

Penerapan Metode K-Means Clustering untuk Pengelompokan Peminatan Unit Kegiatan Mahasiswa

Syafri Aprudi¹, Endang Etriyanti²

¹Ilmu Komputer, Fakultas Teknik, Universitas Musi Rawas

²Sistem Informasi, STMIK Bina Nusantara Jaya Lubuklinggau

¹syafri1971@gmail.com, ²endang.eriyanti@gmail.com

Abstract

This Research aims to group the interests of Student Activity Units (UKM) using the K-Means Clustering method as a data-based approach to support decision-making in a higher education environment. The data used consists of 20 students with six variables, namely sports interest, entrepreneurship interest, IT Club interest, LDK interest, Anti-Drug Student Movement (GAMAN) interest, and art interest, which are measured using a numeric scale of 1–10. The research stages include data collection, preprocessing (cleaning and normalization), determining the number of clusters using the Elbow method, the clustering process with the K-Means algorithm, and evaluation using the Silhouette Score. The results of the Elbow method show that the optimal number of clusters is $K = 2$. The K-Means process produces two main clusters, namely the first cluster dominated by students with high interests in sports and arts, and the second cluster dominated by interests in entrepreneurship, IT, LDK, and GAMAN. The evaluation results show an average Silhouette Score of 0.73, which is in the very good category, indicating that the clustering model has high quality and is able to separate the data clearly. This study proves that the K-Means method is effective in grouping students based on UKM interests and can be used as a basis for decision-making, such as channeling interests, developing UKM programs, and improving student services. For further research, it is recommended to add other variables so that the results obtained are more comprehensive and accurate.

Keywords: K-Means Clustering, Unit Kegiatan Mahasiswa, Clustering, Silhouette Score

Abstrak

Penelitian ini bertujuan untuk mengelompokkan peminatan Unit Kegiatan Mahasiswa (UKM) menggunakan metode K-Means Clustering sebagai pendekatan berbasis data dalam mendukung pengambilan keputusan di lingkungan perguruan tinggi. Data yang digunakan terdiri dari 20 mahasiswa dengan enam variabel, yaitu minat olahraga, minat wirausaha, minat IT Club, minat LDK, minat Gerakan Mahasiswa Anti Narkoba (GAMAN), dan minat seni, yang diukur menggunakan skala numerik 1–10. Tahapan penelitian meliputi pengumpulan data, preprocessing (cleaning dan normalisasi), penentuan jumlah cluster menggunakan metode Elbow, proses clustering dengan algoritma K-Means, serta evaluasi menggunakan Silhouette Score. Hasil metode Elbow menunjukkan bahwa jumlah cluster optimal adalah $K=2$. Proses K-Means menghasilkan dua kelompok utama, yaitu cluster pertama yang didominasi oleh mahasiswa dengan minat tinggi pada bidang olahraga dan seni, serta cluster kedua yang didominasi oleh minat pada bidang wirausaha, IT, LDK, dan GAMAN. Hasil evaluasi menunjukkan nilai Silhouette Score rata-rata sebesar 0,73 yang termasuk dalam kategori cukup baik, sehingga menunjukkan bahwa model clustering memiliki kualitas yang tinggi dan mampu memisahkan data secara jelas. Penelitian ini membuktikan bahwa metode K-Means efektif dalam mengelompokkan mahasiswa berdasarkan minat UKM dan dapat digunakan sebagai dasar dalam pengambilan keputusan, seperti penyaluran minat, pengembangan program UKM, serta peningkatan layanan kemahasiswaan. Untuk penelitian selanjutnya, disarankan menambahkan variabel lain agar hasil yang diperoleh lebih komprehensif dan akurat.

Kata kunci: K-Means Clustering, Unit Kegiatan Mahasiswa, Clustering, Silhouette Score.



