

## SEGMENTASI KEAKTIFAN MAHASISWA UNIVERSITAS ISLAM NEGERI SUMATERA UTARA DALAM KEGIATAN KAMPUS MENGUNAKAN K-MEANS CLUSTERING

Nazwa Aliya Muthmainnah Hasibuan<sup>1</sup>, Dodyk Fahlome<sup>2</sup>, Putri Salsa Nabila<sup>3</sup>,  
Said Arrahman<sup>4</sup>, Mhd. Furqan<sup>5</sup>

<sup>1-5</sup> Program Studi Ilmu Komputer, Universitas Islam Negeri Sumatera Utara, Indonesia  
Email: nazwa0701232104@uinsu.ac.id1, dodyk0701232093@uinsu.ac.id2, putri0701232090@uinsu.ac.id3,  
said0701232091@uinsu.ac.id4, mfurqan@uinsu.ac.id5  
\*Penulis Korespondensi: nazwa0701232104@uinsu.ac.id

### ABSTRAK

Kegiatan kemahasiswaan berperan penting dalam pengembangan kompetensi mahasiswa, namun tingkat keaktifan pada berbagai aktivitas seperti organisasi, seminar, kepanitiaan, lomba, dan pengembangan diri menunjukkan variasi yang signifikan sehingga diperlukan pendekatan berbasis data untuk mengidentifikasi pola keterlibatan secara lebih objektif. Penelitian ini menerapkan K-Means Clustering pada data 100 responden mahasiswa UINSU yang diperoleh melalui Google Forms, melalui tahapan preprocessing, konversi skala ordinal, serta analisis menggunakan Python (Google Colab). Jumlah cluster optimal ditentukan menggunakan metode Elbow berbasis Within Cluster Sum of Squares (WCSS). Hasil penelitian menunjukkan terbentuk tiga cluster ( $k=3$ ), yaitu C0 (30 mahasiswa) dengan karakteristik aktif organisasi dan kepanitiaan yang ditandai skor panitia 2.43 dan organisasi 1.63, C1 (38 mahasiswa) sebagai kelompok sangat aktif/multitalenta dengan dominasi pengembangan diri 2.03 dan lomba 1.76, serta C2 (32 mahasiswa) sebagai kelompok kurang aktif dengan skor terendah pada organisasi 0.31 dan lomba 0.47. Visualisasi PCA memperkuat pemisahan cluster yang terbentuk, sehingga menunjukkan bahwa K-Means efektif dalam mengungkap heterogenitas tingkat keaktifan mahasiswa dan dapat digunakan sebagai dasar pengambilan keputusan berbasis data dalam pengelolaan program kemahasiswaan.

**Kata Kunci**— Kegiatan Kampus, Keaktifan Mahasiswa, Klasterisasi; K-Means, Segmentasi

### ABSTRACT

*Student activities play a crucial role in developing students' competencies; however, participation levels in various activities—such as student organizations, seminars, event committees, competitions, and personal development—show significant variation, necessitating a data-driven approach to identify patterns of engagement more objectively. This study applied K-Means Clustering to data from 100 UINSU student respondents collected via Google Forms, through stages of preprocessing, ordinal scale conversion, and analysis using Python (Google Colab). The optimal number of clusters was determined using the Elbow method based on the Within Cluster Sum of Squares (WCSS). The results indicate the formation of three clusters ( $k=3$ ): C0 (30 students) characterized by active involvement in organizations and committees, marked by a committee score of 2.43 and an organizational score of 1.63; C1 (38 students) as a highly active/multitalented group dominated by personal development (2.03) and competitions (1.76), and C2 (32 students) as a less active group with the lowest scores in organizational activities (0.31) and competitions (0.47). PCA visualization reinforces the separation of the formed clusters, indicating that K-Means is effective in revealing the heterogeneity of student activity levels and can serve as a basis for data-driven decision-making in the management of student programs.*

**Keywords**— Campus Activities, Clustering, K-Means, Segmentation, Student Activity

## I. PENDAHULUAN

Di lingkungan perguruan tinggi, kegiatan kemahasiswaan menjadi salah satu aspek penting dalam mendukung pengembangan kompetensi mahasiswa secara menyeluruh [1]. Selain kegiatan akademik di dalam kelas, mahasiswa juga dituntut untuk aktif dalam berbagai aktivitas seperti organisasi, seminar, kepanitiaan, lomba, serta program pengembangan diri. Keterlibatan dalam kegiatan tersebut tidak hanya memberikan pengalaman praktis, tetapi juga berkontribusi dalam pembentukan kemampuan kepemimpinan, komunikasi, serta manajemen waktu yang menjadi nilai tambah bagi mahasiswa di dunia kerja [2].

