	<b>Jarak C0</b>	<b>Jarak C1</b>	<b>Jarak C2</b>	<b>Jarak C3</b>	<b>Min</b>	<b>C0</b>	<b>C1</b>	<b>C2</b>	<b>C3</b>	<b>Klaster</b>
S1	0.613	1.830	1.720	1.609	0.613	1	0	0	0	0
S2	0.613	1.883	1.908	2.062	0.613	1	0	0	0	0
S3	1.753	0.000	1.433	1.356	0.000	0	1	0	0	1
S4	1.965	1.276	1.537	0.569	0.569	0	0	0	1	3
S5	1.710	1.433	0.000	1.729	0.000	0	0	1	0	2
S6	1.696	1.643	2.065	0.569	0.569	0	0	0	1	3

Skenario percobaan yang telah dilakukan memberikan *output* berupa 3 klasterisasi pada siswa. Kualitas klaster dari masing-masing percobaan yang telah dilakukan dapat diukur menggunakan Evaluasi *Silhouette Score*.

### 3.1.6 Evaluasi *Silhouette Score*

*Silhouette score* merupakan salah satu teknik pengujian yang kerap digunakan untuk mengukur validitas sebuah klaster. Pada penelitian ini, *silhouette score* akan digunakan sebagai pengujian untuk mengetahui kualitas dari klaster yang dibuat.

Perhitungan *silhouette score* akan dilakukan pada skenario 1. Untuk menghitung *silhouette score*, kita perlu mendapatkan jarak rata – rata antara titik (*i*) dengan semua titik data pada klaster yang lain atau  $b(i)$  yang disajikan pada Persamaan 3.4.

$$b(i) = \frac{1}{|C_j|} \sum_{j \in C_j} d(i, j) \quad (3.4)$$

Dimana  $C_j$  merupakan jumlah data pada klaster lain dan  $d(i, j)$  merupakan jarak antara titik data *i* dan titik data *j* pada klaster lain.

Berikut merupakan contoh perhitungan  $b(i)$  pada titik data S1 dengan S3.

Jarak S1 dengan S3

$$\begin{aligned}
 &= \sqrt{\frac{(0.714 - 0)^2 + (0.286 - 1)^2 + (0.333 - 1)^2 + (0.667 - 0)^2 +}{(0.750 - 0.375)^2 + (0.250 - 0.625)^2 + (0.250 - 0.500)^2 + (0.750 - 0.500)^2}} \\
 &= 1.522
 \end{aligned}$$

Selanjutnya, nilai  $b(i)$  merupakan

$$b(i) = \frac{1.522 + 1.418 + 1.080}{3}$$

$$b(i) = 1.340$$

Langkah selanjutnya, kita perlu mendapatkan jarak rata – rata antara titik ( $i$ ) dengan semua titik data pada klaster yang sama atau  $a(i)$  yang disajikan pada Persamaan 3.5.

$$a(i) = \frac{1}{|C_i|-1} \sum_{j \in C_i, j \neq i} d(i, j) \quad (3.5)$$

Dimana  $C_i$  merupakan jumlah data pada klaster yang sama dan  $d(i, j)$  merupakan jarak antara titik data  $i$  dengan data  $j$  pada klaster yang sama.

Berikut merupakan contoh perhitungan  $a(i)$  pada titik data S1 dengan S2.

Jarak S1 dengan S2

$$\begin{aligned}
 &= \sqrt{\frac{(0.714 - 0.286)^2 + (0.286 - 0.714)^2 + (0.333 - 0)^2 + (0.667 - 1)^2 +}{(0.750 - 1)^2 + (0.250 - 0)^2 + (0.250 - 0)^2 + (0.750 - 1)^2}} \\
 &= 0.916
 \end{aligned}$$

Selanjutnya, nilai  $a(i)$  merupakan

$$a(i) = \frac{0.916+1.458}{2}$$

$$a(i) = 1.187$$

Setelah nilai  $b(i)$  dan  $a(i)$  yang telah didapatkan, nilai  $s(i)$  dapat dihitung menggunakan rumus yang disajikan pada Persamaan 3.6.

$$s(i) = \frac{b(i)-a(i)}{\max(a(i), b(i))} \quad (3.6)$$

Berikut merupakan contoh perhitungan  $s(i)$  pada titik data S1.

$$s(i) = \frac{1.187-1.340}{1.340}$$

$$s(i) = 0.114$$

Perhitungan tersebut dilakukan pada seluruh titik data yang ada. Dari perhitungan yang dilakukan, rata – rata dari *silhouette score* pada *clustering* tersebut akan didapatkan. Hasil perhitungan *silhouette score* pada *clustering* siswa dapat dilihat pada Tabel 3.11.

Tabel 3.11. Hasil Perhitungan *Silhouette Score* pada Skenario 1

<b>Id</b>	<b><math>a(i)</math></b>	<b><math>b(i)</math></b>	<b><math>s(i)</math></b>
S1	1.187	1.340	0.114
S2	0.821	1.919	0.572
S3	0.801	1.540	0.480
S4	0.722	1.648	0.562
S5	0.979	1.596	0.386
S6	0.943	1.667	0.434
Rata – rata			0.423

Pengujian *silhouette score* akan dilakukan pada seluruh skenario yang ada. Hal ini dilakukan dengan tujuan untuk mengetahui skenario dengan jumlah klaster tertentu yang memiliki tingkat klasterisasi data yang lebih baik seperti yang dicantumkan pada Tabel 3.12.

Tabel 3.12. Skenario Uji Coba

Skenario uji coba	Jumlah klaster	Pengujian
1	2	<i>Silhouette Score</i>
2	3	
3	4	
4	5	
5	6	
6	7	
7	8	
8	9	
9	10	
10	11	

## **BAB IV**

### **HASIL DAN PEMBAHASAN**

Bagian ini menjelaskan mengenai hasil serta pembahasan yang didapatkan pada penelitian setelah data telah melalui proses pengolahan data serta serangkaian skenario yang telah ditentukan.

#### **4.1 Hasil Penelitian**

Pada bagian ini, akan dilakukan implementasi program menggunakan Algoritma K-Means *Clustering* pada beberapa skenario percobaan yang ada.

##### **4.1.1 Data Penelitian**

Data yang digunakan pada penelitian ini merupakan data siswa yang berasal dari Bimbingan Belajar Tesla Sidoarjo. Data yang berupa hasil kuesioner ILS akan melalui proses normalisasi terlebih dahulu. Hal ini dilakukan supaya tidak ada salah satu atribut yang lebih menonjol. Setelah normalisasi, data akan diolah menggunakan Algoritma K-Means *Clustering* serta diuji menggunakan *Silhouette Score*.

###### **a. Sumber Data**

Data yang digunakan pada penelitian ini merupakan data siswa yang meliputi data hasil kuesioner ILS yang telah dikerjakan oleh siswa. Data tersebut akan disajikan pada Tabel 4.1.

Tabel 4.1. Tabel Data

<b>Id</b>	<b>active</b>	<b>reflective</b>	<b>sense</b>	<b>intuitive</b>	<b>visual</b>	<b>verbal</b>	<b>sequential</b>	<b>global</b>
S1	5	6	6	5	10	1	3	8
S2	3	8	9	2	5	6	5	6
S3	6	5	8	3	2	9	6	5
S4	4	7	8	3	10	1	7	4
S5	10	1	8	3	3	8	4	7
S6	8	3	6	5	8	3	8	3
S7	5	6	7	4	6	5	6	5
S8	5	6	6	5	10	1	4	7
S9	6	5	8	3	5	6	7	4
S10	7	4	11	0	10	1	5	6
...	...	...	...	...	...	...	...	...
S51	7	4	4	7	7	4	6	5
S52	4	7	10	1	7	4	6	5
S53	8	3	9	2	8	3	9	2
S54	9	2	4	7	6	5	4	7
S55	8	3	10	1	8	3	5	6
S56	11	0	7	4	7	4	7	4
S57	6	5	7	4	8	3	4	7
S58	5	6	9	2	6	5	6	5

Data tersebut disimpan dengan format *csv* dan akan diinput kedalam sistem menggunakan *data frame* pada *library pandas*. Berikut merupakan implementasi input data pada sistem.

Tabel 4.3. Implementasi *Source Code* Normalisasi

<b>No</b>	<b>Source Code</b>
1	import pandas as pd
2	self.df = pd.read_csv(file_path)
3	self.filename_label.config(text=f"File: {file_path.split('/')[-1]}")
4	data_origin = self.df.drop(columns=['id', 'cluster'], errors='ignore')

Kolom *id* pada data merupakan atribut yang tidak digunakan, atribut ini tidak memiliki relevansi langsung pada proses klasterisasi siswa. Untuk itu, kolom *id* akan diabaikan tidak dimasukkan pada proses perhitungan selanjutnya. Sama halnya dengan kolom *id*, kolom *cluster* akan diabaikan supaya tidak mengganggu

proses *clustering* yang dijalankan. Kolom ini berasal dari *output* yang didapatkan dari proses *clustering* yang dilakukan.

### b. Praproses Data

Sebelum data diolah menggunakan Algoritma K-Means *Clustering*, data terlebih dahulu akan melalui proses normalisasi. Hal ini dilakukan dengan tujuan mendapatkan rentang data yang sama, sehingga tidak ada salah satu atribut yang lebih mendominasi atribut yang lain. Data akan melalui proses normalisasi, dimana data akan diubah menjadi skala 0 sampai dengan 1. Berikut merupakan implementasi normalisasi data pada python.

Tabel 4.4. Implementasi *Source Code* Normalisasi

No	Source Code
1	def scaler(x, xmax, xmin):
2	x_scaled = (x - xmin) / (xmax - xmin)
3	return x_scaled

Fungsi diatas merupakan penerapan rumus normalisasi pada python. Selanjutnya, fungsi tersebut akan dipanggil kedalam fungsi lain untuk melakukan normalisasi pada setiap data yang ada. Berikut merupakan implementasi normalisasi data pada python.

Tabel 4.5. Implementasi *Source Code* Normalisasi pada Data

No	Source Code
1	def normalize(df):
2	df_normalized = df.copy()
3	for i in df.columns:
4	xmax = df[i].max()
5	xmin = df[i].min()
6	df_normalized[i] = df[i].apply(lambda x: scaler(x, xmax, xmin))
7	return df_normalized

Fungsi tersebut akan mengubah setiap data yang berada pada variabel `df.copy()` menggunakan fungsi `scaler()` yang didefinisikan sebelumnya. Data yang telah melalui proses normalisasi kemudian disimpan pada variabel `df_normalized`. Data tersebut disajikan pada Tabel 4.6.

Tabel 4.6. Tabel Normalisasi Data

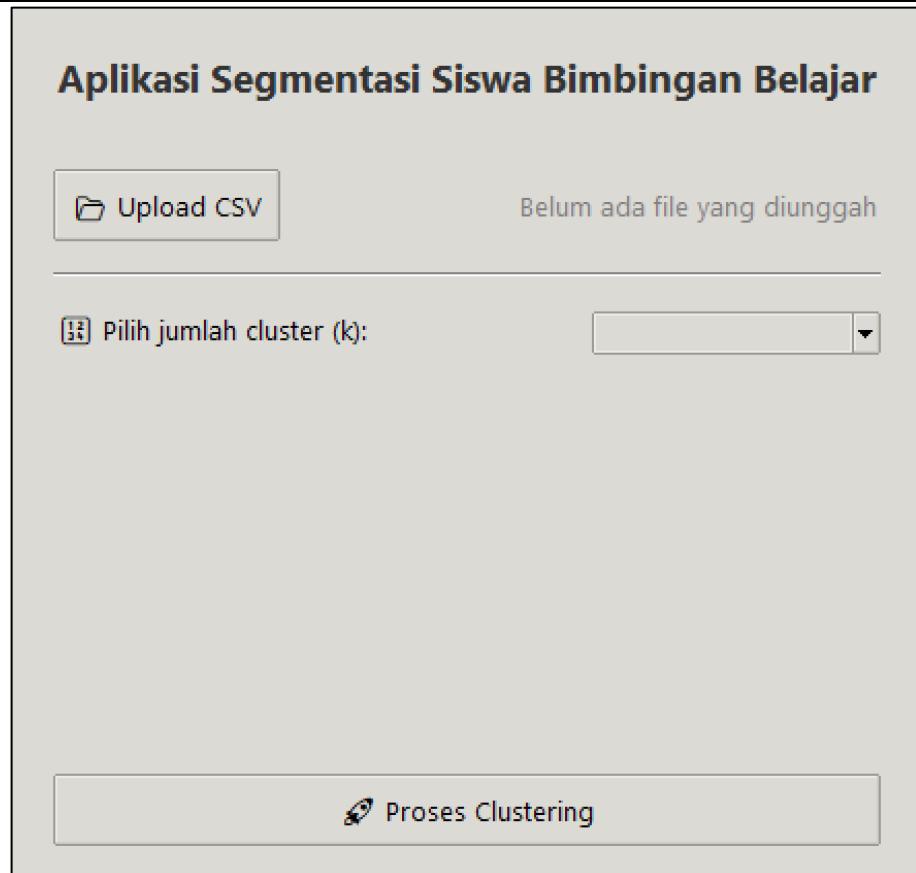
<b>Id</b>	<b>active</b>	<b>reflective</b>	<b>sense</b>	<b>intuitive</b>	<b>visual</b>	<b>verbal</b>	<b>sequential</b>	<b>global</b>
S1	0.250	0.750	0.444	0.556	1.000	0.000	0.222	0.778
S2	0.000	1.000	0.778	0.222	0.375	0.625	0.444	0.556
S3	0.375	0.625	0.667	0.333	0.000	1.000	0.556	0.444
S4	0.125	0.875	0.667	0.333	1.000	0.000	0.667	0.333
S5	0.875	0.125	0.667	0.333	0.125	0.875	0.333	0.667
S6	0.625	0.375	0.444	0.556	0.750	0.250	0.778	0.222
S7	0.250	0.750	0.556	0.444	0.500	0.500	0.556	0.444
S8	0.250	0.750	0.444	0.556	1.000	0.000	0.333	0.667
S9	0.375	0.625	0.667	0.333	0.375	0.625	0.667	0.333
S10	0.500	0.500	1.000	0.000	1.000	0.000	0.444	0.556
...	...	...	...	...	...	...	...	...
S51	0.500	0.500	0.222	0.778	0.625	0.375	0.556	0.444
S52	0.125	0.875	0.889	0.111	0.625	0.375	0.556	0.444
S53	0.625	0.375	0.778	0.222	0.750	0.250	0.889	0.111
S54	0.750	0.250	0.222	0.778	0.500	0.500	0.333	0.667
S55	0.625	0.375	0.889	0.111	0.750	0.250	0.444	0.556
S56	1.000	0.000	0.556	0.444	0.625	0.375	0.667	0.333
S57	0.375	0.625	0.556	0.444	0.750	0.250	0.333	0.667
S58	0.250	0.750	0.778	0.222	0.500	0.500	0.556	0.444

Setelah data melalui proses normalisasi, data tersebut akan diolah menggunakan Algoritma K-Means *Clustering* pada beberapa skenario uji coba yang diterapkan sebelumnya.

#### 4.1.2 K-Means Clustering

Data yang telah melalui proses normalisasi kemudian diolah menggunakan Algoritma K-Means *Clustering*. Implementasi Algoritma K-Means *Clustering*

dilakukan menggunakan Bahasa Pemrograman *Python*. Berikut merupakan implementasi sistem pada data.



Gambar 4.1. Tampilan Awal Sistem

Sistem dimulai dengan mengunggah *dataset* yang akan dikelompokkan sesuai dengan jumlah  $k$  yang dipilih. Setelah *dataset* dan jumlah  $k$  dipilih, proses *clustering* akan dimulai dengan menekan tombol proses *clustering*. Berikut merupakan sistem setelah proses *clustering* dilakukan yang disajikan pada Gambar 4.2.