1. Pendahuluan

Unit Kegiatan Mahasiswa (UKM) memainkan peran penting dalam mengembangkan soft skill mahasiswa, seperti komunikasi, kerja tim, kepemimpinan, dan kreativitas[1], yang sangat penting untuk pertumbuhan pribadi dan profesional mereka. Namun, metode manual yang saat ini digunakan untuk mengelompokkan mahasiswa berdasarkan minat mereka dalam kegiatan UKM seringkali kurang efisien dan objektif, sehingga menyebabkan manajemen dan dukungan yang kurang optimal untuk pengembangan mahasiswa. Untuk mengatasi tantangan ini, pendekatan berbasis data seperti algoritma clustering dapat digunakan untuk mengelompokkan mahasiswa secara sistematis sesuai dengan minat dan pola partisipasi mereka. Algoritma K-Means sangat cocok untuk tujuan ini karena kemampuannya untuk membagi data menjadi cluster yang bermakna berdasarkan kesamaan, sehingga memfasilitasi desain program dan alokasi sumber daya yang tepat sasaran.

Penelitian menunjukkan bahwa pengelompokan minat siswa menggunakan K-Means dapat mengungkap kelompok-kelompok berbeda dengan tingkat keterlibatan dan preferensi yang beragam [2], sehingga memungkinkan pengelola UKM untuk menyesuaikan kegiatan yang lebih sesuai dengan kebutuhan siswa dan meningkatkan hasil partisipasi. Misalnya, studi telah berhasil menerapkan K-Means Clustering untuk mengkategorikan siswa berdasarkan keterlibatan ekstrakurikuler dan minat belajar mereka, menghasilkan strategi yang lebih efektif untuk mendorong pengembangan keterampilan lunak. Selain itu, wawasan yang diperoleh dari pengelompokan dapat membantu administrator UKM mengoptimalkan proses pengambilan keputusan dengan mengidentifikasi area minat utama dan menyelarasukannya dengan tujuan institusional.

Selama ini, proses penentuan peminatan UKM umumnya masih dilakukan secara manual, seperti melalui pengisian formulir sederhana atau berdasarkan persepsi pengelola dan mahasiswa itu sendiri. Metode ini memiliki beberapa kelemahan, di antaranya bersifat subjektif, tidak terstruktur, serta sulit mengolah data dalam jumlah besar. Selain itu, pendekatan manual tidak mampu menangkap pola tersembunyi dari data minat mahasiswa yang sebenarnya dapat dimanfaatkan untuk pengambilan keputusan yang lebih akurat. Akibatnya, rekomendasi yang dihasilkan cenderung kurang tepat sasaran dan tidak mencerminkan karakteristik mahasiswa secara menyeluruh.

Untuk mengatasi permasalahan tersebut, diperlukan pendekatan berbasis data yang mampu mengelompokkan mahasiswa berdasarkan pola minat secara objektif. Salah satu metode yang dapat digunakan adalah K-Means Clustering, yang merupakan bagian dari teknik data mining untuk melakukan pengelompokan data tanpa label (unsupervised learning) [3]. Metode K-Means dipilih karena memiliki keunggulan dalam kesederhanaan algoritma, efisiensi dalam pengolahan data berukuran besar, serta kemampuannya dalam menghasilkan cluster yang mudah diinterpretasikan. Dibandingkan dengan metode clustering lainnya seperti Hierarchical Clustering atau DBSCAN, K-Means lebih cepat dalam proses komputasi dan lebih mudah diimplementasikan, sehingga cocok digunakan dalam penelitian ini. Oleh karena itu, penelitian ini bertujuan untuk menerapkan metode K-Means Clustering dalam mengelompokkan peminatan UKM mahasiswa guna mendukung pengambilan keputusan yang lebih efektif dan berbasis data.

Penerapan sistem berbasis K-Means untuk mengelompokkan minat siswa dalam UKM tidak hanya menyederhanakan proses manajemen tetapi juga mendukung terciptanya klaster representatif yang mencerminkan beragam profil siswa. Pendekatan ini pada akhirnya berkontribusi untuk memaksimalkan dampak UKM pada peningkatan soft skill dengan mempromosikan keterlibatan aktif dalam unit aktivitas yang sesuai. Oleh karena itu, penerapan K-Means Clustering menawarkan solusi yang menjanjikan untuk meningkatkan baik organisasi UKM maupun pengembangan holistik siswa melalui pengambilan keputusan berbasis data.