Namun demikian, tingkat keaktifan mahasiswa dalam mengikuti kegiatan kampus menunjukkan variasi yang cukup signifikan [3]. Sebagian mahasiswa aktif terlibat dalam berbagai kegiatan, sementara sebagian lainnya cenderung memiliki tingkat partisipasi yang terbatas. Kondisi ini menyebabkan sulitnya memperoleh gambaran yang jelas mengenai pola keterlibatan mahasiswa secara keseluruhan. Tanpa adanya pengelompokan yang sistematis, pihak kampus akan mengalami kendala dalam memahami karakteristik mahasiswa serta dalam merancang program kegiatan yang sesuai dengan kebutuhan mereka.

Seiring dengan meningkatnya pemanfaatan teknologi informasi, data yang berkaitan dengan aktivitas mahasiswa sebenarnya dapat diolah untuk menghasilkan informasi yang lebih bermakna [4]. Data tersebut dapat digunakan untuk mengidentifikasi pola keaktifan mahasiswa sehingga menghasilkan segmentasi yang lebih terstruktur. Pendekatan ini memungkinkan analisis dilakukan secara objektif dan berbasis data, dibandingkan dengan pengamatan manual yang cenderung bersifat subjektif dan kurang representatif.

Salah satu pendekatan yang dapat diterapkan untuk melakukan pengelompokan data adalah dengan menggunakan algoritma K-Means yang dikombinasikan dengan metode Elbow untuk menentukan jumlah cluster yang optimal [5]. Selain itu, hasil *clustering* dapat divisualisasikan menggunakan *Principal Component Analysis* (PCA) untuk mempermudah interpretasi pola penyebaran data [6]. Dengan demikian, penerapan metode ini memungkinkan data mahasiswa dikelompokkan berdasarkan tingkat keaktifan mereka secara lebih jelas dan sistematis.

Beberapa penelitian sebelumnya telah menerapkan K-Means untuk pengelompokan data mahasiswa. [7] menerapkan K-Means pada pengelompokan indeks prestasi akademik mahasiswa dan berhasil

mengategorikan mahasiswa ke dalam dua kluster, yaitu kategori baik dan tidak baik. [8] menggabungkan K-Means dengan Random Forest untuk memprediksi kelulusan mahasiswa tepat waktu, dengan mengelompokkan mahasiswa berdasarkan IPK, jumlah kredit, dan lama studi. [9] melakukan klusterisasi kinerja akademik mahasiswa menggunakan K-Means dengan metode Elbow untuk menentukan jumlah kluster optimal, menghasilkan tiga kategori kinerja: tinggi, sedang, dan rendah. Meskipun berbagai penelitian tersebut telah berhasil menerapkan K-Means, mayoritas fokus pada aspek performa akademik atau kelulusan mahasiswa. Penelitian yang berfokus pada segmentasi keaktifan organisasi non-akademik dengan optimasi jumlah cluster menggunakan metode Elbow dan visualisasi pola penyebaran data melalui PCA masih sangat terbatas, sehingga menjadi celah penelitian yang perlu diisi.

Berdasarkan permasalahan tersebut, penelitian ini dilakukan untuk melakukan segmentasi keaktifan mahasiswa Universitas Islam Negeri Sumatera Utara (UINSU) dalam kegiatan kampus menggunakan algoritma K-Means Clustering. Hasil dari penelitian ini diharapkan dapat membantu pihak kampus dalam memahami pola keterlibatan mahasiswa serta menjadi dasar dalam pengambilan kebijakan yang lebih tepat sasaran dalam pengembangan kegiatan kemahasiswaan.