Hasil Clustering										
Jumlah Cluster (k) : 3										
Jumlah Iterasi : 5 iterasi										
Silhouette Score : 0.2443461171295426										
----- Persebaran Data -----										
id	active	reflective	sense	intuitive	visual	verbal	sequential	global	cluster	
s1	5	6	6	5	10	1	3	8	1	
s2	3	8	9	2	5	6	5	6	2	
s3	6	5	8	3	2	9	6	5	2	
s4	4	7	8	3	10	1	7	4	1	
s5	10	1	8	3	3	8	4	7	0	
s6	8	3	6	5	8	3	8	3	0	
s7	5	6	7	4	6	5	6	5	2	
s8	5	6	6	5	10	1	4	7	1	
s9	6	5	8	3	5	6	7	4	2	
s10	7	4	11	0	10	1	5	6	1	
s11	4	7	6	5	3	8	5	6	2	
s12	6	5	7	4	9	2	8	3	1	
s13	8	3	8	3	4	7	1	10	2	
s14	5	6	7	4	8	3	6	5	1	
s15	10	1	9	2	7	4	9	2	0	
s16	5	6	10	1	5	6	5	6	2	
s17	7	4	6	5	4	7	5	6	2	
s18	8	3	9	2	8	3	7	4	0	
s19	6	5	5	6	8	3	4	7	1	
s20	5	6	9	2	8	3	8	3	1	
s21	10	1	9	2	8	3	6	5	0	
s22	9	2	9	2	9	2	3	8	1	

Gambar 4.2. Tampilan Hasil *Clustering*

Sistem akan menampilkan *window* berisi hasil *clustering* data siswa yang diunggah sebelumnya. Tampilan ini berisi persebaran data siswa setelah *clustering* beserta nilai dari *Silhouette Score* yang didapatkan. Sebagai contoh, pada gambar diatas, data siswa berhasil dikelompokkan menjadi 3 *cluster* dengan 5 iterasi yang memiliki nilai *silhouette* 0.244. Selain hasil *clustering* siswa, sistem juga menampilkan analisa pada setiap *cluster* yang disajikan pada Gambar 4.3.

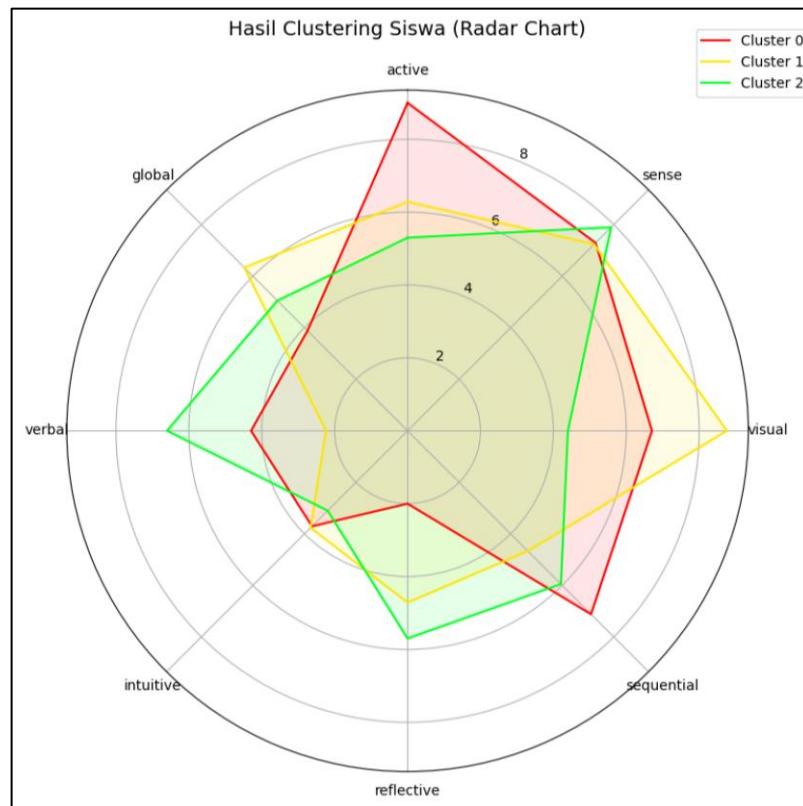
Hasil Clustering											
s37	8	3	5	6	8	3	5	6	1		
s38	5	6	7	4	2	9	6	5	2		
s39	10	1	7	4	8	3	8	3	0		
s40	3	8	9	2	7	4	8	3	2		
s41	6	5	8	3	6	5	8	3	2		
s42	6	5	6	5	9	2	6	5	1		
s43	6	5	10	1	3	8	9	2	2		
s44	9	2	4	7	8	3	6	5	0		
s45	7	4	7	4	2	9	8	3	2		
s46	6	5	9	2	3	8	6	5	2		
s47	3	8	5	6	3	8	7	4	2		
s48	5	6	6	5	9	2	2	9	1		
s49	10	1	8	3	6	5	9	2	0		
s50	8	3	7	4	9	2	3	8	1		
s51	7	4	4	7	7	4	6	5	1		
s52	4	7	10	1	7	4	6	5	2		
s53	8	3	9	2	8	3	9	2	0		
s54	9	2	4	7	6	5	4	7	0		
s55	8	3	10	1	8	3	5	6	1		
s56	11	0	7	4	7	4	7	4	0		
s57	6	5	7	4	8	3	4	7	1		
s58	5	6	9	2	6	5	6	5	2		
----- Analisa Cluster -----											
cluster	active	reflective	sense	intuitive	visual	verbal	sequential	global	jumlah		
0	9.000000	2.000000	7.294118	3.705882	6.705882	4.294118	7.117647	3.882353	17		
1	6.285714	4.714286	7.238095	3.761905	8.761905	2.238095	4.666667	6.333333	21		
2	5.300000	5.700000	7.900000	3.100000	4.400000	6.600000	5.950000	5.050000	20		

Gambar 4.3. Analisa Cluster

Analisa *cluster* berisi rata – rata dari setiap variabel yang digunakan pada *clustering* serta jumlah anggota dari masing – masing *cluster*. Sebagai contoh, pada gambar diatas, data siswa dikelompokkan menjadi 3 *cluster*. Sistem akan menampilkan analisa cluster dari masing – masing *cluster* yang ada. *Cluster* 0 memiliki nilai yang dominan pada variabel *active* dengan rata – rata 9.0. Hal yang sama juga ditunjukkan pada variabel lain, seperti variabel *sense* dengan rata – rata 7.29, variabel *visual* dengan rata – rata 6.70, dan variabel *sequential* dengan rata – rata 7.11.

Analisa ini juga menampilkan jumlah anggota yang tergolong pada *cluster* tersebut. Pada gambar diatas, ditunjukkan bahwa *cluster* 0 terdiri dari 17 siswa, *cluster* 1 terdiri dari 21 siswa, serta *cluster* 2 terdiri dari 20 siswa. Selain persebaran

data, hasil clustering juga akan ditampilkan melalui visualisasi *radar chart* yang disajikan pada Gambar 4.4.



Gambar 4.4. Visualisasi *Radar Chart* Hasil *Clustering*

Visualisasi tersebut menampilkan karakteristik yang dimiliki oleh masing – masing cluster. Pada penelitian ini, akan dilakukan serangkaian percobaan pada sistem sesuai dengan skenario uji coba yang tertera pada Tabel 3.12.

#### 4.1.3 Skenario Uji Coba

Terdapat 10 skenario percobaan yang akan dilakukan pada penelitian ini. Skenario tersebut memiliki perbedaan jumlah  $k$  yang akan digunakan pada proses *clustering*.

### a. Skenario 1

Pada skenario 1, proses *clustering* akan menggunakan jumlah  $k$  dengan nilai 2 sebagai nilai  $k$  inisial dari *clustering*. Pada skenario ini, data akan dikelompokkan menjadi 2 klaster. Berikut merupakan hasil *clustering* pada skenario 1 yang disajikan pada Gambar 4.5.

Hasil Clustering										
----- Persebaran Data -----										
id	active	reflective	sense	intuitive	visual	verbal	sequential	global	cluster	
s1	5	6	6	5	10	1	3	8	1	
s2	3	8	9	2	5	6	5	6	0	
s3	6	5	8	3	2	9	6	5	0	
s4	4	7	8	3	10	1	7	4	1	
s5	10	1	8	3	3	8	4	7	0	
s6	8	3	6	5	8	3	8	3	0	
s7	5	6	7	4	6	5	6	5	0	
s8	5	6	6	5	10	1	4	7	1	
s9	6	5	8	3	5	6	7	4	0	
s10	7	4	11	0	10	1	5	6	1	
s11	4	7	6	5	3	8	5	6	0	
s12	6	5	7	4	9	2	8	3	1	
s13	8	3	8	3	4	7	1	10	0	
s14	5	6	7	4	8	3	6	5	1	
s15	10	1	9	2	7	4	9	2	0	
s16	5	6	10	1	5	6	5	6	0	
s17	7	4	6	5	4	7	5	6	0	
s18	8	3	9	2	8	3	7	4	0	
s19	6	5	5	6	8	3	4	7	1	
s20	5	6	9	2	8	3	8	3	0	
s21	10	1	9	2	8	3	6	5	0	
s22	9	2	9	2	9	2	3	8	1	

Gambar 4.5. Hasil *Clustering* Skenario 1

Pada skenario 1, terdapat 2 klaster yang keluar sebagai hasil *clustering*. Proses ini berjalan sebanyak 2 iterasi dengan nilai *Silhouette Score* sebesar 0.1922551004321032. Nilai *Silhouette Score* didapatkan dengan cara menghitung jarak rata – rata antara titik data pada klaster yang sama dan klaster tetangga. Persebaran data pada setiap klaster disajikan pada Gambar 4.6.

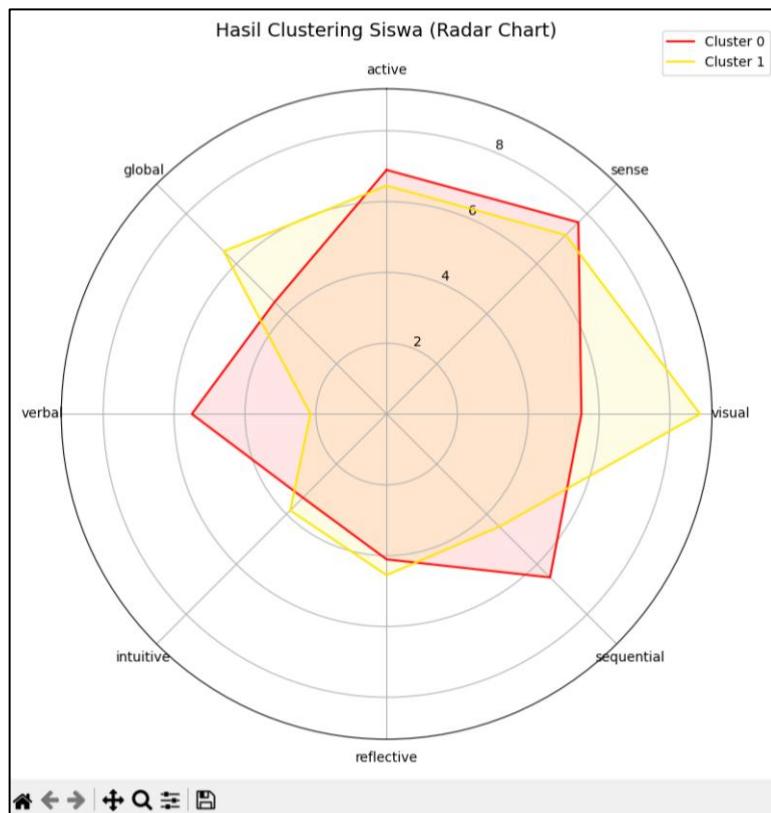
Hasil Clustering												
s36	6	5	8	3	7	4	10	1	0			
s37	8	3	5	6	8	3	5	6	1			
s38	5	6	7	4	2	9	6	5	0			
s39	10	1	7	4	8	3	8	3	0			
s40	3	8	9	2	7	4	8	3	0			
s41	6	5	8	3	6	5	8	3	0			
s42	6	5	6	5	9	2	6	5	1			
s43	6	5	10	1	3	8	9	2	0			
s44	9	2	4	7	8	3	6	5	1			
s45	7	4	7	4	2	9	8	3	0			
s46	6	5	9	2	3	8	6	5	0			
s47	3	8	5	6	3	8	7	4	0			
s48	5	6	6	5	9	2	2	9	1			
s49	10	1	8	3	6	5	9	2	0			
s50	8	3	7	4	9	2	3	8	1			
s51	7	4	4	7	7	4	6	5	0			
s52	4	7	10	1	7	4	6	5	0			
s53	8	3	9	2	8	3	9	2	0			
s54	9	2	4	7	6	5	4	7	0			
s55	8	3	10	1	8	3	5	6	1			
s56	11	0	7	4	7	4	7	4	0			
s57	6	5	7	4	8	3	4	7	1			
s58	5	6	9	2	6	5	6	5	0			
----- Analisa Cluster -----												
cluster	active	reflective	sense	intuitive	visual	verbal	sequential	global	jumlah			
0	6.894737	4.105263	7.657895	3.342105	5.50	5.50	6.526316	4.473684	38			
1	6.450000	4.550000	7.150000	3.850000	8.85	2.15	4.500000	6.500000	20			

Gambar 4.6. Analisa Cluster Skenario 1

Dari *clustering* yang dilakukan pada skenario 1, didapatkan hasil berupa 2 klaster. Klaster 0 terdiri dari 38 siswa dengan rata – rata yang cukup tinggi pada atribut *sense* dengan nilai 7.65, atribut *active* dengan nilai 6.89, atribut *sequential* dengan nilai 6.52, serta rata – rata yang seimbang pada atribut *visual* dan *verbal* dengan nilai 5.50.

Sedangkan klaster 1 terdiri dari 20 siswa dengan rata – rata yang cukup tinggi pada atribut *visual* dengan nilai 8.85, atribut *active* dengan nilai 6.45, atribut *sense* dengan nilai 7.15, serta atribut *global* dengan nilai 6.5.

Berikut merupakan visualisasi dari *clustering* pada skenario 1 dalam bentuk *radar plot* perbandingan dari masing masing kategori pada seluruh dimensi pada skenario 1 yang disajikan pada Gambar 4.7.



Gambar 4.7. Visualisasi Skenario 1

Dari visualisasi tersebut, ditunjukkan bahwa klaster 0 memiliki skor rata – rata bernilai 7.65 pada atribut *sense*, yang menunjukkan kecenderungan pada gaya belajar *sense* yang lebih cenderung menyukai data yang sesuai dengan fakta. Sedangkan klaster 1 memiliki skor rata – rata 8.85 pada atribut *visual* yang menunjukkan kecenderungan tinggi pada gaya belajar *visual* dimana siswa lebih baik dalam mengingat gambar, diagram dan penjelasan dalam bentuk visualisasi.

### b. Skenario 2

Pada skenario 2, proses *clustering* akan menggunakan jumlah  $k$  dengan nilai 3 sebagai nilai  $k$  inisial dari *clustering*. Pada skenario ini, data akan dikelompokkan

menjadi 3 klaster. Berikut merupakan hasil *clustering* pada skenario 2 yang disajikan pada Gambar 4.8.

Hasil Clustering										
Persebaran Data										
id	active	reflective	sense	intuitive	visual	verbal	sequential	global	cluster	
s1	5	6	6	5	10	1	3	8	1	
s2	3	8	9	2	5	6	5	6	2	
s3	6	5	8	3	2	9	6	5	2	
s4	4	7	8	3	10	1	7	4	1	
s5	10	1	8	3	3	8	4	7	0	
s6	8	3	6	5	8	3	8	3	0	
s7	5	6	7	4	6	5	6	5	2	
s8	5	6	6	5	10	1	4	7	1	
s9	6	5	8	3	5	6	7	4	2	
s10	7	4	11	0	10	1	5	6	1	
s11	4	7	6	5	3	8	5	6	2	
s12	6	5	7	4	9	2	8	3	1	
s13	8	3	8	3	4	7	1	10	2	
s14	5	6	7	4	8	3	6	5	1	
s15	10	1	9	2	7	4	9	2	0	
s16	5	6	10	1	5	6	5	6	2	
s17	7	4	6	5	4	7	5	6	2	
s18	8	3	9	2	8	3	7	4	0	
s19	6	5	5	6	8	3	4	7	1	
s20	5	6	9	2	8	3	8	3	1	
s21	10	1	9	2	8	3	6	5	0	
s22	9	2	9	2	9	2	3	8	1	

Gambar 4.8. Hasil *Clustering* Skenario 2

Pada skenario 2, terdapat 3 klaster yang keluar sebagai hasil *clustering*. Proses ini berjalan sebanyak 3 iterasi dengan nilai *Silhouette Score* sebesar 0.2443461171295426. Nilai *Silhouette Score* didapatkan dengan cara menghitung jarak rata – rata antara titik data pada klaster yang sama dan klaster tetangga. Persebaran data pada setiap klaster disajikan pada Gambar 4.9.