2. Metode Penelitian

2.1. Clustering

clustering adalah proses pengelompokan suatu pola yang belum memiliki label dan dilakukan tanpa supervisi menjadi sebuah kelompok yang memiliki karakteristik tertentu[4]. Metode pengelompokan (clustering) berguna untuk mengelompokkan objek serupa di berbagai bidang, tetapi efektivitasnya terbatas oleh ukuran kumpulan data dan kompleksitas algoritma yang digunakan[5].

2.2. K-Means Clustering

K-Means Clustering merupakan algoritma non-hierarki yang partisi data menjadi k kelompok atau klaster, di mana setiap observasi termasuk ke klaster dengan mean terdekat, berfungsi optimal pada dataset dengan atribut numerik yang terdefinisi dengan baik[6]. Metode ini iteratif, dimulai dengan penentuan jumlah klaster (K) secara acak dan

penetapan centroid awal, kemudian menghitung jarak Euclidean setiap titik data ke masing-masing centroid untuk mengelompokkan data ke klaster terdekat[7]. Selanjutnya, centroid klaster diperbarui berdasarkan rata-rata nilai data di setiap klaster, dan proses ini berulang hingga tidak ada perubahan signifikan pada penempatan klaster atau centroid, atau jumlah iterasi maksimum tercapai[8].

2.3. Unit Kegiatan Mahasiswa

Unit Kegiatan Mahasiswa adalah wadah kegiatan mahasiswa untuk mengembangkan minat, bakat, kegemaran, dan kreativitas, serta kerohanian mahasiswa di tingkat Perguruan Tinggi[9], Unit Kegiatan Mahasiswa dapat memberikan kontribusi positif terhadap keberhasilan siswa, tetapi menghadapi kendala proses dan alokasi pendanaan yang tidak memadai[10].

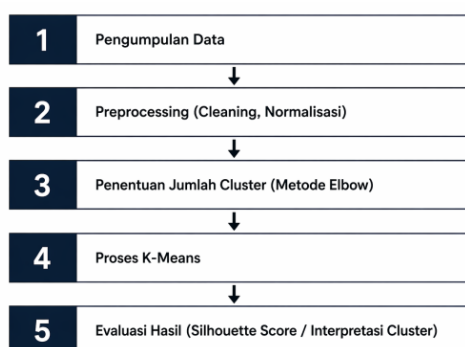
2.4. Tools dan Software

Penelitian ini menggunakan bahasa pemrograman Python sebagai alat utama dalam proses pengolahan dan analisis data. Python dipilih karena memiliki sintaks yang sederhana, fleksibel, serta didukung oleh berbagai pustaka (library) yang kuat dalam bidang data science dan machine learning. Dalam penelitian ini, beberapa library yang digunakan antara lain NumPy untuk pengolahan numerik, Pandas untuk manipulasi dan pengelolaan dataset, serta Scikit-learn untuk implementasi algoritma K-Means Clustering dan perhitungan evaluasi seperti Silhouette Score. Selain itu, digunakan juga Matplotlib untuk visualisasi data, seperti pembuatan grafik Elbow dan grafik evaluasi clustering.

Proses pengolahan data dilakukan menggunakan lingkungan pengembangan Google Colab yang memudahkan peneliti dalam menjalankan kode secara interaktif serta melakukan eksplorasi data secara bertahap. Melalui tools ini, seluruh tahapan penelitian mulai dari preprocessing data, penentuan jumlah cluster, proses clustering, hingga evaluasi hasil dapat dilakukan secara terintegrasi dan efisien.

2.5. Tahapan Penelitian

Tahapan penelitian dapat dilihat pada gambar 1. Dimulai dari pengumpulan data, processing, penentuan jumlah cluster, Proses K-Means, dan Evaluasi hasil.



Gambar 1. Tahapan Penelitian

2.5.1 Pengumpulan Data

Teknik pengumpulan data merupakan bagian yang terpenting dalam desain penelitian, karena jika judul karya ilmiah dari desain penelitian sudah disetujui untuk diteliti, maka peneliti sudah dapat mulai mengumpulkan data[11] dalam penelitian ini peneliti menggunakan data mahasiswa dan peminatan unit kegiatan mahasiswanya, data ini diperoleh dari data bidang kemasiswa bagian unit kegiatan mahasiswa yang diukur menggunakan skala numerik 1 hingga 10 untuk merepresentasikan tingkat ketertarikan mahasiswa dengan metode self assessment.