## II. KAJIAN TEORITIS

### A. *Clustering*

*Clustering* merupakan teknik dalam *data mining* untuk mengelompokkan data berdasarkan tingkat kemiripan karakteristik tertentu [10]. Data dengan pola atau atribut serupa akan berada dalam satu cluster, sedangkan data dengan perbedaan signifikan ditempatkan pada cluster yang berbeda, sehingga memungkinkan identifikasi struktur alami tanpa memerlukan label sebelumnya.

Dalam penerapannya, *clustering* banyak digunakan untuk segmentasi data di berbagai bidang, termasuk pendidikan [11]. Analisis data keaktifan mahasiswa dapat mengungkap pola keterlibatan secara alami, sehingga menghasilkan informasi yang lebih terstruktur dan mendukung evaluasi serta pengambilan keputusan yang lebih tepat.

### B. *Algoritma K-Means*

Algoritma K-Means merupakan salah satu metode *clustering* yang banyak digunakan karena kesederhanaan konsep serta efisiensinya dalam mengelompokkan data [5]. Metode ini bekerja dengan membagi data ke dalam sejumlah cluster berdasarkan kedekatan antara data dengan centroid [12], di mana setiap data akan ditempatkan pada cluster dengan jarak

terdekat sehingga terbentuk kelompok yang relatif homogen.

Proses K-Means dilakukan secara iteratif melalui tahapan penentuan jumlah cluster, inisialisasi centroid awal, pengelompokan data berdasarkan jarak terdekat, serta pembaruan centroid hingga mencapai kondisi konvergen [13]. Dalam proses tersebut, perhitungan jarak antara data dan centroid menjadi komponen utama yang umumnya menggunakan Euclidean Distance sebagai berikut:

$$d(x_i, c_j) = \sqrt{\sum_{k=1}^n (x_{ik} - c_{jk})^2}$$

Keterangan:

- a.  $d(x_i, c_j)$  = jarak antara data ke- $i$  dan centroid ke- $j$
- b.  $x_{ik}$  = nilai atribut ke- $k$  dari data ke- $i$
- c.  $c_{jk}$  = nilai atribut ke- $k$  dari centroid ke- $j$
- d.  $k$  = indeks atribut (variabel)
- e.  $n$  = jumlah atribut

### C. Metode Elbow

Metode Elbow digunakan untuk menentukan jumlah cluster optimal dengan membandingkan variasi dalam cluster pada beberapa jumlah cluster yang berbeda [14]. Pendekatan ini membantu menghindari pemilihan cluster yang terlalu sedikit atau terlalu banyak sehingga hasil pengelompokan lebih representatif terhadap karakteristik data.

Penentuan jumlah cluster dilakukan dengan mengamati penurunan nilai *Within Cluster Sum of Squares* (WCSS) yang divisualisasikan dalam grafik, di mana titik elbow menunjukkan jumlah cluster optimal saat penurunan mulai melambat [15]. Nilai WCSS mencerminkan tingkat penyebaran data dalam cluster yang semakin kecil nilainya, semakin baik kualitas pengelompokan. Perhitungan WCSS dapat dirumuskan sebagai berikut:

$$WCSS = \sum_{j=1}^K \sum_{i=1}^n (x_i - c_j)^2$$

Keterangan:

- a. WCSS= total variasi dalam cluster
- b.  $K$ = jumlah cluster
- c.  $x_i$ = data ke- $i$
- d.  $c_j$ = centroid ke- $j$
- e.  $n$ = jumlah data

### D. Principal Component Analysis (PCA)

*Principal Component Analysis* (PCA) merupakan teknik analisis data yang digunakan untuk mereduksi dimensi data dengan mengubah sejumlah variabel yang saling berkorelasi menjadi beberapa komponen utama yang tidak berkorelasi [6]. Teknik ini bertujuan untuk menyederhanakan struktur data tanpa menghilangkan informasi penting sehingga mempermudah proses analisis.

Dalam konteks *clustering*, PCA digunakan untuk memproyeksikan data berdimensi tinggi ke dalam dua atau tiga dimensi agar pola penyebaran cluster dapat diamati dengan lebih jelas [16]. Dengan demikian, PCA membantu dalam menginterpretasikan hasil pengelompokan data secara lebih intuitif dan informatif.