Hasil Clustering										
s37	8	3	5	6	8	3	5	6	1	
s38	5	6	7	4	2	9	6	5	2	
s39	10	1	7	4	8	3	8	3	0	
s40	3	8	9	2	7	4	8	3	2	
s41	6	5	8	3	6	5	8	3	2	
s42	6	5	6	5	9	2	6	5	1	
s43	6	5	10	1	3	8	9	2	2	
s44	9	2	4	7	8	3	6	5	0	
s45	7	4	7	4	2	9	8	3	2	
s46	6	5	9	2	3	8	6	5	2	
s47	3	8	5	6	3	8	7	4	2	
s48	5	6	6	5	9	2	2	9	1	
s49	10	1	8	3	6	5	9	2	0	
s50	8	3	7	4	9	2	3	8	1	
s51	7	4	4	7	7	4	6	5	1	
s52	4	7	10	1	7	4	6	5	2	
s53	8	3	9	2	8	3	9	2	0	
s54	9	2	4	7	6	5	4	7	0	
s55	8	3	10	1	8	3	5	6	1	
s56	11	0	7	4	7	4	7	4	0	
s57	6	5	7	4	8	3	4	7	1	
s58	5	6	9	2	6	5	6	5	2	
----- Analisa Cluster -----										
cluster	active	reflective	sense	intuitive	visual	verbal	sequential	global	jumlah	
0	9.000000	2.000000	7.294118	3.705882	6.705882	4.294118	7.117647	3.882353	17	
1	6.285714	4.714286	7.238095	3.761905	8.761905	2.238095	4.666667	6.333333	21	
2	5.300000	5.700000	7.900000	3.100000	4.400000	6.600000	5.950000	5.050000	20	

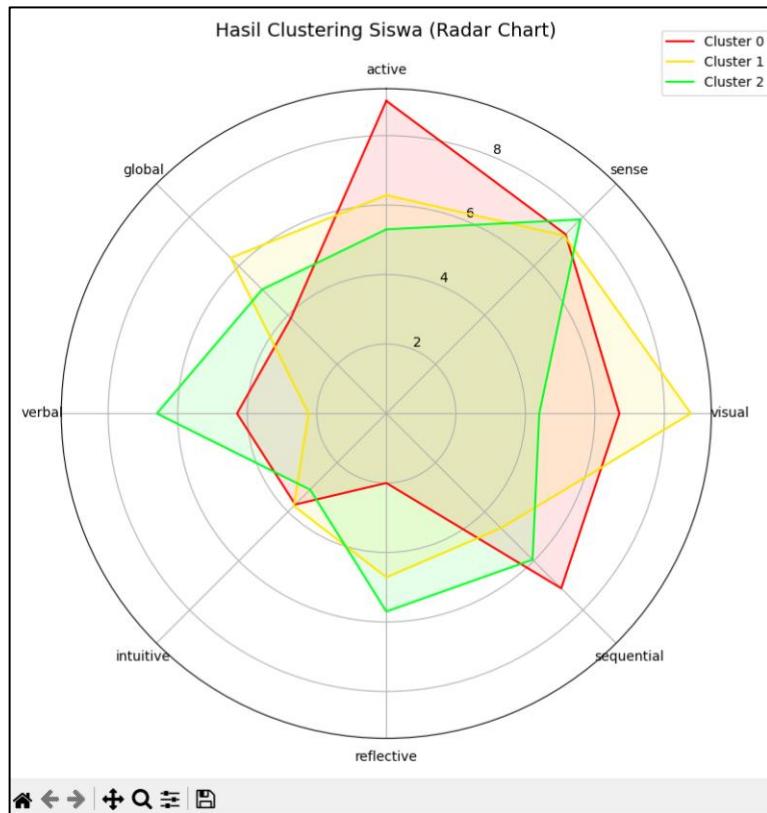
Exit

Gambar 4.9. Analisa Cluster Skenario 2

Dari *clustering* yang dilakukan pada skenario 2, didapatkan hasil berupa 3 klaster. Klaster 0 terdiri dari 17 siswa dengan rata – rata yang cukup tinggi pada atribut *active* dengan nilai 9.0, atribut *sense* dengan nilai 7.29, atribut *visual* dengan nilai 6.70, serta atribut *sequential* dengan nilai 7.11. Klaster 1 terdiri dari 21 siswa dengan rata – rata yang cukup tinggi pada atribut *visual* dengan nilai 8.76 serta atribut *sense* dengan nilai 7.23.

Klaster 2 terdiri dari 20 siswa dengan rata – rata tinggi pada atribut *sense* dengan nilai 7.90, atribut *verbal* dengan nilai 6.60, serta rata – rata yang cenderung seimbang pada atribut *active* dan *reflective* serta atribut *global* dan *sequential*.

Berikut merupakan visualisasi dari *clustering* pada skenario 2 dalam bentuk *radar plot* perbandingan dari masing masing kategori pada seluruh dimensi pada skenario 2 yang disajikan pada Gambar 4.10.



Gambar 4.10. Visualisasi Skenario 2

Dari visualisasi tersebut, ditunjukkan bahwa klaster 0 memiliki rata – rata 9.0 pada atribut *active*, yang menunjukkan kecenderungan tinggi pada gaya belajar aktif dimana siswa lebih menyukai belajar bersama kelompok belajar dan terlibat secara langsung. Sedangkan klaster 1 memiliki rata – rata 8.76 pada atribut *visual* yang menunjukkan kecenderungan tinggi pada gaya belajar *visual* dimana siswa lebih baik dalam mengingat gambar, diagram dan penjelasan dalam bentuk visualisasi. Klaster 2 memiliki rata – rata 7.9 pada atribut *sense*, yang menunjukkan kecenderungan pada gaya belajar *sense* dimana siswa cenderung menyukai data yang sesuai dengan fakta.

### c. Skenario 3

Pada skenario 3, proses *clustering* akan menggunakan jumlah  $k$  dengan nilai 4 sebagai nilai  $k$  inisial dari *clustering*. Pada skenario ini, data akan dikelompokkan menjadi 4 klaster. Berikut merupakan hasil *clustering* pada skenario 3 yang disajikan pada Gambar 4.11.

Hasil Clustering										
----- Persebaran Data -----										
id	active	reflective	sense	intuitive	visual	verbal	sequential	global	cluster	
s1	5	6	6	5	10	1	3	8	1	
s2	3	8	9	2	5	6	5	6	2	
s3	6	5	8	3	2	9	6	5	2	
s4	4	7	8	3	10	1	7	4	1	
s5	10	1	8	3	3	8	4	7	3	
s6	8	3	6	5	8	3	8	3	0	
s7	5	6	7	4	6	5	6	5	2	
s8	5	6	6	5	10	1	4	7	1	
s9	6	5	8	3	5	6	7	4	2	
s10	7	4	11	0	10	1	5	6	1	
s11	4	7	6	5	3	8	5	6	2	
s12	6	5	7	4	9	2	8	3	1	
s13	8	3	8	3	4	7	1	10	3	
s14	5	6	7	4	8	3	6	5	1	
s15	10	1	9	2	7	4	9	2	0	
s16	5	6	10	1	5	6	5	6	2	
s17	7	4	6	5	4	7	5	6	2	
s18	8	3	9	2	8	3	7	4	0	
s19	6	5	5	6	8	3	4	7	1	
s20	5	6	9	2	8	3	8	3	1	
s21	10	1	9	2	8	3	6	5	0	
s22	9	2	9	2	9	2	3	8	3	

Gambar 4.11. Hasil *Clustering* Skenario 3

Pada skenario 3, terdapat 4 klaster yang keluar sebagai hasil *clustering*. Proses ini berjalan sebanyak 5 iterasi dengan nilai *Silhouette Score* sebesar 0.2452330344105687. Nilai *Silhouette Score* didapatkan dengan cara menghitung jarak rata – rata antara titik data pada klaster yang sama dan klaster tetangga. Persebaran data pada setiap klaster disajikan pada Gambar 4.12.

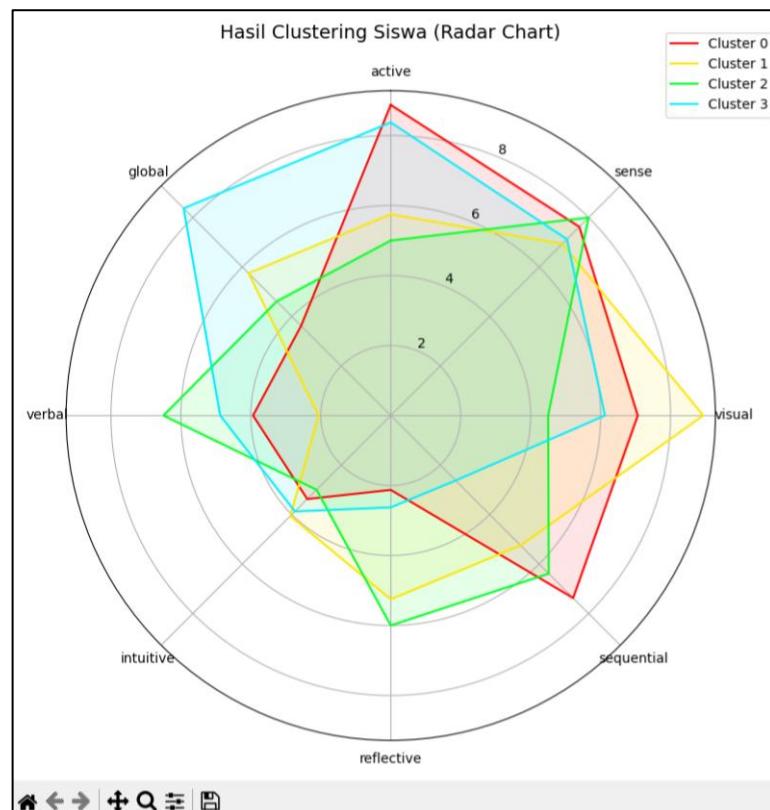
Hasil Clustering												
s38	5	6	7	4	2	9	6	5	2			
s39	10	1	7	4	8	3	8	3	0			
s40	3	8	9	2	7	4	8	3	2			
s41	6	5	8	3	6	5	8	3	2			
s42	6	5	6	5	9	2	6	5	1			
s43	6	5	10	1	3	8	9	2	2			
s44	9	2	4	7	8	3	6	5	0			
s45	7	4	7	4	2	9	8	3	2			
s46	6	5	9	2	3	8	6	5	2			
s47	3	8	5	6	3	8	7	4	2			
s48	5	6	6	5	9	2	2	9	1			
s49	10	1	8	3	6	5	9	2	0			
s50	8	3	7	4	9	2	3	8	3			
s51	7	4	4	7	7	4	6	5	1			
s52	4	7	10	1	7	4	6	5	2			
s53	8	3	9	2	8	3	9	2	0			
s54	9	2	4	7	6	5	4	7	3			
s55	8	3	10	1	8	3	5	6	0			
s56	11	0	7	4	7	4	7	4	0			
s57	6	5	7	4	8	3	4	7	1			
s58	5	6	9	2	6	5	6	5	2			
----- Analisa Cluster -----												
cluster	active	reflective	sense	intuitive	visual	verbal	sequential	global	jumlah			
0	8.875	2.125	7.6250	3.3750	7.0625	3.9375	7.375000	3.625000	16			
1	5.750	5.250	6.9375	4.0625	8.9375	2.0625	5.250000	5.750000	16			
2	5.000	6.000	8.0000	3.0000	4.5000	6.5000	6.388889	4.611111	18			
3	8.375	2.625	7.1250	3.8750	6.1250	4.8750	2.625000	8.375000	8			

Gambar 4.12. Analisa Cluster Skenario 3

Dari *clustering* yang dilakukan pada skenario 3, didapatkan hasil berupa 4 klaster. Klaster 0 terdiri dari 16 siswa dengan rata – rata yang cukup tinggi pada atribut *active* dengan nilai 8.87. Klaster 1 juga terdiri dari 16 siswa yang memiliki rata – rata yang cukup tinggi pada atribut *visual* dengan nilai 8.93, serta rata – rata yang cenderung seimbang pada atribut *active* dan *reflective* serta atribut *global* dan *sequential*.

Klaster 2 merupakan klaster dengan anggota paling banyak dibandingkan klaster lainnya, terdiri dari 18 siswa yang memiliki rata – rata yang cukup tinggi pada atribut *sense* dengan nilai 8.0. Sedangkan klaster 3 merupakan klaster terkecil yang terdiri dari 8 siswa dengan rata – rata yang cukup tinggi pada atribut *active* dengan nilai 8.37.

Berikut merupakan visualisasi dari *clustering* pada skenario 3 dalam bentuk *radar plot* perbandingan dari masing masing kategori pada seluruh dimensi pada skenario 3 yang disajikan pada Gambar 4.13.



Gambar 4.13. Visualisasi Skenario 3

Dari visualisasi tersebut, ditunjukkan bahwa atribut *global* memiliki nilai rata – rata 8.37 pada klaster 3 yang menunjukkan kecenderungan tinggi pada gaya belajar *global*, dimana siswa cenderung belajar secara bebas tanpa urutan. Atribut *active* memiliki rata – rata 8.87 pada klaster 0 dan 8.37 pada klaster 3, yang menunjukkan kecenderungan tinggi pada gaya belajar *active*. Klaster 1 memiliki rata – rata 8.93 pada atribut *visual*, yang menunjukkan kecenderungan tinggi pada

gaya belajar *visual*. Sedangkan klaster 2 memiliki rata – rata 8.0 pada atribut *sense* yang menunjukkan kecenderungan pada gaya belajar *sense*.

#### d. Skenario 4

Pada skenario 4, proses *clustering* akan menggunakan jumlah  $k$  dengan nilai 5 sebagai nilai  $k$  inisial dari *clustering*. Pada skenario ini, data akan dikelompokkan menjadi 5 klaster. Berikut merupakan hasil *clustering* pada skenario 4 yang disajikan pada Gambar 4.14.

Hasil Clustering											
===== Hasil Clustering =====											
Jumlah Cluster (k) : 5 Jumlah Iterasi : 5 iterasi Silhouette Score : 0.19880873001579144											
----- Persebaran Data -----											
id active reflective sense intuitive visual verbal sequential global cluster											
s1	5	6	6	5	10	1	3	8	1		
s2	3	8	9	2	5	6	5	6	2		
s3	6	5	8	3	2	9	6	5	2		
s4	4	7	8	3	10	1	7	4	1		
s5	10	1	8	3	3	8	4	7	3		
s6	8	3	6	5	8	3	8	3	4		
s7	5	6	7	4	6	5	6	5	2		
s8	5	6	6	5	10	1	4	7	1		
s9	6	5	8	3	5	6	7	4	0		
s10	7	4	11	0	10	1	5	6	1		
s11	4	7	6	5	3	8	5	6	2		
s12	6	5	7	4	9	2	8	3	1		
s13	8	3	8	3	4	7	1	10	3		
s14	5	6	7	4	8	3	6	5	1		
s15	10	1	9	2	7	4	9	2	4		
s16	5	6	10	1	5	6	5	6	2		
s17	7	4	6	5	4	7	5	6	0		
s18	8	3	9	2	8	3	7	4	4		
s19	6	5	5	6	8	3	4	7	1		
s20	5	6	9	2	8	3	8	3	1		
s21	10	1	9	2	8	3	6	5	4		
s22	9	2	9	2	9	2	3	8	3		

Gambar 4.14. Hasil *Clustering* Skenario 4

Pada skenario 4, terdapat 5 klaster yang keluar sebagai hasil *clustering*. Proses ini berjalan sebanyak 5 iterasi dengan nilai *Silhouette Score* sebesar 0.19880873001579144. Nilai *Silhouette Score* didapatkan dengan cara menghitung jarak rata – rata antara titik data pada klaster yang sama dan klaster tetangga. Persebaran data pada setiap klaster disajikan pada Gambar 4.15.