2.5.2 Preprocessing

Tahapan ini adalah untuk membersihkan, menormalkan, dan mempersiapkan data sehingga dapat diolah lebih efektif oleh algoritma analisis atau model machine learning. Dataset ini terlebih dahulu dilakukan pembersihan, perbaikan, dan transformasi tanpa merubah validitas isi yang terkandung dalam dataset.[12]

2.5.3 Penentuan Jumlah Cluster

Penentuan Cluster menggunakan metode Elbow, adalah untuk menghasilkan informasi dengancara melihat perbandingan hasil antara jumlah cluster yang akan membentuk suatu titik terakhir[13]. Pada tahap ini dilakukan perhitungan nilai Within Cluster Sum of Squares (WCSS) untuk beberapa jumlah cluster (K), dengan rumus (1):

$$WCSS = \sum_{k=1}^K \sum_{i \in ck} ||xi - ck||^2 \quad (1)$$

keterangan:

xi = data ke-i

ck = centroid cluster ke-k

Hitung jarak → lalu dikuadratkan → dijumlahkan kemudian divisualisasikan dalam bentuk grafik. Titik siku (elbow) pada grafik digunakan sebagai acuan untuk menentukan jumlah cluster yang paling optimal.

2.5.4 Penghitungan K-Means

Proses ini dimulai dengan inisialisasi centroid awal, kemudian dilanjutkan dengan perhitungan jarak antara data dan centroid menggunakan Euclidean Distance (2):

$$d = \sqrt{\sum_{i=1}^n (xi - ci)^2} \quad (2)$$

Setiap data akan dikelompokkan ke dalam cluster terdekat. Selanjutnya dilakukan pembaruan centroid berdasarkan rata-rata anggota cluster, dan proses ini diulang hingga kondisi konvergen atau tidak terjadi perubahan cluster.

2.5.5 Evaluasi Hasil

Yang dilakukan menggunakan metode Silhouette Score, dalah teknik yang menggambarkan seberapa dekat suatu objek dengan klaster nyadalam kaitannya dengan klaster lainnya[13]. dan untuk mengukur kualitas pengelompokan (3):

$$s(i) = \frac{b(i)-a(i)}{\max(a(i)),b(i)} \quad (3)$$

Keterangan:

a(i) = rata-rata jarak data ke anggota dalam cluster yang sama

b(i) = rata-rata jarak data ke cluster terdekat lainnya

Nilai s(i) antara -1 sampai 1

Nilai Silhouette digunakan untuk mengetahui seberapa baik suatu data berada dalam clusternya dibandingkan dengan cluster lain. Selain itu, dilakukan juga interpretasi terhadap setiap cluster yang terbentuk untuk memahami karakteristik masing-masing kelompok mahasiswa berdasarkan minat UKM.

3. Hasil dan Pembahasan

3.1. Deskripsi Data

Data yang digunakan dalam penelitian ini terdiri dari 20 mahasiswa yang dijadikan sebagai sampel untuk proses pengelompokan peminatan Unit Kegiatan Mahasiswa (UKM). Setiap data mahasiswa memiliki enam variabel utama, yaitu minat olahraga, minat wirausaha, minat IT Club, minat LDK, minat GAMAN (Gerakan Mahasiswa Anti Narkoba), dan minat seni, yang diukur menggunakan skala numerik 1 hingga 10 untuk merepresentasikan tingkat ketertarikan mahasiswa dengan metode self assessment oleh mahasiswa itu sendiri.

Secara umum, karakteristik data menunjukkan adanya variasi minat yang cukup signifikan antar mahasiswa, di mana sebagian mahasiswa cenderung memiliki minat tinggi pada bidang olahraga dan seni, sementara sebagian lainnya lebih dominan pada bidang akademik dan organisasi seperti IT Club, Wirausaha, LDK, dan GAMAN. Variasi ini menjadi dasar yang penting dalam proses clustering karena memungkinkan terbentuknya kelompok-kelompok mahasiswa dengan pola minat yang serupa sehingga dapat dianalisis lebih lanjut untuk mendukung pengambilan keputusan dalam pengelolaan UKM, data mahasiswa data dilihat pada tabel 1.