### E. Penelitian Terdahulu

Penelitian terdahulu menunjukkan bahwa teknik *clustering*, khususnya algoritma K-Means, efektif digunakan untuk menghasilkan segmentasi data yang terstruktur dalam bidang pendidikan. Metode ini mampu mengelompokkan data berdasarkan tingkat kemiripan karakteristik sehingga memudahkan proses analisis pola. Penerapannya pada data mahasiswa mampu mengelompokkan gaya hidup ke dalam beberapa kategori sehingga memudahkan identifikasi pola perilaku serta memahami kecenderungan aktivitas mahasiswa secara lebih sistematis [17].

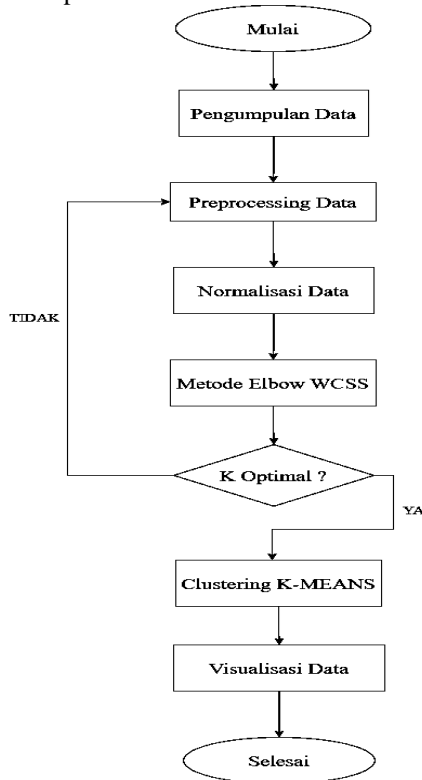
Selain itu, kombinasi K-Means dengan metode Elbow terbukti mampu menentukan jumlah cluster optimal dalam pengelompokan data pendidikan [18]. Metode *clustering* juga dapat mengelompokkan wilayah berdasarkan tingkat kualitas pendidikan sehingga membantu mengidentifikasi kesenjangan antar daerah serta memberikan dasar yang lebih kuat dalam mendukung pengambilan kebijakan berbasis data [19].

## III. METODE PENELITIAN

Penelitian ini merupakan penelitian kuantitatif dengan pendekatan analisis data menggunakan teknik *clustering*. Pendekatan ini digunakan karena penelitian berfokus pada pengolahan data numerik yang diperoleh dari hasil kuesioner untuk menghasilkan segmentasi mahasiswa berdasarkan tingkat keaktifan dalam kegiatan kampus secara objektif dan terstruktur. Tahapan penelitian meliputi pengumpulan data, *preprocessing*, normalisasi, penentuan jumlah cluster optimal menggunakan metode Elbow, proses *clustering* menggunakan algoritma K-Means, serta visualisasi hasil menggunakan *Principal Component Analysis* (PCA).

Penelitian ini dilaksanakan melalui serangkaian tahapan yang terstruktur untuk menghasilkan segmentasi mahasiswa berdasarkan tingkat keaktifan dalam kegiatan kampus. Adapun alur pelaksanaan

penelitian mulai dari pengumpulan data, pengolahan data, proses clustering, hingga visualisasi hasil ditunjukkan pada Gambar 1.



Gbr. 1 Diagram Alur Proses Tahapan Penelitian

### A. Pengumpulan Data

Pengumpulan data dalam penelitian ini dilakukan menggunakan kuesioner berbasis Google Forms sebagai alat bantu survei untuk memperoleh informasi terkait keaktifan mahasiswa dalam kegiatan kampus selama satu semester terakhir. Responden dalam penelitian ini berjumlah 100 mahasiswa Universitas Islam Negeri Sumatera Utara (UINSU) yang dipilih menggunakan teknik *purposive sampling*, dengan kriteria mahasiswa aktif yang telah mengikuti kegiatan kampus minimal selama satu semester terakhir.