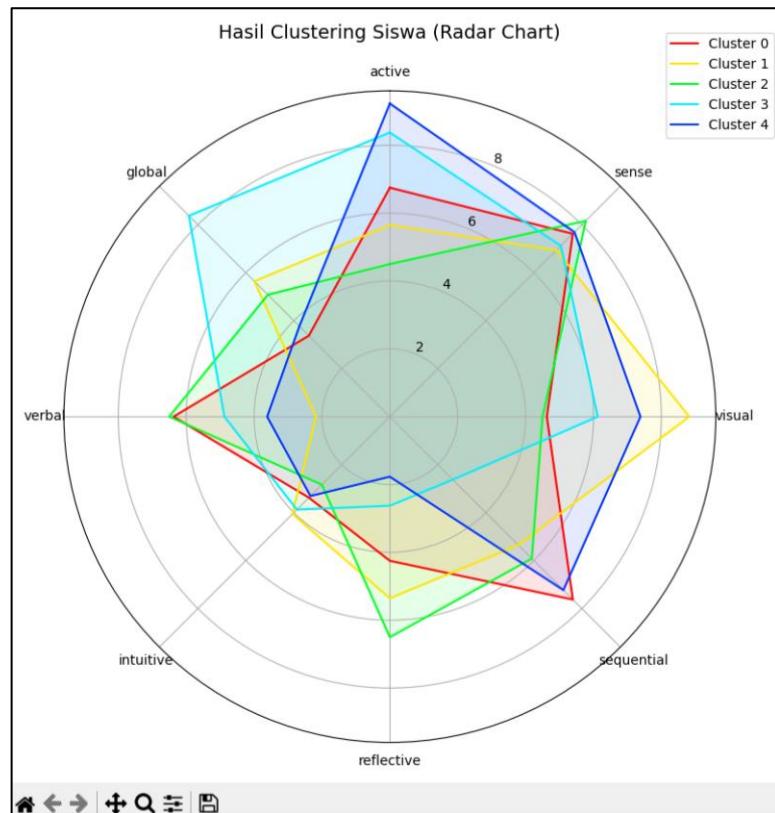
Hasil Clustering											
s39	10	1	7	4	8	3	8	3	4		
s40	3	8	9	2	7	4	8	3	2		
s41	6	5	8	3	6	5	8	3	0		
s42	6	5	6	5	9	2	6	5	1		
s43	6	5	10	1	3	8	9	2	0		
s44	9	2	4	7	8	3	6	5	4		
s45	7	4	7	4	2	9	8	3	0		
s46	6	5	9	2	3	8	6	5	2		
s47	3	8	5	6	3	8	7	4	2		
s48	5	6	6	5	9	2	2	9	1		
s49	10	1	8	3	6	5	9	2	4		
s50	8	3	7	4	9	2	3	8	3		
s51	7	4	4	7	7	4	6	5	1		
s52	4	7	10	1	7	4	6	5	2		
s53	8	3	9	2	8	3	9	2	4		
s54	9	2	4	7	6	5	4	7	3		
s55	8	3	10	1	8	3	5	6	4		
s56	11	0	7	4	7	4	7	4	4		
s57	6	5	7	4	8	3	4	7	1		
s58	5	6	9	2	6	5	6	5	2		
----- Analisa Cluster -----											
cluster	active	reflective	sense	intuitive	visual	verbal	sequential	global	jumlah		
0	6.750000	4.250000	7.625000	3.375000	4.625000	6.375000	7.625000	3.375000	8		
1	5.647059	5.352941	6.941176	4.058824	8.823529	2.176471	5.352941	5.647059	17		
2	4.500000	6.500000	8.166667	2.833333	4.500000	6.500000	5.916667	5.083333	12		
3	8.375000	2.625000	7.125000	3.875000	6.125000	4.875000	2.625000	8.375000	8		
4	9.230769	1.769231	7.692308	3.307692	7.384615	3.615385	7.230769	3.769231	13		

Gambar 4.15. Analisa Cluster Skenario 4

Dari *clustering* yang dilakukan pada skenario 4, didapatkan hasil berupa 5 klaster. Klaster 0 dan klaster 3 merupakan klaster terkecil yang terdiri dari 8 siswa yang memiliki rata – rata cukup tinggi pada atribut *sequential* dengan nilai 7.62 pada klaster 0, serta atribut *active* dengan nilai 8.37 pada klaster 3. Sedangkan klaster 1 merupakan klaster terbesar yang terdiri dari 17 siswa yang memiliki rata – rata cukup tinggi pada atribut *visual* dengan nilai 8.82.

Klaster 2 terdiri dari 12 siswa yang memiliki rata – rata cukup tinggi pada atribut *sense* dengan nilai 8.16, serta rata – rata yang cenderung seimbang pada atribut *global* dan *sequential*. Klaster 4 terdiri dari 13 siswa yang memiliki rata – rata cukup tinggi pada atribut *active* dengan nilai 9.23.

Berikut merupakan visualisasi dari *clustering* pada skenario 3 dalam bentuk *radar plot* perbandingan dari masing masing kategori pada seluruh dimensi pada skenario 3 yang disajikan pada Gambar 4.16.



Gambar 4.16. Visualisasi Skenario 4

Dari visualisasi tersebut, klaster 3 dan 4 memiliki rata – rata yang tinggi pada atribut *active* yang menunjukkan kecenderungan pada gaya belajar *active*. Klaster 1 memiliki rata – rata yang tinggi pada atribut *visual* yang menunjukkan kecenderungan pada gaya belajar *visual*, seperti melalui gambar atau diagram. Klaster 2 memiliki rata – rata yang tinggi pada atribut *sense* yang menunjukkan kecenderungan pada gaya belajar *sense*, dimana siswa cenderung menyukai data

yang sesuai dengan fakta. Sedangkan klaster 0 memiliki rata – rata yang tinggi pada atribut *sequential*, dimana siswa cenderung belajar sesuai dengan urutan yang ada.

#### e. Skenario 5

Pada skenario 5, proses *clustering* akan menggunakan jumlah  $k$  dengan nilai 6 sebagai nilai  $k$  inisial dari *clustering*. Pada skenario ini, data akan dikelompokkan menjadi 6 klaster. Berikut merupakan hasil *clustering* pada skenario 5 yang disajikan pada Gambar 4.17.

Hasil Clustering										
----- Persebaran Data -----										
id	active	reflective	sense	intuitive	visual	verbal	sequential	global	cluster	
s1	5	6	6	5	10	1	3	8	1	
s2	3	8	9	2	5	6	5	6	2	
s3	6	5	8	3	2	9	6	5	5	
s4	4	7	8	3	10	1	7	4	1	
s5	10	1	8	3	3	8	4	7	5	
s6	8	3	6	5	8	3	8	3	4	
s7	5	6	7	4	6	5	6	5	0	
s8	5	6	6	5	10	1	4	7	1	
s9	6	5	8	3	5	6	7	4	0	
s10	7	4	11	0	10	1	5	6	1	
s11	4	7	6	5	3	8	5	6	2	
s12	6	5	7	4	9	2	8	3	1	
s13	8	3	8	3	4	7	1	10	5	
s14	5	6	7	4	8	3	6	5	1	
s15	10	1	9	2	7	4	9	2	4	
s16	5	6	10	1	5	6	5	6	0	
s17	7	4	6	5	4	7	5	6	5	
s18	8	3	9	2	8	3	7	4	4	
s19	6	5	5	6	8	3	4	7	3	
s20	5	6	9	2	8	3	8	3	0	
s21	10	1	9	2	8	3	6	5	4	
s22	9	2	9	2	9	2	3	8	3	

Gambar 4.17. Hasil *Clustering* Skenario 5

Pada skenario 5, terdapat 6 klaster yang keluar sebagai hasil *clustering*. Proses ini berjalan sebanyak 5 iterasi dengan nilai *Silhouette Score* sebesar 0.2040150748088004. Nilai *Silhouette Score* didapatkan dengan cara menghitung jarak rata – rata antara titik data pada klaster yang sama dan klaster tetangga. Persebaran data pada setiap klaster disajikan pada Gambar 4.18.

Hasil Clustering											
s40	3	8	9	2	7	4	8	3	0		
s41	6	5	8	3	6	5	8	3	0		
s42	6	5	6	5	9	2	6	5	1		
s43	6	5	10	1	3	8	9	2	0		
s44	9	2	4	7	8	3	6	5	4		
s45	7	4	7	4	2	9	8	3	5		
s46	6	5	9	2	3	8	6	5	5		
s47	3	8	5	6	3	8	7	4	2		
s48	5	6	6	5	9	2	2	9	3		
s49	10	1	8	3	6	5	9	2	4		
s50	8	3	7	4	9	2	3	8	3		
s51	7	4	4	7	7	4	6	5	3		
s52	4	7	10	1	7	4	6	5	0		
s53	8	3	9	2	8	3	9	2	4		
s54	9	2	4	7	6	5	4	7	3		
s55	8	3	10	1	8	3	5	6	4		
s56	11	0	7	4	7	4	7	4	4		
s57	6	5	7	4	8	3	4	7	3		
s58	5	6	9	2	6	5	6	5	0		
----- Analisa Cluster -----											
cluster	active	reflective	sense	intuitive	visual	verbal	sequential	global	jumlah		
0	5.000000	6.000000	8.666667	2.333333	6.000000	5.000000	7.083333	3.916667	12		
1	5.444444	5.555556	7.444444	3.555556	9.555556	1.444444	5.777778	5.222222	9		
2	3.750000	7.250000	6.750000	4.250000	3.250000	7.750000	5.750000	5.250000	4		
3	7.181818	3.818182	6.363636	4.636364	8.000000	3.000000	3.363636	7.636364	11		
4	9.142857	1.857143	7.714286	3.285714	7.214286	3.785714	7.285714	3.714286	14		
5	7.500000	3.500000	7.250000	3.750000	3.250000	7.750000	4.875000	6.125000	8		

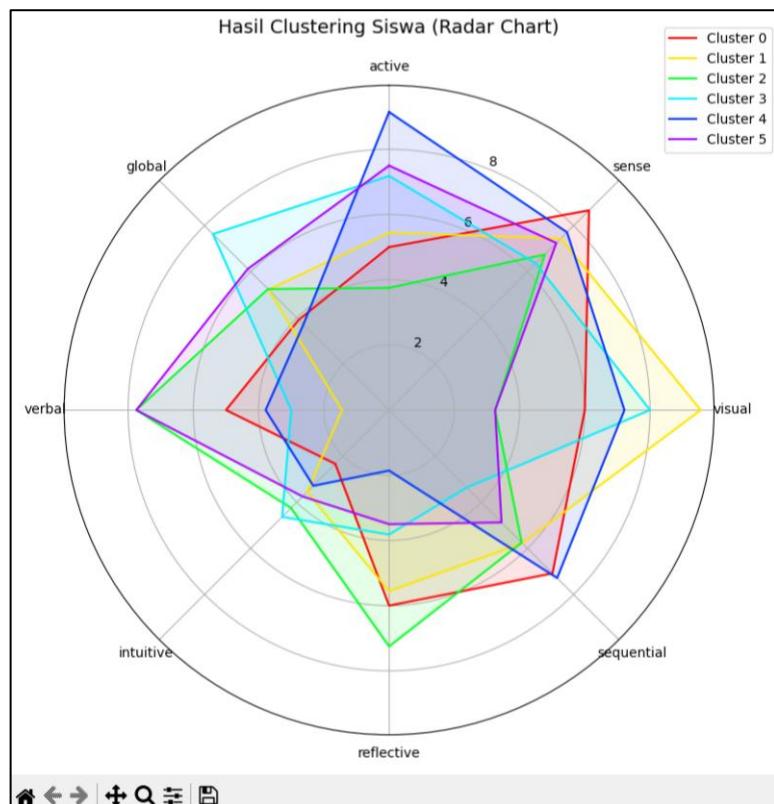
Gambar 4.18. Analisa Cluster Skenario 5

Dari *clustering* yang dilakukan pada skenario 5, didapatkan hasil berupa 6 klaster. Klaster 0 terdiri dari 12 siswa yang memiliki rata – rata yang cukup tinggi pada atribut *sense* dengan nilai 8.66. Klaster 1 terdiri dari 9 siswa yang memiliki rata – rata yang cukup tinggi pada atribut *visual* dengan nilai 9.55, serta rata – rata yang cenderung seimbang pada atribut *active* dan *reflective*.

Klaster 2 merupakan klaster terkecil yang terdiri dari 4 siswa, memiliki rata – rata yang cukup tinggi pada atribut *verbal* dengan nilai 7.75. Sedangkan klaster 4 merupakan klaster terbesar yang terdiri dari 14 siswa, memiliki rata – rata yang cukup tinggi pada atribut *active* dengan nilai 9.14.

Klaster 3 terdiri dari 11 siswa yang memiliki rata – rata yang cukup tinggi pada atribut *visual* dengan nilai 8.0. Klaster 5 terdiri dari 8 siswa yang memiliki rata – rata yang cukup tinggi pada atribut *verbal* dengan nilai 7.75.

Berikut merupakan visualisasi dari *clustering* pada skenario 5 dalam bentuk *radar plot* perbandingan dari masing masing kategori pada seluruh dimensi pada skenario 5 yang disajikan pada Gambar 4.19.



Gambar 4.19. Visualisasi Skenario 5

Dari visualisasi tersebut, klaster 0 memiliki rata – rata tinggi pada atribut *sense*, dimana siswa lebih menyukai belajar berdasarkan data dan fakta yang ada. Klaster 1 memiliki rata – rata tinggi pada atribut *visual*, dimana siswa lebih menyukai belajar menggunakan ilustrasi berupa gambar atau diagram. Klaster 2 memiliki rata – rata tinggi pada atribut *verbal* dan *reflective*, dimana siswa dengan gaya belajar *verbal* menunjukkan kecenderungan pada penjelasan secara deskriptif

dan siswa dengan gaya belajar *reflective* menunjukkan kecenderungan untuk melakukan diskusi secara mandiri.

Klaster 3 memiliki rata – rata tinggi pada atribut *global* dan *visual*, dimana siswa cenderung menyukai belajar secara bebas dan menggunakan berbagai ilustrasi seperti gambar atau diagram. Klaster 4 memiliki rata – rata tinggi pada atribut *active* dan *sequential*, dimana siswa cenderung menyukai belajar secara diskusi dengan kelompok dan sesuai dengan rangkaian materi yang ada. Sedangkan klaster 5 memiliki rata – rata tinggi pada atribut *active* dan *verbal*, dimana siswa cenderung menyukai belajar secara diskusi dengan kelompok dan menyukai penjelasan secara deskriptif.

#### **f. Skenario 6**

Pada skenario 6, proses *clustering* akan menggunakan jumlah  $k$  dengan nilai 7 sebagai nilai  $k$  inisial dari *clustering*. Pada skenario ini, data akan dikelompokkan menjadi 7 klaster. Berikut merupakan hasil *clustering* pada skenario 6 yang disajikan pada Gambar 4.20.

Hasil Clustering											
===== Hasil Clustering =====											
Jumlah Cluster (k) : 7											
Jumlah Iterasi : 4 iterasi											
Silhouette Score : 0.18933045635578755											
----- Persebaran Data -----											
id	active	reflective	sense	intuitive	visual	verbal	sequential	global	cluster		
s1	5	6	6	5	10	1	3	8	1		
s2	3	8	9	2	5	6	5	6	2		
s3	6	5	8	3	2	9	6	5	5		
s4	4	7	8	3	10	1	7	4	1		
s5	10	1	8	3	3	8	4	7	5		
s6	8	3	6	5	8	3	8	3	6		
s7	5	6	7	4	6	5	6	5	0		
s8	5	6	6	5	10	1	4	7	1		
s9	6	5	8	3	5	6	7	4	0		
s10	7	4	11	0	10	1	5	6	4		
s11	4	7	6	5	3	8	5	6	2		
s12	6	5	7	4	9	2	8	3	1		
s13	8	3	8	3	4	7	1	10	5		
s14	5	6	7	4	6	3	6	5	1		
s15	10	1	9	2	7	4	9	2	6		
s16	5	6	10	1	5	6	5	6	0		
s17	7	4	6	5	4	7	5	6	5		
s18	8	3	9	2	8	3	7	4	4		
s19	6	5	5	6	8	3	4	7	1		
s20	5	6	9	2	8	3	8	3	0		
s21	10	1	9	2	8	3	6	5	4		
s22	9	2	9	2	9	2	3	8	3		

Gambar 4.20. Hasil *Clustering* Skenario 6

Pada skenario 6, terdapat 7 klaster yang keluar sebagai hasil *clustering*.

Proses ini berjalan sebanyak 4 iterasi dengan nilai *Silhouette Score* sebesar 0.18933045635578755. Nilai *Silhouette Score* didapatkan dengan cara menghitung jarak rata – rata antara titik data pada klaster yang sama dan klaster tetangga. Persebaran data pada setiap klaster disajikan pada Gambar 4.21.