Tabel 1. Data Peminatan Mahasiswa

Mhs	Olahraga	Wirausaha	IT	LDK	GAMAN	Seni
Eka Dea	8	3	2	4	3	7
Aliya Sapitri	7	2	3	3	2	8
Edris Dahyuti Aprino	2	8	7	6	7	3
neni asmara	3	9	8	7	8	2
Restalia	6	4	3	5	4	6
Putri Utama	9	2	2	3	2	8
bahrul bintang	2	7	8	6	7	2
Afrilian Imelda	3	8	7	7	8	3

Mhs	Olahraga	Wirausaha	IT	LDK	GAMAN	Seni
Aulia Lathifah Sirri Latif	7	3	2	4	3	7
Dyah Chairunnisa	6	4	3	5	4	6
Sella Wulandari	2	9	8	7	8	2
Sasta Maria	3	7	7	6	7	3
Wulan Novita Sari	8	3	2	4	3	7
Elidiyah	7	2	3	3	2	8
Naina Nazwa	2	8	7	6	7	3
Valentino Mercurizky	3	9	8	7	8	2
Putri Jogil	6	4	3	5	4	6
Risky Intan Melitasari	9	2	2	3	2	8
Meika Shintia	2	7	8	6	7	2
Sarfia Karira	3	8	7	7	8	3

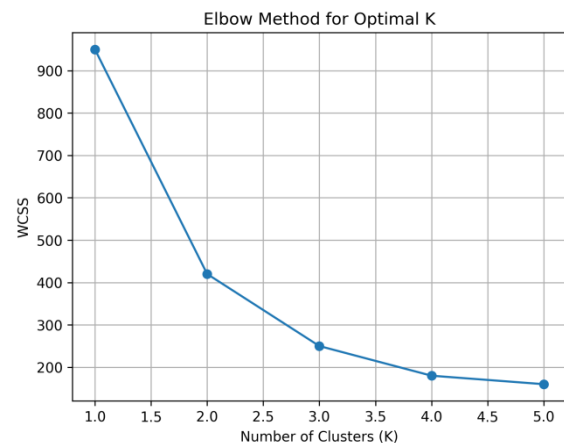
3.2. Hasil Penentuan Cluster

Metode Elbow menggunakan nilai WCSS (Within Cluster Sum of Squares), yaitu total jarak kuadrat antara data dengan centroid dalam satu cluster menggunakan rumus (1) sehingga didapatkan hasil seperti tabel 2.

Tabel 2. Data Perhitungan Cluster

Jumlah Cluster (K)	Nilai WCSS
1	950
2	420
3	250
4	180
5	160

Sehingga dapat divisualisasikan pada gambar 2.



Gambar 2. Hasil Metode Elbow

Dari gambar dan tabel diatas dapat disimpulkan titik siku adalah cluster ke-2.

3.3. K-Means Clustering

Tahapan awal dari metode clustering adalah menentukan centroid awal yang sipilih secara acak, disini peneliti memilih Eka Dea, dan Edris Dahyuti Aprino sebagai centroid awal, seperti pada tabel 3.

Tabel 3. Data Centroid Awal

Mhs	Olahraga	Wirausaha	IT	LDK	GAMAN	Seni
Eka Dea	8	3	2	4	3	7
Edris Dahyuti Aprino	2	8	7	6	7	3

Tahapan selanjutnya adalah menghitung jarak seluruh data dengan menggunakan rumus (2)

Aliya Sapitri ke Centroid 1

$$\sqrt{((7 - 8)^2 + (2 - 3)^2 + (3 - 2)^2 \dots + (8 - 7)^2)}$$

$$\sqrt{(1 + 1 + 1 + 1 + 1 + 1)}$$

$$\sqrt{6}$$

2,45

Aliya Sapitri ke Centroid 2

$$\sqrt{((7 - 2)^2 + (2 - 8)^2 + (3 - 7)^2 \dots + (8 - 3)^2)}$$

$$\sqrt{(25 + 36 + 16 + 9 + 25 + 25)}$$

$$\sqrt{136}$$

11,66

Dari perhitungan diatas dapat diperoleh kesimpulan aliya sapitri masuk ke Centroid 1 karena lebih minimum jaraknya, dari seluruh data dihitung jaraknya sehingga didapatkan hasil seperti pada tabel 4.