Skala pengukuran yang digunakan adalah skala ordinal yang dikonversi menjadi nilai numerik, yaitu 0 untuk tidak pernah, 1 untuk pernah, 2 untuk cukup sering, dan 3 untuk sangat sering. Meskipun skala yang digunakan bersifat ordinal, nilai tersebut diperlakukan sebagai data numerik karena merepresentasikan tingkat intensitas keaktifan mahasiswa, sehingga masih relevan untuk digunakan dalam perhitungan jarak pada algoritma K-Means [20]. Data yang terkumpul selanjutnya diperiksa untuk memastikan tidak terdapat data kosong (*missing value*), sehingga siap untuk diproses pada tahap berikutnya.

### B. Variabel dan Analisis Data

Variabel yang digunakan dalam penelitian ini terdiri dari lima aspek keaktifan mahasiswa, yaitu organisasi, seminar, kepanitiaan, lomba, dan kegiatan pengembangan diri. Kelima variabel tersebut digunakan untuk merepresentasikan tingkat keterlibatan mahasiswa dalam berbagai kegiatan kampus. Adapun variabel penelitian dapat dilihat pada Tabel 1.

TABEL I. VARIABEL DAN INDIKATOR PENELITIAN

Variabel	Indikator
Organisasi	Keaktifan struktural
Seminar	Keaktifan ilmiah
Kepanitiaan	Keaktifan kontribusi
Lomba	Keaktifan kompetitif
Pengembangan Diri	Keaktifan akademik non-kelas

Data yang telah dikumpulkan kemudian melalui tahap *preprocessing* untuk memastikan seluruh data berada dalam bentuk numerik dan siap dianalisis. Proses pembersihan data serta konversi data ke dalam bentuk numerik dilakukan menggunakan Microsoft Excel. Selanjutnya dilakukan proses normalisasi data menggunakan metode *Min-Max Scaling* untuk menyamakan rentang nilai antar variabel sehingga tidak terjadi bias dalam perhitungan jarak pada algoritma K-Means [13].

Proses analisis data dalam penelitian ini meliputi beberapa tahapan, yaitu penentuan jumlah cluster optimal menggunakan metode Elbow dengan pendekatan *Within Cluster Sum of Squares* (WCSS) [15], penerapan algoritma K-Means untuk proses pengelompokan data, serta visualisasi hasil *clustering* menggunakan *Principal Component Analysis* (PCA). Seluruh tahapan tersebut dilakukan menggunakan platform Google Colab berbasis bahasa pemrograman Python, yang mendukung pengolahan data secara efisien serta memudahkan dalam eksplorasi, pemodelan, dan interpretasi hasil analisis.

Penentuan jumlah cluster optimal dilakukan dengan menghitung nilai WCSS untuk beberapa variasi jumlah cluster, kemudian divisualisasikan dalam bentuk grafik untuk menentukan titik optimal. Proses *clustering* dilakukan secara iteratif yang diawali dengan inisialisasi centroid awal secara acak, dilanjutkan dengan pengelompokan data berdasarkan jarak Euclidean, serta pembaruan centroid hingga mencapai kondisi konvergen [13]. Hasil *clustering* selanjutnya divisualisasikan menggunakan PCA untuk mempermudah interpretasi pola penyebaran cluster [16].

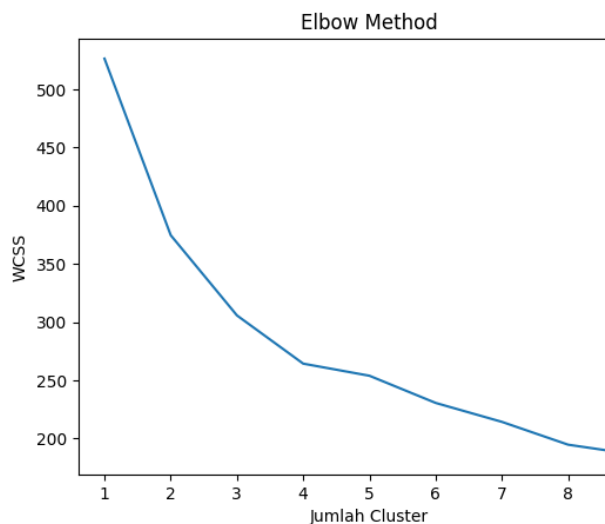
## IV. HASIL DAN PEMBAHASAN

Data dalam penelitian ini diperoleh dari 100 responden mahasiswa Universitas Islam Negeri