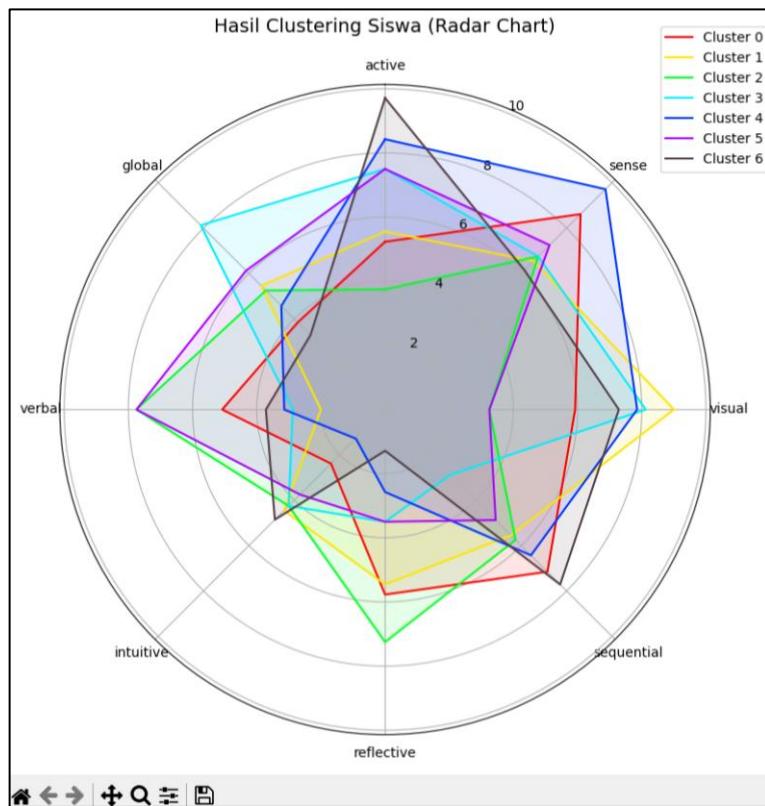
Hasil Clustering												
s41	6	5	8	3	6	5	8	3	0			
s42	6	5	6	5	9	2	6	5	1			
s43	6	5	10	1	3	8	9	2	0			
s44	9	2	4	7	8	3	6	5	6			
s45	7	4	7	4	2	9	8	3	5			
s46	6	5	9	2	3	8	6	5	5			
s47	3	8	5	6	3	8	7	4	2			
s48	5	6	6	5	9	2	2	9	3			
s49	10	1	8	3	6	5	9	2	6			
s50	8	3	7	4	9	2	3	8	3			
s51	7	4	4	7	7	4	6	5	1			
s52	4	7	10	1	7	4	6	5	0			
s53	8	3	9	2	8	3	9	2	4			
s54	9	2	4	7	6	5	4	7	3			
s55	8	3	10	1	8	3	5	6	4			
s56	11	0	7	4	7	4	7	4	6			
s57	6	5	7	4	8	3	4	7	1			
s58	5	6	9	2	6	5	6	5	0			
----- Analisa Cluster -----												
cluster	active	reflective	sense	intuitive	visual	verbal	sequential	global	jumlah			
0	5.230769	5.769231	8.615385	2.384615	5.923077	5.076923	7.153846	3.846154	13			
1	5.545455	5.454545	6.545455	4.454545	9.000000	2.000000	5.545455	5.454545	11			
2	3.750000	7.250000	6.750000	4.250000	3.250000	7.750000	5.750000	5.250000	4			
3	7.500000	3.500000	6.750000	4.250000	8.125000	2.875000	2.875000	8.125000	8			
4	8.428571	2.571429	9.714286	1.285714	7.857143	3.142857	6.428571	4.571429	7			
5	7.500000	3.500000	7.250000	3.750000	3.250000	7.750000	4.875000	6.125000	8			
6	9.714286	1.285714	6.142857	4.857143	7.285714	3.714286	7.714286	3.285714	7			

Gambar 4.21. Analisa Cluster Skenario 6

Dari *clustering* yang dilakukan pada skenario 6, didapatkan hasil berupa 7 klaster. Klaster 0 merupakan klaster terbesar yang terdiri dari 13 siswa, memiliki rata – rata yang cukup tinggi pada atribut *sense* dengan nilai 8.61. Sedangkan klaster 2 merupakan klaster terkecil yang terdiri dari 4 siswa, memiliki rata – rata yang cukup tinggi pada atribut *verbal* dengan nilai 7.75.

Klaster 1 merupakan klaster terbesar kedua yang terdiri dari 11 siswa, memiliki rata – rata yang cukup tinggi pada atribut *visual* dengan nilai 9.0, serta rata – rata yang cenderung seimbang pada atribut *active* dan *reflective* serta atribut *global* dan *sequential*.

Berikut merupakan visualisasi dari *clustering* pada skenario 6 dalam bentuk *radar plot* perbandingan dari masing masing kategori pada seluruh dimensi pada skenario 6 yang disajikan pada Gambar 4.22.



Gambar 4.22. Visualisasi Skenario 6

Dari visualisasi tersebut, pada klaster 3, 4, 5, dan 6 memiliki rata – rata tinggi pada atribut *active*, dimana siswa cenderung menyukai belajar dengan diskusi secara berkelompok. Klaster 0 dan 4 memiliki rata – rata tinggi pada atribut *sense*, dimana siswa cenderung menyukai belajar berdasarkan data dan fakta yang ada.

Klaster 2 dan 5 memiliki rata – rata tinggi pada atribut *verbal*, dimana siswa cenderung menyukai belajar dengan menggunakan penjelasan secara deskriptif. Sedangkan klaster 1, 3, dan 4 memiliki rata – rata tinggi pada atribut *visual*, dimana siswa cenderung menyukai belajar dengan penjelasan berupa ilustrasi seperti gambar atau diagram.

### g. Skenario 7

Pada skenario 7, proses *clustering* akan menggunakan jumlah  $k$  dengan nilai 8 sebagai nilai  $k$  inisial dari *clustering*. Pada skenario ini, data akan dikelompokkan menjadi 8 klaster. Berikut merupakan hasil *clustering* pada skenario 7 yang disajikan pada Gambar 4.23.

Hasil Clustering										
----- Persebaran Data -----										
id	active	reflective	sense	intuitive	visual	verbal	sequential	global	cluster	
s1	5	6	6	5	10	1	3	8	1	
s2	3	8	9	2	5	6	5	6	5	
s3	6	5	8	3	2	9	6	5	7	
s4	4	7	8	3	10	1	7	4	1	
s5	10	1	8	3	3	8	4	7	7	
s6	8	3	6	5	8	3	8	3	6	
s7	5	6	7	4	6	5	6	5	0	
s8	5	6	6	5	10	1	4	7	1	
s9	6	5	8	3	5	6	7	4	7	
s10	7	4	11	0	10	1	5	6	4	
s11	4	7	6	5	3	8	5	6	2	
s12	6	5	7	4	9	2	8	3	0	
s13	8	3	8	3	4	7	1	10	7	
s14	5	6	7	4	8	3	6	5	1	
s15	10	1	9	2	7	4	9	2	4	
s16	5	6	10	1	5	6	5	6	5	
s17	7	4	6	5	4	7	5	6	7	
s18	8	3	9	2	8	3	7	4	4	
s19	6	5	5	6	8	3	4	7	1	
s20	5	6	9	2	8	3	8	3	0	
s21	10	1	9	2	8	3	6	5	4	
s22	9	2	9	2	9	2	3	8	3	

Gambar 4.23. Hasil *Clustering* Skenario 7

Pada skenario 6, terdapat 7 klaster yang keluar sebagai hasil *clustering*. Proses ini berjalan sebanyak 7 iterasi dengan nilai *Silhouette Score* sebesar 0.19841597698077995. Nilai *Silhouette Score* didapatkan dengan cara menghitung jarak rata – rata antara titik data pada klaster yang sama dan klaster tetangga. Persebaran data pada setiap klaster disajikan pada Gambar 4.24.

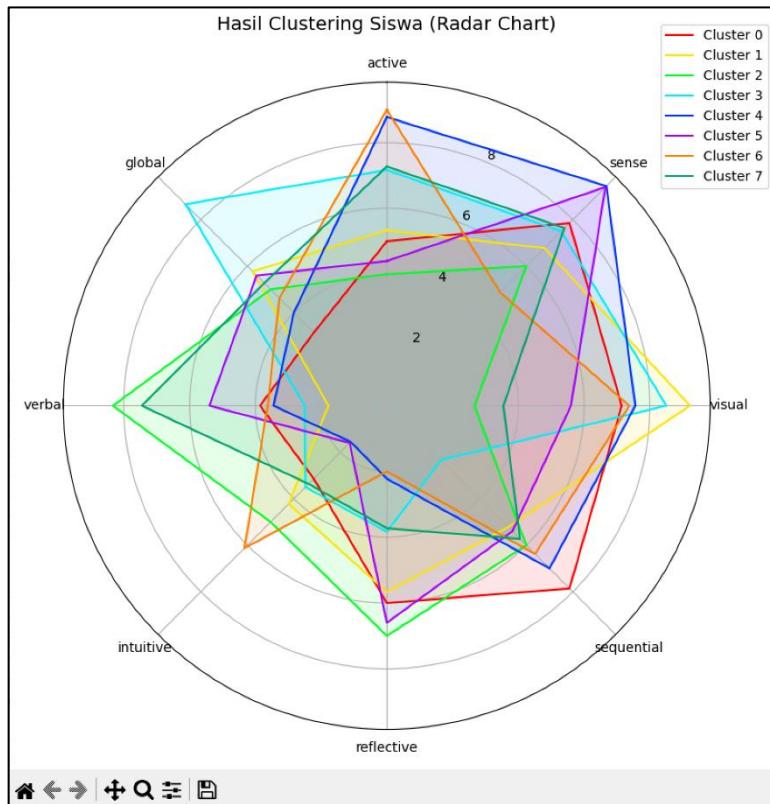
Hasil Clustering										
Analisa Cluster										
cluster	active	reflective	sense	intuitive	visual	verbal	sequential	global	jumlah	
0	5.000000	6.000000	7.857143	3.142857	7.142857	3.857143	7.857143	3.142857	7	
1	5.333333	5.666667	6.777778	4.222222	9.222222	1.777778	5.222222	5.777778	9	
2	4.000000	7.000000	6.000000	5.000000	2.666667	8.333333	6.000000	5.000000	3	
3	7.166667	3.833333	7.500000	3.500000	8.500000	2.500000	2.333333	8.666667	6	
4	8.777778	2.222222	9.444444	1.555556	7.555556	3.444444	7.000000	4.000000	9	
5	4.400000	6.600000	9.400000	1.600000	5.600000	5.400000	5.400000	5.600000	5	
6	9.000000	2.000000	4.875000	6.125000	7.375000	3.625000	6.375000	4.625000	8	
7	7.272727	3.727273	7.636364	3.363636	3.545455	7.454545	5.727273	5.272727	11	

Gambar 4.24. Analisa Cluster Skenario 7

Dari *clustering* yang dilakukan pada skenario 7, didapatkan hasil berupa 8 klaster. Klaster 7 merupakan klaster terbesar yang terdiri dari 11 siswa, memiliki rata – rata yang cukup tinggi pada atribut *sense* dengan nilai 7.63. Sedangkan klaster 2 merupakan klaster terkecil yang terdiri dari 3 siswa, memiliki rata – rata yang cukup tinggi pada atribut *verbal* dengan nilai 8.33.

Klaster 4 merupakan klaster terbesar kedua yang terdiri dari 9 siswa, memiliki rata – rata yang cukup tinggi pada atribut *sense* dengan nilai 9.4. Diikuti dengan klaster 1 yang terdiri dari 9 siswa juga, memiliki rata – rata yang cukup tinggi pada atribut *visual* dengan nilai 9.2.

Berikut merupakan visualisasi dari *clustering* pada skenario 7 dalam bentuk *radar plot* perbandingan dari masing masing kategori pada seluruh dimensi pada skenario 7 yang disajikan pada Gambar 4.25.



Gambar 4.25. Visualisasi Skenario 7

Dari visualisasi tersebut, klaster 0, 1, 2, 3, 4, 5 serta 7 menunjukkan rata – rata yang cukup tinggi pada atribut *sense*, dimana siswa cenderung menyukai belajar berdasarkan data dan fakta yang ada atau data yang bersifat faktual. Hal ini berbanding terbalik dengan klaster 6 yang memiliki rata – rata yang cukup tinggi pada atribut *intuitive*, dimana siswa lebih menyukai informasi yang bersifat teoritis.

Klaster 0, 1, 3, 4 serta 6 memiliki rata – rata yang cukup tinggi pada atribut *visual*, dimana siswa lebih menyukai informasi sensorik berbasis demonstrasi, seperti gambar atau diagram. Berlawanan dengan klaster 2 dan 7 yang memiliki rata – rata yang cukup tinggi pada atribut *verbal*, dimana siswa lebih menyukai informasi berupa penjelasan yang bersifat deskriptif.

### h. Skenario 8

Pada skenario 8, proses *clustering* akan menggunakan jumlah  $k$  dengan nilai 9 sebagai nilai  $k$  inisial dari *clustering*. Pada skenario ini, data akan dikelompokkan menjadi 9 klaster. Berikut merupakan hasil *clustering* pada skenario 8 yang disajikan pada Gambar 4.26.

Hasil Clustering										
----- Persebaran Data -----										
id	active	reflective	sense	intuitive	visual	verbal	sequential	global	cluster	
s1	5	6	6	5	10	1	3	8	1	
s2	3	8	9	2	5	6	5	6	5	
s3	6	5	8	3	2	9	6	5	7	
s4	4	7	8	3	10	1	7	4	1	
s5	10	1	8	3	3	8	4	7	7	
s6	8	3	6	5	8	3	8	3	4	
s7	5	6	7	4	6	5	6	5	0	
s8	5	6	6	5	10	1	4	7	1	
s9	6	5	8	3	5	6	7	4	7	
s10	7	4	11	0	10	1	5	6	8	
s11	4	7	6	5	3	8	5	6	2	
s12	6	5	7	4	9	2	8	3	0	
s13	8	3	8	3	4	7	1	10	7	
s14	5	6	7	4	8	3	6	5	1	
s15	10	1	9	2	7	4	9	2	6	
s16	5	6	10	1	5	6	5	6	5	
s17	7	4	6	5	4	7	5	6	7	
s18	8	3	9	2	8	3	7	4	8	
s19	6	5	5	6	8	3	4	7	1	
s20	5	6	9	2	8	3	8	3	0	
s21	10	1	9	2	8	3	6	5	6	
s22	9	2	9	2	9	2	3	8	8	

Gambar 4.26. Hasil *Clustering* Skenario 8

Pada skenario 8, terdapat 9 klaster yang keluar sebagai hasil *clustering*. Proses ini berjalan sebanyak 5 iterasi dengan nilai *Silhouette Score* sebesar 0.20213839213565232. Nilai *Silhouette Score* didapatkan dengan cara menghitung jarak rata – rata antara titik data pada klaster yang sama dan klaster tetangga. Persebaran data pada setiap klaster disajikan pada Gambar 4.27.

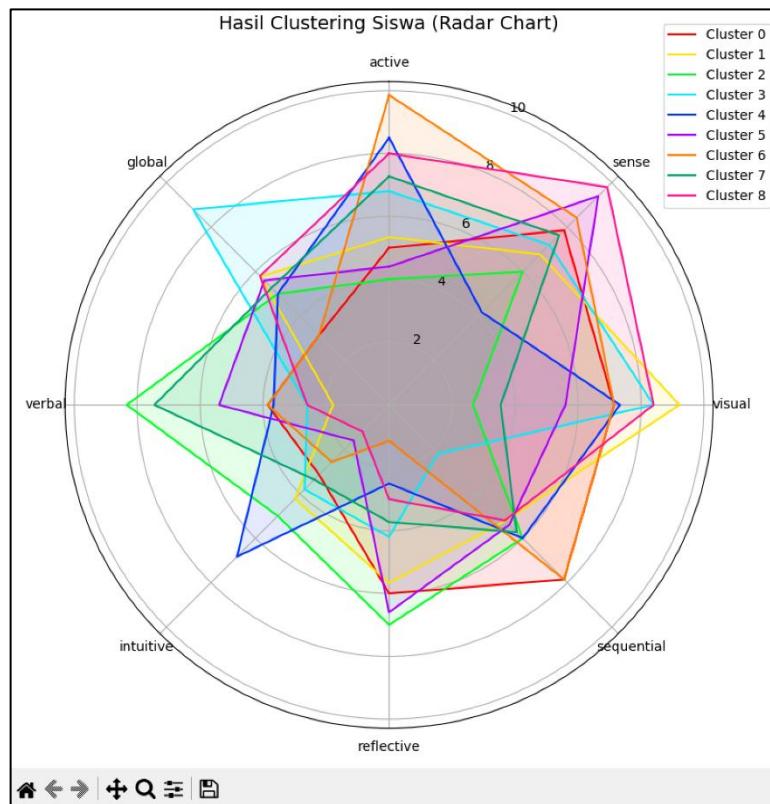
Hasil Clustering											
Analisa Cluster											
cluster	active	reflective	sense	intuitive	visual	verbal	sequential	global	jumlah		
0	5.000000	6.000000	7.857143	3.142857	7.142857	3.857143	7.857143	3.142857	7		
1	5.333333	5.666667	6.777778	4.222222	9.222222	1.777778	5.222222	5.777778	9		
2	4.000000	7.000000	6.000000	5.000000	2.666667	8.333333	6.000000	5.000000	3		
3	6.800000	4.200000	7.200000	3.800000	8.400000	2.600000	2.200000	8.800000	5		
4	8.500000	2.500000	4.166667	6.833333	7.333333	3.666667	6.000000	5.000000	6		
5	4.400000	6.600000	9.400000	1.600000	5.600000	5.400000	5.400000	5.600000	5		
6	9.857143	1.142857	8.428571	2.571429	7.142857	3.857143	7.857143	3.142857	7		
7	7.272727	3.727273	7.636364	3.363636	3.545455	7.454545	5.727273	5.272727	11		
8	8.000000	3.000000	9.800000	1.200000	8.400000	2.600000	5.200000	5.800000	5		

Gambar 4.27. Analisa Cluster Skenario 8

Dari *clustering* yang dilakukan pada skenario 8, didapatkan hasil berupa 9 klaster. Klaster 7 merupakan klaster terbesar yang terdiri dari 11 siswa, memiliki rata – rata yang cukup tinggi pada atribut *sense* dengan nilai 7.63. Sedangkan klaster 2 merupakan klaster terkecil yang terdiri dari 3 siswa, memiliki rata – rata yang cukup tinggi pada atribut *verbal* dengan nilai 8.33.