Tabel 4. Hasil Data Centroid

Mhs	Centroid 1	Centroid 2	Hasil
Eka Dea	0	11.05	C1
Aliya Sapitri	2.45	11.66	C1
Edris Dahyuti Aprino	11.05	0	C2
neni asmara	12.49	2.45	C2
Restalia	3	8.19	C1
Putri Utama	2.24	13	C1
bahrul bintang	11.53	1.73	C2
Afrilian Imelda	11.18	1.73	C2
Aulia Lathifah Sirri Latif	1	10.54	C1
Dyah Chairunnisa	3	8.19	C1
Sella Wulandari	12.92	2.24	C2
Sasta Maria	10.10	1.41	C2
Wulan Novita Sari	0	11.05	C1
Elidiyah	2.45	11.66	C1
Naina Nazwa	11.05	0	C2
Valentino	12.49	2.45	C2
Mercurizky			
Putri Jogil	3	8.19	C1

Mhs	Centroid 1	Centroid 2	Hasil
Risky Intan Melitasari	2.24	13	C1
Meika Shintia	11.53	1.73	C2
Sarfia Karira	11.18	1.73	C2

Sehingga dari tabel diatas didapatkan hasil klasteriterisasi (Iterasi 1), kemudian langkah selanjutnya adalah menghitung centroid baru

Centroid C1 Baru

Rata-rata (10 data):

$$\text{Olahraga} = (8+7+6+9+7+6+8+7+6+9)/10 = 7.3$$

$$\text{Wirausaha} = (3+2+4+2+3+4+3+2+4+2)/10 = 2.9$$

$$\text{IT Club} = (2+3+3+2+2+3+2+3+3+2)/10 = 2.5$$

$$\text{LDK} = (4+3+5+3+4+5+4+3+5+3)/10=4.0$$

$$\text{GAMAN} = (3+2+4+2+3+4+3+2+4+2)/10= 2.9$$

$$\text{Seni} = (7+8+6+8+7+6+7+8+6+8)/10= 7.3$$

Sehingga didapatkan hasil C1 = (7.3, 2.9, 2.5, 4.0, 2.9, 7.3)

Centroid C2 Baru

Rata-rata (10 data):

$$\text{Olahraga} = (2+3+2+3+2+3+2+3+2+3)/10 = 2.5$$

$$\text{Wirausaha} = (8+9+7+8+9+7+8+9+7+8)/10 = 8.0$$

$$\text{IT Club} = (7+8+8+7+8+7+7+8+8+7)/10 = 7.5$$

$$\text{LDK} = (6+7+6+7+7+6+6+7+6+7)/10=6.5$$

$$\text{GAMAN} = (7+8+7+8+8+7+7+8+7+8)/10=7.5$$

$$\text{Seni} = (3+2+2+3+2+3+3+2+2+3)/10= 2.5$$

3.4 Interpretasi Hasil

Cluster 1:

Tinggi: Olahraga & Seni

Cluster 2:

Tinggi: Wirausaha, IT, LDK, GAMAN

3.5 Evaluasi Hasil

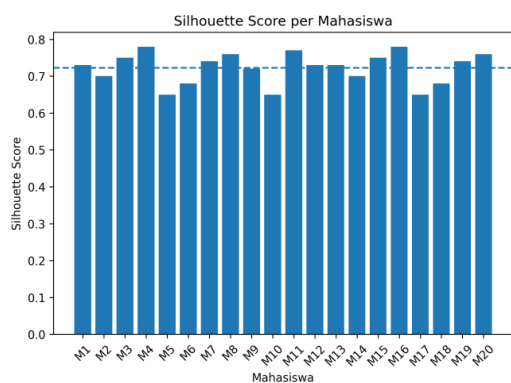
Evaluasi dilakukan dengan Silhouette Score, Silhouette Score mengukur seberapa baik suatu data berada dalam cluster-nya dibandingkan dengan cluster lain, perhitungan dilakukan dengan rumus (3) sehingga didapatkan ahasil seperti padda tabel 5.