Sumatera Utara (UINSU) melalui penyebaran kuesioner berbasis Google Form. Data yang dikumpulkan mencakup tingkat keaktifan mahasiswa dalam kegiatan organisasi, seminar, kepanitiaan, lomba, dan pengembangan diri selama satu semester terakhir. Seluruh data yang diperoleh telah melalui proses *preprocessing*, di mana setiap jawaban dikonversi ke dalam bentuk numerik sesuai dengan skala yang telah ditentukan sehingga siap untuk dianalisis menggunakan metode *clustering*, serta memastikan data lebih terstruktur dan mudah diolah dalam proses analisis selanjutnya.

#### A. Penentuan Jumlah Cluster dengan Metode Elbow

Penentuan jumlah cluster optimal dalam penelitian ini dilakukan menggunakan metode Elbow dengan menghitung nilai *Within Cluster Sum of Squares* (WCSS) pada beberapa variasi jumlah cluster [15]. Metode ini digunakan untuk mengidentifikasi jumlah cluster yang tepat berdasarkan pola penurunan nilai WCSS seiring bertambahnya jumlah cluster, seperti yang ditunjukkan pada Gambar 1



Gbr. 2 Grafik Metode Elbow

Berdasarkan Gambar 1, terlihat bahwa nilai WCSS mengalami penurunan yang signifikan hingga jumlah cluster  $k = 3$ , kemudian setelah itu penurunannya cenderung melambat. Hal ini menunjukkan bahwa penambahan jumlah cluster di atas  $k = 3$  tidak memberikan pengurangan variasi yang signifikan. Oleh karena itu, titik elbow diidentifikasi pada  $k = 3$ , sehingga jumlah cluster yang digunakan dalam penelitian ini ditetapkan sebanyak 3 cluster karena dinilai mampu merepresentasikan struktur data secara optimal dengan tingkat kompleksitas yang efisien.

#### B. Hasil Clustering Menggunakan K-Means

Proses *clustering* dilakukan menggunakan algoritma K-Means dengan jumlah cluster sebanyak 3 berdasarkan hasil dari metode Elbow. Hasil pengelompokan ini menghasilkan segmentasi mahasiswa berdasarkan tingkat keaktifan dalam berbagai kegiatan kampus. Distribusi jumlah anggota pada masing-masing cluster disajikan pada Tabel 2.

TABEL II. JUMLAH ANGGOTA SETIAP CLUSTER

Cluster	Jumlah Mahasiswa
C0	30
C1	38
C2	32

Berdasarkan hasil *clustering* pada Tabel 2, distribusi jumlah mahasiswa pada masing-masing cluster menunjukkan bahwa cluster C1 memiliki jumlah anggota terbanyak yaitu 38 mahasiswa, diikuti oleh cluster C2 sebanyak 32 mahasiswa, dan cluster C0 sebanyak 30 mahasiswa. Distribusi yang relatif seimbang ini menunjukkan bahwa keterlibatan mahasiswa tidak terpusat pada satu kelompok tertentu, melainkan tersebar pada berbagai tingkat partisipasi. Hal ini mengindikasikan bahwa aktivitas mahasiswa dalam kegiatan kampus bersifat heterogen, sehingga diperlukan pendekatan program yang berbeda untuk setiap kelompok. Untuk mengetahui karakteristik masing-masing cluster, dilakukan perhitungan nilai rata-rata pada setiap variabel, seperti yang ditampilkan pada Tabel 3.