Klaster 1 merupakan klaster terbesar kedua yang terdiri dari 9 siswa, memiliki rata – rata yang cukup tinggi pada atribut *visual* dengan nilai 9.2, serta rata – rata yang cenderung seimbang pada atribut *active* dan *reflective* serta atribut *global* dan *sequential*.

Berikut merupakan visualisasi dari *clustering* pada skenario 8 dalam bentuk *radar plot* perbandingan dari masing masing kategori pada seluruh dimensi pada skenario 8 yang disajikan pada Gambar 4.28.



Gambar 4.28. Visualisasi Skenario 8

Dari visualisasi tersebut, atribut *sense* didominasi dengan siswa yang termasuk kedalam klaster 0, 1, 3, 7, 6, 5 serta 8, dimana siswa cenderung menyukai belajar berdasarkan data dan fakta yang ada atau data yang bersifat faktual. Hal ini berlawanan dengan atribut *intuitive* yang didominasi oleh siswa yang termasuk kedalam klaster 4, dimana siswa lebih menyukai informasi yang bersifat teoritis.

Atribut *active* didominasi oleh siswa yang termasuk pada klaster 6, 4, 8, 7, serta 3, dimana siswa lebih suka untuk terlibat kedalam diskusi untuk memproses sebuah informasi. Berlawanan dengan atribut *reflective* yang didominasi oleh siswa yang ada pada klaster 0, 1, 2 dan 5, dimana siswa lebih menyukai untuk berdiskusi secara mandiri untuk memproses informasi yang diterima.

### i. Skenario 9

Pada skenario 9, proses *clustering* akan menggunakan jumlah  $k$  dengan nilai 10 sebagai nilai  $k$  inisial dari *clustering*. Pada skenario ini, data akan dikelompokkan menjadi 9 klaster. Berikut merupakan hasil *clustering* pada skenario 9 yang disajikan pada Gambar 4.29.

Hasil Clustering										
===== Hasil Clustering =====										
Jumlah Cluster (k) : 10 Jumlah Iterasi : 7 iterasi Silhouette Score : 0.1997484449800557										
----- Persebaran Data -----										
id	active	reflective	sense	intuitive	visual	verbal	sequential	global	cluster	
s1	5	6	6	5	10	1	3	8	3	
s2	3	8	9	2	5	6	5	6	2	
s3	6	5	8	3	2	9	6	5	7	
s4	4	7	8	3	10	1	7	4	1	
s5	10	1	8	3	3	8	4	7	7	
s6	8	3	6	5	8	3	8	3	4	
s7	5	6	7	4	6	5	6	5	5	
s8	5	6	6	5	10	1	4	7	1	
s9	6	5	8	3	5	6	7	4	7	
s10	7	4	11	0	10	1	5	6	8	
s11	4	7	6	5	3	8	5	6	9	
s12	6	5	7	4	9	2	8	3	1	
s13	8	3	8	3	4	7	1	10	5	
s14	5	6	7	4	8	3	6	5	1	
s15	10	1	9	2	7	4	9	2	6	
s16	5	6	10	1	5	6	5	6	5	
s17	7	4	6	5	4	7	5	6	7	
s18	8	3	9	2	8	3	7	4	8	
s19	6	5	5	6	8	3	4	7	3	
s20	5	6	9	2	8	3	8	3	0	
s21	10	1	9	2	8	3	6	5	6	
s22	9	2	9	2	9	2	3	8	8	

Gambar 4.29. Hasil *Clustering* Skenario 9

Pada skenario 9, terdapat 10 klaster yang keluar sebagai hasil *clustering*. Proses ini berjalan sebanyak 7 iterasi dengan nilai *Silhouette Score* sebesar 0.1997484449800557. Nilai *Silhouette Score* didapatkan dengan cara menghitung jarak rata – rata antara titik data pada klaster yang sama dan klaster tetangga. Persebaran data pada setiap klaster disajikan pada Gambar 4.30.

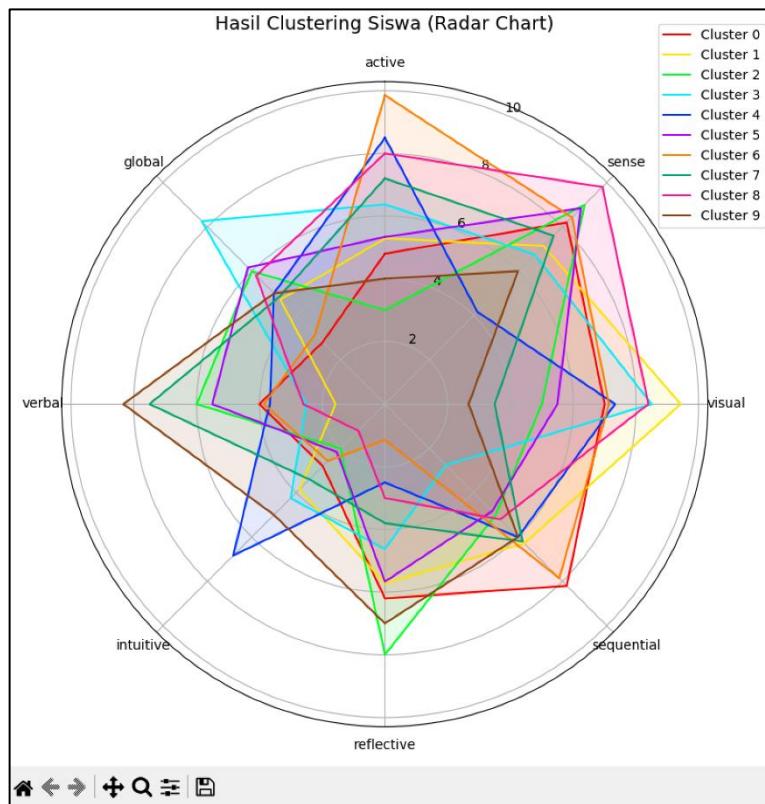
Hasil Clustering										
Analisa Cluster										
cluster	active	reflective	sense	intuitive	visual	verbal	sequential	global	jumlah	
0	4.800000	6.200000	8.200000	2.800000	7.000000	4.000000	8.200000	2.800000	5	
1	5.285714	5.714286	7.142857	3.857143	9.428571	1.571429	6.285714	4.714286	7	
2	3.000000	8.000000	9.000000	2.000000	5.000000	6.000000	5.000000	6.000000	1	
3	6.375000	4.625000	6.750000	4.250000	8.500000	2.500000	2.750000	8.250000	8	
4	8.500000	2.500000	4.166667	6.833333	7.333333	3.666667	6.000000	5.000000	6	
5	5.333333	5.666667	8.833333	2.166667	5.500000	5.500000	4.833333	6.166667	6	
6	9.857143	1.142857	8.428571	2.571429	7.142857	3.857143	7.857143	3.142857	7	
7	7.200000	3.800000	7.600000	3.400000	3.500000	7.500000	6.200000	4.800000	10	
8	8.000000	3.000000	9.800000	1.200000	8.400000	2.600000	5.200000	5.800000	5	
9	4.000000	7.000000	6.000000	5.000000	2.666667	8.333333	6.000000	5.000000	3	

Gambar 4.30. Analisa Cluster Skenario 9

Dari *clustering* yang dilakukan pada skenario 9, didapatkan hasil berupa 10 klaster. Klaster 7 merupakan klaster terbesar yang terdiri dari 10 siswa, memiliki rata – rata yang cukup tinggi pada atribut *sense* dengan nilai 7.60. Sedangkan klaster 2 merupakan klaster terkecil yang terdiri dari 1 siswa, memiliki rata – rata yang cukup tinggi pada atribut *sense* dengan nilai 9.00.

Klaster 1 merupakan klaster terbesar kedua yang terdiri dari 9 siswa, memiliki rata – rata yang cukup tinggi pada atribut *visual* dengan nilai 9.2, serta rata – rata yang cenderung seimbang pada atribut *active* dan *reflective* serta atribut *global* dan *sequential*.

Berikut merupakan visualisasi dari *clustering* pada skenario 9 dalam bentuk *radar plot* perbandingan dari masing masing kategori pada seluruh dimensi pada skenario 9 yang disajikan pada Gambar 4.31.



Gambar 4.31. Visualisasi Skenario 9

Dari visualisasi tersebut, klaster 7 dan 9 memiliki rata – rata yang cukup tinggi pada atribut *global*, yang menunjukkan bahwa siswa cenderung lebih lambat untuk memahami materi yang diberikan. Klaster 0, 2 dan 9 memiliki rata – rata yang cukup tinggi pada atribut *reflective*, dimana siswa lebih suka untuk memproses informasi dengan cara berdiskusi secara mandiri.

Dari visualisasi tersebut juga, dapat dilihat bahwa sebagian besar klaster tergolong pada gaya belajar dengan atribut *active*, *sense* serta *visual*. Hal ini ditunjukkan dengan rata – rata yang cukup tinggi pada atribut tersebut.

### j. Skenario 10

Pada skenario 10, proses *clustering* akan menggunakan jumlah  $k$  dengan nilai 11 sebagai nilai  $k$  inisial dari *clustering*. Pada skenario ini, data akan dikelompokkan menjadi 11 klaster. Berikut merupakan hasil *clustering* pada skenario 11 yang disajikan pada Gambar 4.32.

Hasil Clustering											
===== Hasil Clustering =====											
Jumlah Cluster (k) : 11											
Jumlah Iterasi : 6 iterasi											
Silhouette Score : 0.1799362180935965											
----- Persebaran Data -----											
id	active	reflective	sense	intuitive	visual	verbal	sequential	global	cluster		
s1	5	6	6	5	10	1	3	8	3		
s2	3	8	9	2	5	6	5	6	2		
s3	6	5	8	3	2	9	6	5	7		
s4	4	7	8	3	10	1	7	4	1		
s5	10	1	8	3	3	8	4	7	7		
s6	8	3	6	5	8	3	8	3	4		
s7	5	6	7	4	6	5	6	5	5		
s8	5	6	6	5	10	1	4	7	1		
s9	6	5	8	3	5	6	7	4	7		
s10	7	4	11	0	10	1	5	6	8		
s11	4	7	6	5	3	8	5	6	9		
s12	6	5	7	4	9	2	8	3	1		
s13	8	3	8	3	4	7	1	10	5		
s14	5	6	7	4	8	3	6	5	1		
s15	10	1	9	2	7	4	9	2	6		
s16	5	6	10	1	5	6	5	6	5		
s17	7	4	6	5	4	7	5	6	7		
s18	8	3	9	2	8	3	7	4	4		
s19	6	5	5	6	8	3	4	7	3		
s20	5	6	9	2	8	3	8	3	0		
s21	10	1	9	2	8	3	6	5	6		
s22	9	2	9	2	9	2	3	8	8		

Gambar 4.32. Hasil *Clustering* Skenario 10

Pada skenario 10, terdapat 11 klaster yang keluar sebagai hasil *clustering*. Proses ini berjalan sebanyak 6 iterasi dengan nilai *Silhouette Score* sebesar 0.1799362180935965. Nilai *Silhouette Score* didapatkan dengan cara menghitung jarak rata – rata antara titik data pada klaster yang sama dan klaster tetangga. Persebaran data pada setiap klaster disajikan pada Gambar 4.33.

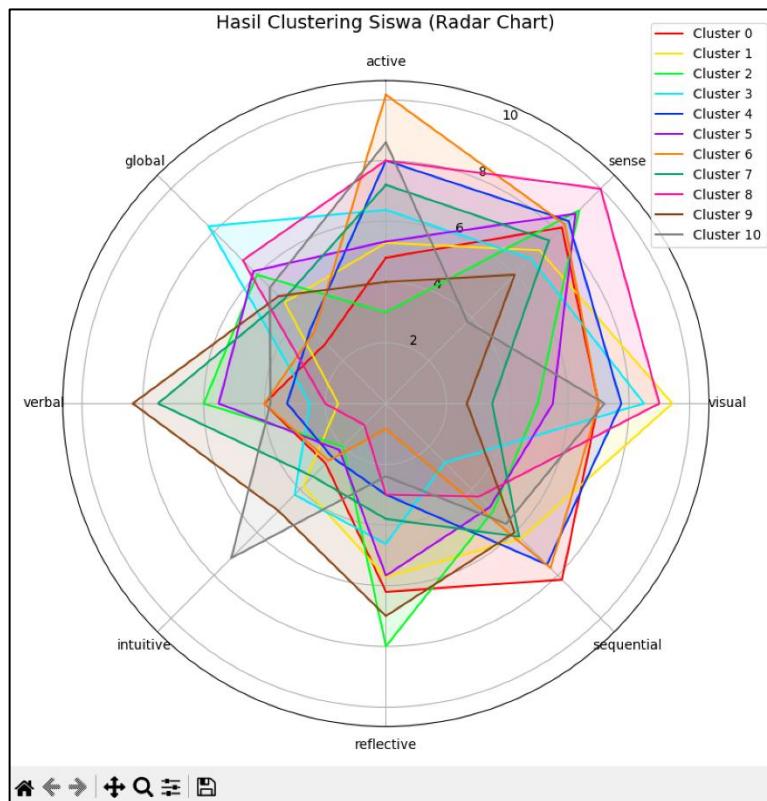
Hasil Clustering										
Analisa Cluster										
cluster	active	reflective	sense	intuitive	visual	verbal	sequential	global	jumlah	
0	4.800000	6.200000	8.200000	2.800000	7.000000	4.000000	8.200000	2.800000	5	
1	5.285714	5.714286	7.142857	3.857143	9.428571	1.571429	6.285714	4.714286	7	
2	3.000000	8.000000	9.000000	2.000000	5.000000	6.000000	5.000000	6.000000	1	
3	6.375000	4.625000	6.750000	4.250000	8.500000	2.500000	2.750000	8.250000	8	
4	8.000000	3.000000	8.500000	2.500000	7.750000	3.250000	7.500000	3.500000	4	
5	5.333333	5.666667	8.833333	2.166667	5.500000	5.500000	4.833333	6.166667	6	
6	10.166667	0.833333	8.333333	2.666667	7.000000	4.000000	7.666667	3.333333	6	
7	7.200000	3.800000	7.600000	3.400000	3.500000	7.500000	6.200000	4.800000	10	
8	8.000000	3.000000	10.000000	1.000000	9.000000	2.000000	4.333333	6.666667	3	
9	4.000000	7.000000	6.000000	5.000000	2.666667	8.333333	6.000000	5.000000	3	
10	8.600000	2.400000	3.800000	7.200000	7.200000	3.800000	5.600000	5.400000	5	

Gambar 4.33. Analisa Cluster Skenario 10

Dari *clustering* yang dilakukan pada skenario 10, didapatkan hasil berupa 11 klaster. Klaster 7 merupakan klaster terbesar yang terdiri dari 10 siswa, memiliki rata – rata yang cukup tinggi pada atribut *sense* dengan nilai 7.60. Sedangkan klaster 2 merupakan klaster terkecil yang terdiri dari 1 siswa, memiliki rata – rata yang cukup tinggi pada atribut *sense* dengan nilai 9.00.