Tabel 5. Hasil Silhouette Score

Mhs	Nilai s(i)
Eka Dea	0.73
Aliya Sapitri	0.7
Edris Dahyuti Aprino	0.75
neni asmara	0.78

Mhs	Nilai s(i)
Restalia	0.65
Putri Utama	0.68
bahrul bintang	0.74
Afrilian Imelda	0.76
Aulia Lathifah Sirri Latif	0.72
Dyah Chairunnisa	0.65
Sella Wulandari	0.77
Sasta Maria	0.73
Wulan Novita Sari	0.73
Elidiyah	0.7
Naina Nazwa	0.75
Valentino Mercurizky	0.78
Putri Jogil	0.65
Risky Intan Melitasari	0.68
Meika Shintia	0.74
Sarfia Karira	0.76

Jika divisualisasikan maka hasilnya adalah seperti pada gambar 3.



Gambar 3. Grafik Silhouette Score

Grafik Silhouette Score menunjukkan nilai evaluasi untuk masing-masing mahasiswa terhadap hasil clustering yang terbentuk. Garis putus-putus merepresentasikan nilai rata-rata Silhouette Score sebesar ± 0.73 . Secara umum, seluruh data memiliki nilai di atas 0.6 yang menunjukkan bahwa setiap objek telah terkelompok dengan baik dan memiliki tingkat kesamaan yang tinggi dalam cluster-nya dibandingkan dengan cluster lain. Hal ini mengindikasikan bahwa metode K-Means menghasilkan kualitas clustering yang cukup baik.

4. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, metode K-Means Clustering terbukti mampu mengelompokkan mahasiswa berdasarkan pola minat terhadap Unit Kegiatan Mahasiswa (UKM) secara efektif. Hasil pengolahan data menunjukkan bahwa terbentuk dua cluster utama dengan karakteristik yang berbeda, yaitu kelompok mahasiswa yang dominan pada minat olahraga dan seni, serta kelompok mahasiswa yang lebih cenderung pada minat wirausaha, IT Club, LDK, dan GAMAN. Nilai evaluasi menggunakan Silhouette

Score yang tinggi juga menunjukkan bahwa kualitas clustering yang dihasilkan tergolong cukup baik, sehingga dapat merepresentasikan kondisi data secara akurat.

Selain itu, hasil clustering yang diperoleh dapat dimanfaatkan sebagai dasar dalam pengambilan keputusan oleh pihak kampus, khususnya dalam pengelolaan dan pengembangan UKM. Dengan adanya pengelompokan ini, pihak pengelola dapat lebih mudah dalam menentukan strategi pembinaan, penyaluran minat mahasiswa, serta perencanaan program kerja yang lebih tepat sasaran. Pendekatan berbasis data ini juga membantu meningkatkan efisiensi dalam proses pengelolaan UKM dibandingkan dengan metode konvensional yang bersifat subjektif.

Adapun saran untuk penelitian selanjutnya adalah dengan menambahkan variabel lain yang lebih beragam, seperti aspek psikologis, tingkat keaktifan mahasiswa, latar belakang akademik, serta faktor lingkungan sosial. Selain itu, penelitian berikutnya dapat membandingkan metode K-Means dengan algoritma clustering lainnya seperti DBSCAN atau Hierarchical Clustering untuk memperoleh hasil yang lebih optimal. Pengembangan sistem berbasis aplikasi juga disarankan agar hasil clustering dapat diimplementasikan secara langsung dalam sistem rekomendasi peminatan UKM di lingkungan perguruan tinggi.