Klaster 1 merupakan klaster terbesar kedua yang terdiri dari 9 siswa, memiliki rata – rata yang cukup tinggi pada atribut *visual* dengan nilai 9.2, serta rata – rata yang cenderung seimbang pada atribut *active* dan *reflective* serta atribut *global* dan *sequential*.

Berikut merupakan visualisasi dari *clustering* pada skenario 10 dalam bentuk *radar plot* perbandingan dari masing masing kategori pada seluruh dimensi pada skenario 10 yang disajikan pada Gambar 4.34.



Gambar 4.34. Visualisasi Skenario 10

Dari visualisasi tersebut, terdapat banyak klaster yang tergolong pada gaya belajar dengan atribut *sense*. Hal ini ditunjukkan dengan banyaknya klaster yang memiliki rata – rata yang cukup tinggi pada atribut *sense*, yaitu seluruh klaster yang ada selain dari klaster 9 dan 10.

Hal ini juga didapati pada gaya belajar dengan atribut *active* serta *visual*. Dimana klaster 3, 7, 4, 6, 8 dan 10 memiliki rata – rata yang cukup tinggi pada atribut *active*, yang menunjukkan bahwa siswa lebih mudah untuk memproses informasi melalui diskusi. Sedangkan klaster 0, 1, 3, 4, 6 dan 10 memiliki rata – rata yang cukup tinggi pada atribut *visual*, yang menunjukkan bahwa siswa lebih menyukai informasi berupa ilustrasi seperti gambar atau diagram.

#### 4.1.4 Evaluasi *Silhouette Score*

Setelah data diolah menggunakan Algoritma K-Means Clustering pada serangkaian skenario yang ada, klaster yang keluar sebagai hasil akan diuji validitasnya menggunakan *Silhouette Score*. Implementasi *Silhouette Score* dilakukan pada Bahasa Pemrograman *Python*. *Silhouette Score* didapatkan dengan menghitung nilai dari jarak rata – rata antara titik (*i*) dengan semua titik data pada klaster yang lain atau  $b(i)$  dan jarak rata – rata antara titik (*i*) dengan semua titik data pada klaster yang sama  $a(i)$ . Berikut merupakan implementasi *silhouette score* pada sistem.

Tabel 4.7. Implementasi *Source Code Silhouette Score*

No	Source Code
1	# a(i)
2	if len(own_cluster_points) > 1:
3	a = np.mean([distance(data[i], point) for j, point in enumerate(own_cluster_points) if not np.array_equal(data[i], point)])
4	else:
5	a = 0
6	
7	# b(i)
8	b = float('inf')
9	for other_label in other_cluster_labels:
10	other_points = data[labels == other_label]
11	dist = np.mean([distance(data[i], point) for point in other_points])
12	b = min(b, dist)
13	
14	# silhouette coefficient
15	s = (b - a) / max(a, b) if max(a, b) != 0 else 0
16	silhouette_scores.append(s)

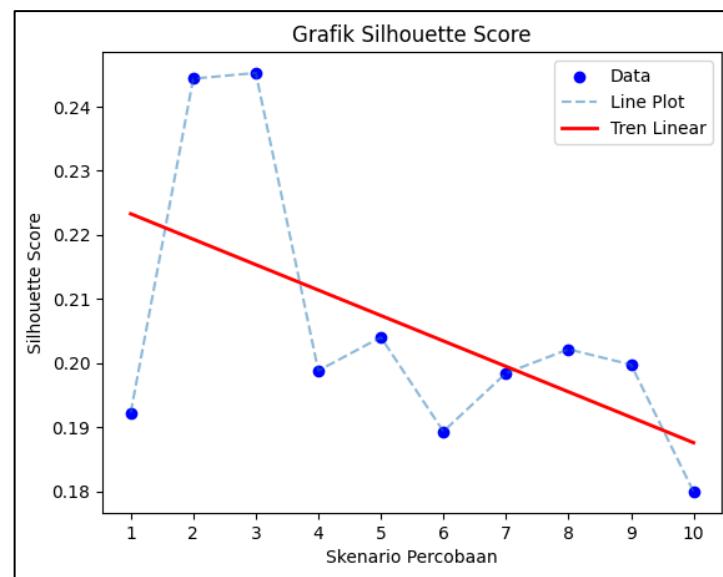
Implementasi *silhouette score* dilakukan pada seluruh skenario yang dilakukan. Hal ini dilakukan dengan tujuan untuk mengetahui skenario yang

memiliki tingkat klasterisasi paling baik. Nilai *silhouette score* pada skenario uji coba disajikan pada Tabel 4.8.

Tabel 4.8. Hasil *Silhouette Score* pada Skenario Percobaan

Skenario percobaan	Jumlah klaster	Silhouette Score
1	2	0.1922551004321032
2	3	0.2443461171295426
3	4	0.24523330344105687
4	5	0.19880873001579144
5	6	0.2040150748088004
6	7	0.18933045635578755
7	8	0.19841597698077995
8	9	0.20213839213565232
9	10	0.1997484449800557
10	11	0.1799362180935965

Berikut merupakan visualisasi dari nilai *silhouette score* pada skenario percobaan dengan jumlah  $k = 2$  sampai dengan 11 yang disajikan pada Gambar 4.35.



Gambar 4.35. Grafik *Silhouette Score*

Dari visualisasi diatas, dapat diketahui bahwa skenario 3 memiliki nilai *silhouette score* yang paling baik. Hal ini menunjukkan bahwa skenario 3 memiliki tingkat klasterisasi yang lebih baik daripada skenario yang lain.

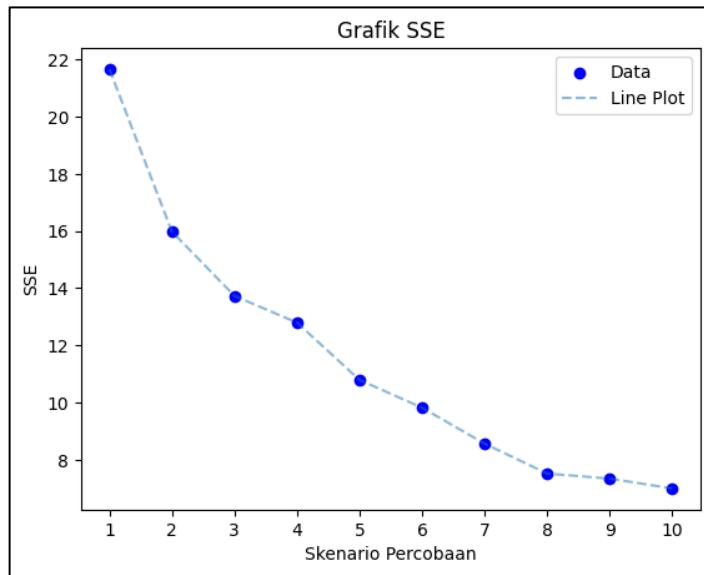
#### 4.1.5 Evaluasi *Elbow Method*

Selain diuji menggunakan *Silhouette Score*, data hasil pengolahan juga akan diuji validitasnya menggunakan *Elbow Method*. Metode ini berjalan dengan cara mencari pola yang berbentuk seperti siku dari nilai SSE yang didapatkan pada masing – masing skenario. Berikut merupakan hasil perhitungan SSE dari masing – masing skenario yang disajikan pada Tabel 4.9.

Tabel 4.9. Nilai SSE pada Skenario Percobaan

Skenario percobaan	Jumlah klaster	Nilai SSE	Penurunan
1	2	21.666581384015593	-
2	3	15.980669891413356	5.6859
3	4	13.730575917352539	2.2501
4	5	12.79545886344804	0.9351
5	6	10.789342109365258	2.0061
6	7	9.817878446784697	0.9715
7	8	8.558976212188018	1.2589
8	9	7.514929664813924	1.0440
9	10	7.337986386684303	0.1769
10	11	6.992199900793651	0.3458

Dari tabel yang didapatkan, terjadi penurunan yang sangat besar pada skenario 1, dimana  $k = 2$  ke skenario 2, dimana  $k = 3$ . Selisih penurunan masih terbilang cukup besar pada skenario 2, dimana  $k = 3$  ke skenario 3, dimana  $k = 4$ . Setelah skenario 3, perubahan penurunan mulai sedikit menurun, walaupun selisih penurunan kembali meningkat. Perubahan ini dapat dilihat lebih jelas pada Gambar 4.36.



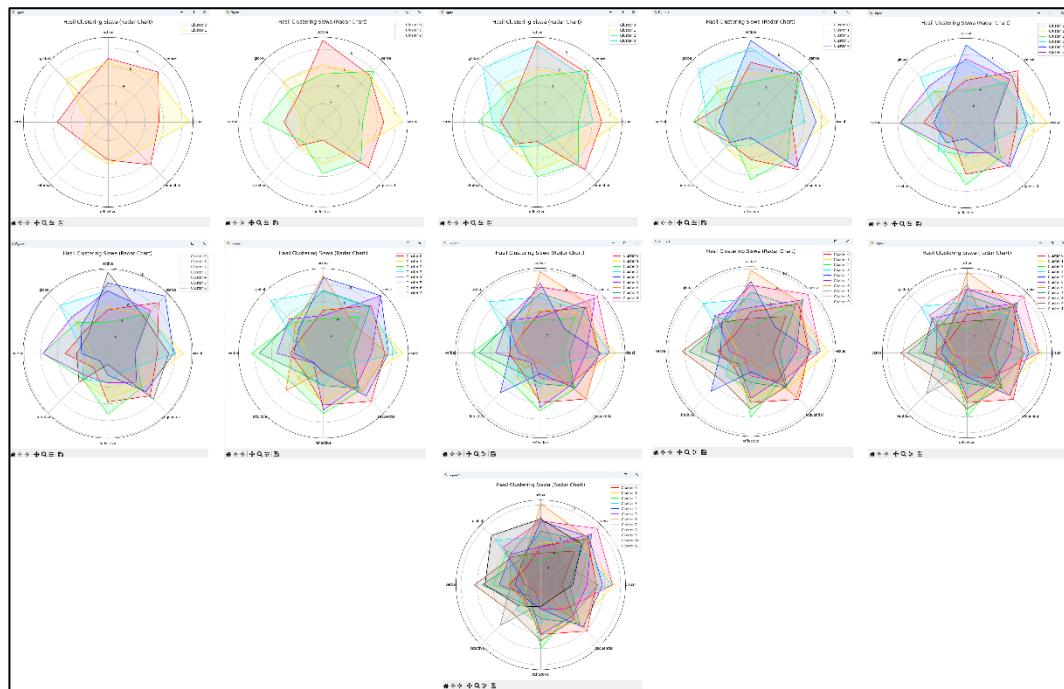
Gambar 4.36. Visualisasi Grafik SSE

Dari visualisasi tersebut, dapat diketahui bahwa perubahan nilai SSE yang terjadi pada skenario 3 menuju skenario 4 menunjukkan selisih yang mengecil. Ini merupakan perubahan pertama yang menunjukkan selisih nilai SSE yang mengecil pada seluruh skenario. Namun, fenomena ini tidak hanya terjadi pada skenario 3. Fenomena perubahan nilai SSE yang mengecil juga terjadi pada skenario 5 menuju skenario 6. Selain itu, fenomena ini juga terjadi pada skenario 8 menuju skenario 9. Dari hasil tersebut, dapat diketahui bahwa titik siku pada proses *clustering* ini berada pada skenario 3, dimana skenario ini menjadi skenario pertama yang memiliki perubahan nilai SSE yang mengecil atau mengalami penurunan yang signifikan.

## 4.2 Pembahasan

Proses *clustering* siswa bimbingan belajar dilakukan dengan beberapa skenario yang telah ditentukan. Dari beberapa skenario yang telah dilakukan,

terdapat sebuah pola berulang yang hampir selalu muncul pada seluruh skenario. Beberapa klaster memiliki nilai yang dominan pada gaya belajar dengan atribut *active*, *sense*, serta *visual*. Pola ini dapat dilihat pada Gambar 4.37.



Gambar 4.37. Visualisasi Seluruh Skenario

Hal ini menunjukkan bahwa mayoritas siswa bimbingan belajar memiliki kecenderungan pada gaya belajar *active*, *sense* dan *visual*, yang ditunjukkan dengan nilai rata – rata yang cukup tinggi pada ketiga atribut tersebut. Ini pertama kali muncul pada skenario 2, dimana masing – masing klaster memiliki ciri berupa nilai rata – rata yang tinggi pada atribut *active*, *sense* dan *visual*.

Kemudian pada skenario 3 muncul sebuah pola baru, dimana klaster 3 pada skenario 3 muncul dengan nilai rata – rata yang tinggi pada atribut *active* dan *global*. Sedangkan pada skenario sebelumnya, atribut *global* ditunjukkan memiliki rata – rata yang tidak terlalu tinggi. Ini merupakan pola yang menunjukkan bahwa

proses pemisahan data pada *clustering* semakin meningkat, dimana pada skenario ini, muncul sebuah klaster baru yang memiliki ciri serta pola yang sebelumnya tidak ada pada skenario sebelumnya. Hal ini juga didukung oleh eskalasi nilai *silhouette score* yang ada pada skenario 3. Selain nilai *silhouette score* yang meningkat, pada skenario ini pula pola seperti siku muncul pada grafik SSE untuk pertama kali. Hal tersebut ditunjukkan dengan penurunan nilai SSE yang kian mengecil.

Pada skenario selanjutnya, tidak ada kemunculan pola baru yang memiliki perbedaan yang signifikan dengan skenario – skenario sebelumnya. Fenomena ini berarti bahwa proses pemisahan data yang terjadi tidak menunjukkan pembentukan kelompok baru, namun hanya memisah kelompok yang sudah homogen menjadi beberapa bagian kecil. Ini juga dipertegas dengan penurunan nilai *silhouette score* pada skenario lanjutan setelah skenario 3.

Hal ini terus berlanjut sampai dengan skenario 7, dimana muncul sebuah klaster yang memiliki dominasi pada atribut *intuitive*. Klaster ini merupakan pola baru yang tidak dimiliki oleh skenario sebelumnya, dimana hal ini juga diperkuat dengan kenaikan nilai *silhouette score* pada skenario 7 dan 8. Kenaikan tersebut mengindikasikan adanya peningkatan kualitas pemisahan data yang terjadi pada skenario tersebut. Selain *silhouette score* yang menunjukkan peningkatan, skenario 8 juga menunjukkan sebuah pola seperti siku yang kembali muncul untuk kedua kalinya. Dimana hal ini menunjukkan bahwa *silhouette score* dan *elbow method* menampilkan hasil yang saling mendukung satu sama lain.

Namun, fenomena kenaikan nilai *silhouette* dan pola siku yang ditunjukkan oleh *elbow method* tersebut hanya bertahan sampai dengan skenario 8, dimana pada

skenario selanjutnya, tidak ada pembentukan pola yang berbeda. Hal tersebut ditunjukkan dengan penuruan nilai *silhouette score* setelah skenario 8.

Fenomena ini juga dipengaruhi oleh keterbatasan variabel yang digunakan pada proses *clustering*. Pada proses *clustering* siswa bimbingan belajar tesla Sidoarjo, variabel yang digunakan berasal dari 8 dimensi berbeda yang berpasangan sekaligus berlawanan pada masing – masing dimensi. Karena keterbatasan yang ada pada variabel yang digunakan, variasi antar siswa juga relatif terbatas.

Hal ini mengakibatkan tidak ada pola baru yang terbentuk secara signifikan setelah pemisahan data mencapai 4 klaster. Karena keempat klaster tersebut sudah cukup untuk menunjukkan pengelompokan data yang memiliki ciri serta karakteristik masing – masing. Pemisahan data dengan tingkat klaster lebih dari 4 hanya akan membagi kelompok – kelompok yang sudah ada menjadi kelompok yang lebih kecil yang memiliki ciri serta karakteristik yang serupa. Hal ini juga didukung oleh penurunan nilai *silhouette score* setelah skenario 3 yang membagi data menjadi 4 klaster.

Pada penelitian ini, proses *clustering* siswa dilakukan pada beberapa skenario percobaan. Pada proses tersebut, peneliti tidak memberi batasan pada jumlah iterasi yang berjalan di proses *clustering* siswa. Berikut merupakan jumlah iterasi pada masing – masing skenario yang disajikan pada Tabel 4.10.

Tabel 4.10. Jumlah Iterasi pada Skenario

Skenario percobaan	Jumlah klaster	Jumlah iterasi
1	2	3
2	3	5
3	4	5
4	5	5
5	6	5
6	7	4
7	8	7
8	9	5
9	10	7
10	11	6

Berdasarkan tabel diatas, tidak ada keterkaitan yang jelas antara jumlah klaster dengan jumlah iterasi yang berjalan pada proses *clustering*. Tabel tersebut menunjukkan perubahan jumlah iterasi tanpa adanya pola yang jelas, yang menunjukkan bahwa tidak ada keterkaitan secara langsung antara jumlah  $k$  atau klaster terhadap jumlah iterasi yang berjalan pada skenario tersebut.

Dari seluruh skenario yang telah dilakukan, skenario 3 menjadi skenario *clustering* paling optimal dengan *silhouette score* tertinggi diantara skenario lainnya. Hasil yang didapatkan dari skenario ini dapat digunakan oleh tenaga pendidik atau guru untuk menyesuaikan strategi serta model pembelajaran yang cocok untuk diaplikasikan pada masing – masing klaster sesuai dengan dimensi gaya belajar yang ada pada *Index of Learning Styles*(Felder & Silverman, 1988).

Skenario 3 terbagi menjadi 4 klaster, klaster 0 didominasi oleh siswa dengan kecenderungan tinggi pada atribut *active*, dimana siswa lebih menyukai gaya belajar aktif seperti melakukan eksperimen. Klaster 1 didominasi oleh siswa dengan

kecenderungan tinggi pada atribut *visual*, dimana siswa lebih menyukai lebih menyukai penjelasan berupa ilustrasi atau gambar.

Klaster 2 didominasi oleh siswa dengan kecenderungan tinggi pada atribut *sense*, dimana siswa lebih menyukai sesuatu yang bersifat faktual dan berdasarkan data. Sedangkan klaster 3 didominasi oleh siswa dengan kecenderungan tinggi pada atribut *active*. Klaster 3 juga didominasi oleh siswa dengan kecenderungan tinggi pada atribut *global*, dimana siswa lebih suka untuk diajak belajar dan berpikir dengan bebas tanpa terpaku oleh beberapa standar yang ada.

Dari beberapa klaster tersebut, guru dapat menyesuaikan strategi pembelajaran dengan identitas serta karakteristik yang ada pada masing – masing klaster. Sebagai contoh, guru dapat melakukan pembelajaran dengan cara diskusi dan melibatkan siswa secara aktif dengan siswa yang tergolong pada klaster 0 dengan kecenderungan tinggi atribut *active*. Guru juga bisa memanfaatkan beberapa ilustrasi sebagai media belajar berupa gambar ataupun diagram kepada siswa yang memiliki kecenderungan tinggi pada atribut *visual* pada klaster 1.

Pada klaster 2, dimana siswa memiliki kecenderungan tinggi pada atribut *sense*, guru dapat menyajikan data yang bersifat faktual serta memberikan siswa beberapa rincian kecil dari sebuah materi pembelajaran. Untuk siswa dengan kecenderungan tinggi pada atribut *global* pada klaster 3, guru dapat memberi siswa kebebasan untuk mencoba menyelesaikan sebuah masalah dengan cara mereka sendiri.

### 4.3 Integrasi Islam

Dari hasil penelitian yang telah dilakukan, siswa yang telah dikelompokkan melalui *clustering* didapatkan sebagai hasil dari penelitian. Hasil ini tidak hanya mencakup lingkup akademis saja, namun juga mencakup lingkup keagamaan. Dalam konteks beragama, pengelompokan seperti *clustering* siswa yang didapatkan sebagai hasil sesuai dengan Firman Allah Swt pada Surat *Hujurat* ayat 13, yang berbunyi

يَأَيُّهَا الْأَنَاسُ إِنَّا خَلَقْنَاكُم مِّنْ ذَرَّةٍ وَأَنْتُمْ وَجَعَلْنَاكُمْ شُعُوبًا وَقَبَائِلَ لِتَعْرِفُوْا إِنَّ أَكْرَمَكُمْ عِنْدَ اللَّهِ أَنْفُسُكُمْ إِنَّ اللَّهَ عَلِيمٌ حَمِيرٌ

“Wahai manusia, sesungguhnya Kami telah 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 dan Mahateliti”(Q.S Hujurat: 13).

Pada tafsir Ibnu Katsir, Allah Swt menciptakan manusia dari laki – laki (Adam) dan perempuan (Hawa) dan menjadikan mereka berbangsa – bangsa. Dari pengertian bangsa Arab, bangsa diartikan sebagai kabilah – kabilah atau kelompok seperti Bani Israil. Allah Swt menjadikan manusia berbangsa – bangsa dengan tujuan untuk saling mengenal diantara sesama mereka (Ghoffar & al-Atsari, 2004). Ayat ini memiliki hubungan erat dengan penelitian ini. Kabilah – kabilah yang disebut sebagai bangsa – bangsa memiliki hubungan erat dengan kelompok klaster yang didapatkan dari hasil penelitian ini.

Klaster – klaster yang didapatkan dari hasil penelitian ini dapat direpresentasikan sebagai kelompok bangsa – bangsa yang memiliki keberagaman. Masing – masing klaster memiliki karakteristik serta ciri yang berbeda diantara satu

klaster satu dengan klaster lainnya. Dari karakteristik serta ciri – ciri yang ada pada klaster tersebut, tenaga pendidik atau guru dapat melakukan analisa serta menentukan sikap serta strategi seperti apa yang harus diaplikasikan pada klaster dengan karakteristik tersebut.

## **BAB V**

### **KESIMPULAN DAN SARAN**

#### **5.1. Kesimpulan**

Dari penelitian yang telah dilakukan, didapatkan kesimpulan sebagai berikut:

1. Algoritma K-Means *Clustering* dapat digunakan untuk mengelompokkan siswa bimbingan belajar berdasarkan kuesioner ILS. Siswa bimbingan belajar dikelompokkan menjadi beberapa klaster sesuai dengan skenario uji coba yang dilakukan. Dari beberapa skenario yang telah dilakukan, skenario 3 keluar sebagai skenario paling baik yang memiliki *silhouette score* sebesar 0.24523330344105687.
2. Skenario ini membagi siswa bimbingan belajar menjadi 4 klaster. Dimana klaster 0 didominasi siswa dengan gaya belajar *active*, klaster 1 didominasi siswa dengan gaya belajar *visual*, klaster 2 didominasi siswa dengan gaya belajar *sense*, dan klaster 3 didominasi siswa dengan gaya belajar *global*.
3. Hasil pengelompokan siswa bimbingan belajar yang didapatkan bisa dimanfaatkan oleh tenaga didik ataupun guru untuk melakukan analisa serta menentukan strategi yang sesuai pada proses pembelajaran. Sebagai contoh,
  - a. Tenaga didik atau guru dapat menerapkan pembelajaran dengan pendekatan diskusi secara aktif untuk siswa dengan gaya belajar *active*.
  - b. Untuk gaya belajar *visual*, guru dapat menerapkan pendekatan pembelajaran dengan mencantumkan gambar atau diagram, dimana

siswa dengan gaya belajar ini lebih mudah untuk mengingat apa yang mereka lihat.

- c. Untuk gaya belajar *sense*, guru dapat menerapkan pendekatan berupa fakta atau prosedur langsung pada proses pembelajaran, dimana siswa dengan gaya belajar ini cenderung berorientasi pada fakta.
- d. Sedangkan untuk gaya belajar *global*, guru dapat menerapkan pembelajaran dimana guru menjelaskan sebuah materi secara utuh dan lengkap, karena siswa dengan gaya belajar ini memerlukan pembahasan yang menyeluruh agar bisa menerapkan materi atau pembelajaran yang didapat.

## 5.2. Saran

Pada penelitian yang dilakukan, peneliti menyadari adanya kekurangan yang dapat ditingkatkan untuk penelitian yang akan datang. Dari kesimpulan yang didapatkan, berikut merupakan saran yang dapat digunakan untuk penelitian mendatang:

1. Nilai *Silhouette Score* pada penelitian ini masih tergolong rendah. Oleh karena itu, peneliti menyarankan untuk melakukan serangkaian optimasi pada proses pengolahan data, seperti reduksi dimensi, optimasi *k-means* atau pemilihan *feature k-means* pada penelitian yang akan datang. Hal ini dilakukan dengan harapan adanya peningkatan kualitas klaster yang terbentuk.
2. Penelitian ini hanya terbatas sampai dengan mengelompokkan siswa kedalam beberapa kelompok. Pada penelitian yang akan datang, penulis

menyarankan untuk melanjutkan proses *clustering* sampai pada tahap pelabelan data.

3. Penelitian selanjutnya juga dapat menggunakan data dengan jumlah responden yang lebih banyak sebagai upaya peningkatan validitas hasil *clustering*.

## DAFTAR PUSTAKA

- Andi Akram Nur Risal, Dyah Darma Andayani, Muh Ilham Suherman, & Andi Baso Kaswar. (2024). Utilizing the K-Means Clustering Algorithm for Analyzing Student Achievement Assessment at SMK Negeri 1 Gowa. *Journal of Embedded Systems, Security and Intelligent Systems*, 60–67. <https://doi.org/10.59562/jessi.v5i1.2178>
- Anggreani, D., Nurmisba, N., Setiawan, D., & Lukman, L. (2024). Optimization of K-Means Clustering Method by Using Elbow Method in Predicting Blood Requirement of Pelamonia Hospital Makassar. *Internet of Things and Artificial Intelligence Journal*, 4(3), 541–550. <https://doi.org/10.31763/iota.v4i3.755>
- Asroni, A., Kurniasari, D., & Kurnianti, A. (2020). The Implementation of Clustering Method With K-Means Algorithm In Grouping Data of Students' Course Scores at Universitas Muhammadiyah Yogyakarta. *Emerging Information Science and Technology*, 1(3), 75–83. <https://doi.org/10.18196/eist.v1i3.13172>
- Ayu Syahfitri, Novriyenni, & Imeldawaty Gultom. (2024). Pengelompokan Data Siswa Berdasarkan Profil Pelajar Pancasila Menggunakan Metode Clustering (Studi Kasus SMK Putra Anda Binjai). *Indonesian Journal of Education And Computer Science*, 2(2), 107–120. <https://doi.org/10.60076/indotech.v2i2.644>
- Azzahrah Putri, R., Magdalena, I., Fauziah, A., & Nur Azizah, F. (2021). Pengaruh Gaya Belajar terhadap Pembelajaran Siswa Sekolah Dasar. *Cerdika: Jurnal Ilmiah Indonesia*, 1(2), 157–163. <https://doi.org/10.36418/cerdika.v1i2.26>
- Chen, L., Zhong, C., & Zhang, Z. (2023). Explanation of clustering result based on multi-objective optimization. *PLOS ONE*, 18(10), e0292960. <https://doi.org/10.1371/journal.pone.0292960>
- Chhabra, A., Masalkovaite, K., & Mohapatra, P. (2021). An Overview of Fairness in Clustering. *IEEE Access*, 9, 130698–130720. <https://doi.org/10.1109/ACCESS.2021.3114099>
- Chi, D. (2021). Research on the Application of K-Means Clustering Algorithm in Student Achievement. *2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE)*, 435–438. <https://doi.org/10.1109/ICCECE51280.2021.9342164>
- Dyulicheva, Yu. Yu., & Kosova, Ye. A. (2020). The Students Group Detection Based on the Learning Styles and Clustering Algorithms: *Proceedings of the*

*2nd International Scientific and Practical Conference on Digital Economy (ISCDE 2020).* 2nd International Scientific and Practical Conference on Digital Economy (ISCDE 2020), Yekaterinburg, Russia. <https://doi.org/10.2991/aebmr.k.201205.017>

- Ediyanto. (2022). Pemetaan Kebutuhan Belajar Berdasarkan Gaya Belajar Peserta Didik. *Seminar Nasional Teknologi Dan Multidisiplin Ilmu (SEMNASTEKMU)*, 2(2), 322–340. <https://doi.org/10.51903/semnastekmu.v2i1.187>
- Felder, R. M., & Brent, R. (2005). Understanding Student Differences. *Journal of Engineering Education*, 94(1), 57–72. <https://doi.org/10.1002/j.2168-9830.2005.tb00829.x>
- Felder, R. M., & Silverman, L. K. (1988). Learning and Teaching Styles in Engineering Education. *Engineering Education*, 78(7), 674–681.
- Felder, R. M., & Spurlin, J. (2005). Applications, Reliability and Validity of the Index of Learning Styles. *Journal of Engineering Education*, 21(1), 103–112. <https://doi.org/10.1037/t43782-000>
- Ghoffar, M. A., & al-Atsari, A. I. (2004). *Tafsir Ibnu Katsir* (Vol. 7). Jakarta: Pustaka Imam asy-Syaffi'i.
- Hidayati, R., Zubair, A., Pratama, A. H., & Indiana, L. (2021). Analisis Silhouette Coefficient pada 6 Perhitungan Jarak K-Means Clustering. *Techno.Com*, 20(2), 186–197. <https://doi.org/10.33633/tc.v20i2.4556>
- Kartina, A. D., & Safii, M. (2021). Implementation of K-Means Clustering on High School Students Management. *Journal of Artificial Intelligence and Engineering Applications (JAIEA)*, 1(1), 15–21. <https://doi.org/10.59934/jaiea.v1i1.47>
- Maryono, D., Budiyanto, C. W., & Pamungkas, A. A. (2022). Implementation of K-Means Clustering for Optimization of Student Grouping Based on Index of Learning Styles in Programming Classes. *IJIE (Indonesian Journal of Informatics Education)*, 6(2), 84. <https://doi.org/10.20961/ijie.v6i2.68151>
- Najm Al Inu, A. N., Fitriani, D., Bani, E. A. S., & Winandar, M. L. (2023). Peran Guru Sebagai Agen Pembaharu dalam Meningkatkan Kualitas Pembelajaran yang Inovatif di Sekolah Dasar. *Journal on Education*, 5(2), 1696–1701. <https://doi.org/10.31004/joe.v5i2.806>
- Pamungkas, A. A. P., Maryono, D., & Budiyanto, C. W. (2021). Cluster Analysis for Student Grouping Based on Index of Learning Styles. *Journal of Physics: Conference Series*, 1808(1), 012023. <https://doi.org/10.1088/1742-6596/1808/1/012023>

- Pavlopoulos, J., Vardakas, G., & Likas, A. (2025). Revisiting Silhouette Aggregation. In D. Pedreschi, A. Monreale, R. Guidotti, R. Pellungrini, & F. Naretto (Eds.), *Discovery Science* (Vol. 15243, pp. 354–368). Springer Nature Switzerland. [https://doi.org/10.1007/978-3-031-78977-9\\_23](https://doi.org/10.1007/978-3-031-78977-9_23)
- Sinaga, K. P., & Yang, M.-S. (2020). Unsupervised K-Means Clustering Algorithm. *IEEE Access*, 8, 80716–80727. <https://doi.org/10.1109/ACCESS.2020.2988796>
- Subakti, H., & Handayani, E. S. (2020). Pengaruh Bimbingan Belajar Terhadap Hasil Belajar Bahasa Indonesia Siswa Kelas Tinggi di Sekolah Dasar. *Jurnal Basicedu*, 5(1), 247–255. <https://doi.org/10.31004/basicedu.v5i1.648>
- Syahputra, Y. H., & Hutagalung, J. (2022). Superior Class to Improve Student Achievement Using the K-Means Algorithm. *SinkrOn*, 7(3), 891–899. <https://doi.org/10.33395/sinkron.v7i3.11458>
- Telaumbanua, E. D. P., & Harefa, A. R. (2024). Pengaruh Gaya Belajar terhadap Hasil Belajar Siswa. *Journal of Education Research*, 5(1), 691–697. <https://doi.org/10.37985/jer.v5i1.873>
- Wahyuningsih, N., Nurbayani, E., & Saugi, W. (2020). Pengaruh Rasio Jumlah Siswa dalam Kelas Terhadap Efektivitas Pembelajaran PAI di SMK Farmasi Samarinda. *Tarbiyah Wa Ta'lim: Jurnal Penelitian Pendidikan Dan Pembelajaran*, 6(1), 47–61. <https://doi.org/10.21093/twt.v6i1.1941>
- Yotaman, N., Osathanunkul, K., Khoenkaw, P., & Pramokchon, P. (2020). Teaching Support System by Clustering Students According to Learning Styles. *2020 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON)*, 137–140. <https://doi.org/10.1109/ECTIDAMTNCON48261.2020.9090729>