Daftar Rujukan

- [1] S. Yolanda, S. M. Sari, I. Ismail, and U. T. Madura, "PERAN ORGANISASI MAHASISWA DALAM MEMBANGUN KARAKTER KEPEMIMPINAN DAN PENINGKATAN SOFT SKILL," vol. 09, 2024.
- [2] A. C. Virgifiyani, E. Nurfajrina, D. S. Aulia, M. H. Dirmansyah, and M. Kom, "ANALISIS PENGARUH GAMIFIKASI DAN KENYAMANAN APLIKASI DUOLINGO TERHADAP EFEKTIVITAS PEMBELAJARAN BAHASA MENGGUNAKAN METODE K-MEANS CLUSTERING," 2026.
- [3] A. D. Ramadani, F. Hilmy Ibrahim, M. Hidayat, A. Habibullah, S. Sumanto, and A. D. Kuswanto, "Klasterisasi Data Produksi Daging Sapi Menggunakan Algoritma K-Means Orange Data Mining," *Jurnal Pustaka Data*, vol. 5, no. 1, pp. 109–114, Jun. 2025, doi: 10.55382/jurnalpustakadata.v5i1.1013.
- [4] S. Paembonan and H. Abduh, "Penerapan Metode Silhouette Coeficient Untuk Evaluasi Clutering Obat," *PENA TEKNIK: Jurnal Ilmiah Ilmu-Ilmu Teknik*, vol. 6, no. 2, 2021.
- [5] T. S. Madhulatha, "AN OVERVIEW ON CLUSTERING METHODS," *IOSRJEN*, vol. 02, no. 04, pp. 719–725, Apr. 2012, doi: 10.9790/3021-0204719725.
- [6] M. H. Fakhriza and K. Umam, "ANALISIS PRODUK TERLARIS MENGGUNAKAN METODE K-MEANS CLUSTERING PADA "PT.SUKANDA DJAYA," *JIKA*, vol. 5, no. 1, p. 8, Apr. 2021, doi: 10.31000/jika.v5i1.3236.
- [7] A. Zaki, I. Irwan, and I. A. Sembe, "Penerapan K-Means Clustering dalam Pengelompokan Data (Studi Kasus Profil Mahasiswa Matematika FMIPA UNM)," *JMathCoS*, vol. 5,

- no. 2, p. 163, Oct. 2022, doi: 10.35580/jmathcos.v5i2.38820.
- [8] A. Asmana, Y. Arie Wijaya, and M. Martanto, "CLUSTERING DATA CALON SISWA BARU MENGGUNAKAN METODE K-MEANS DI SEKOLAH MENENGAH KEJURUAN WAHIDIN KOTA CIREBON," *jati*, vol. 6, no. 2, pp. 552–559, Aug. 2022, doi: 10.36040/jati.v6i2.5236.
- [9] I. Budi Santoso, "PENGARUH KEAKTIFAN ORGANISASI DAN MOTIVASI BELAJAR TERHADAP PRESTASI BELAJAR PENGURUS UNIT KEGIATAN MAHASISWA UNIVERSITAS NEGERI YOGYAKARTA PERIODE 2019-2020 .," *Jurnal Ilmu Manajemen*, vol. 16, no. 1, 2019.
- [10] B. Sugito, J. D. Sasanadjati, A. Hidajad, and B. D. Putra, "Komunitas dan Unit Kegiatan Mahasiswa Dalam Proses Mencapai Prestasi Non Akademik," *JEHSS*, vol. 6, no. 4, pp. 1582–1592, May 2024, doi: 10.34007/jehss.v6i4.1763.
- [11] S. Hamidani and R. Yanto, "Sistem Cerdas Pemilihan Makanan Sehat Berbasis Case-Based Reasoning dan SMART untuk Edukasi Pemenuhan Gizi Masyarakat," *Jurnal Pustaka AI*, vol. 5, no. 2, Aug. 2023, doi: 10.55382/jurnalpustakaai.v5i2.1209.
- [12] F. Putra, H. F. Tahiyat, R. M. Ihsan, R. Rahmaddeni, and L. Efrizoni, "Penerapan Algoritma K-Nearest Neighbor Menggunakan Wrapper Sebagai Preprocessing untuk Penentuan Keterangan Berat Badan Manusia: Application of K-Nearest Neighbor Algorithm Using Wrapper as Preprocessing for Determination of Human Weight Information," *MALCOM*, vol. 4, no. 1, pp. 273–281, Jan. 2024, doi: 10.57152/malcom.v4i1.1085.
- [13] R. Y. Sari, H. Oktavianto, and H. W. Sulistyono, "ALGORITMA K-MEANS DENGAN METODE ELBOW UNTUK MENGELOMPOKKAN KABUPATEN/KOTA DI JAWA TENGAH BERDASARKAN KOMPONEN PEMBENTUK INDEKS PEMBANGUNAN MANUSIA".