TABEL III. RATA-RATA NILAI SETIAP CLUSTER

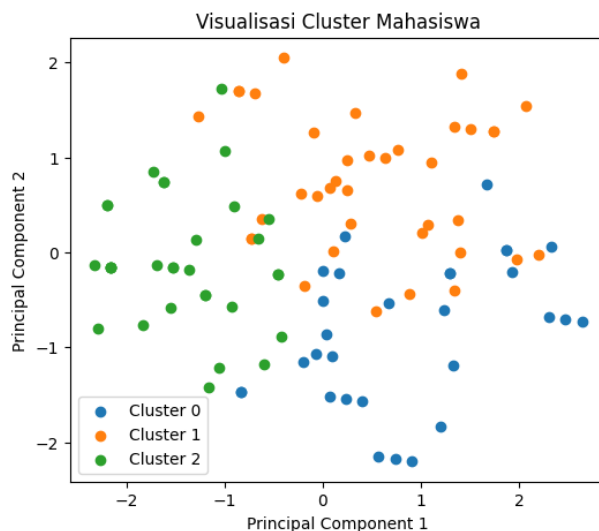
Kategori	C0	C1	C2
Organisasi	1.63	1.74	0.31
Seminar	1.10	1.71	1.38
Panitia	2.43	0.84	0.62
Lomba	1.17	1.76	0.47
Pengembangan Diri	1.67	2.03	0.94

Berdasarkan Tabel 3, terlihat bahwa setiap cluster memiliki karakteristik yang berbeda pada masing-masing variabel keaktifan. Cluster C0 menunjukkan nilai tinggi pada variabel kepanitiaan, yang mengindikasikan keterlibatan aktif dalam kegiatan organisasi sebagai panitia. Cluster C1 memiliki nilai yang relatif tinggi pada hampir semua variabel, terutama pada pengembangan diri dan seminar, yang menunjukkan keterlibatan yang cukup merata pada berbagai kegiatan. Sementara itu, Cluster C2 memiliki nilai rata-rata yang cenderung rendah pada seluruh variabel, yang mengindikasikan tingkat partisipasi yang lebih rendah dibandingkan cluster lainnya. Perbedaan karakteristik antar cluster tersebut selanjutnya

divisualisasikan untuk melihat pola penyebaran data secara lebih jelas.

### C. Visualisasi dan Interpretasi Hasil Clustering

Visualisasi hasil *clustering* dilakukan menggunakan *Principal Component Analysis* (PCA) untuk mereduksi dimensi data ke dalam dua komponen utama sehingga pola penyebaran data dapat lebih mudah diamati [16], seperti yang ditunjukkan pada Gambar 2.



Gbr. 3 Visualisasi Cluster Menggunakan PCA

Berdasarkan Gambar 2, terlihat bahwa data terbagi ke dalam tiga kelompok yang relatif terpisah meskipun masih terdapat beberapa titik yang saling berdekatan antar cluster. Cluster 1 (warna oranye) cenderung berada pada area dengan nilai komponen utama yang lebih tinggi, yang menunjukkan keterlibatan mahasiswa yang lebih merata pada berbagai kegiatan, sehingga dapat dikategorikan sebagai kelompok mahasiswa sangat aktif. Cluster 0 (warna biru) terlihat menyebar pada bagian bawah grafik dengan tingkat partisipasi yang cukup pada beberapa aspek kegiatan, sehingga dapat dikategorikan sebagai kelompok mahasiswa aktif. Sementara itu, Cluster 2 (warna hijau) cenderung berada pada sisi kiri grafik dengan nilai yang lebih rendah pada sebagian besar variabel, yang mencerminkan keterlibatan yang relatif rendah, sehingga dapat dikategorikan sebagai kelompok mahasiswa kurang aktif. Hasil visualisasi ini memperkuat temuan sebelumnya bahwa terdapat perbedaan pola keaktifan mahasiswa yang dapat dikelompokkan menjadi tiga karakteristik utama.

Temuan pembentukan tiga klaster keaktifan mahasiswa pada penelitian ini memiliki kecenderungan yang sejalan dengan hasil penelitian terdahulu yang juga menerapkan metode K-Means pada konteks mahasiswa. [21] melakukan segmentasi kinerja

akademik mahasiswa menggunakan metode K-Means dengan pendekatan Elbow dan menghasilkan tiga kelompok dengan karakteristik kinerja yang berbeda, yaitu tinggi, sedang, dan rendah. Temuan serupa juga ditunjukkan oleh [22] yang menerapkan K-Means dalam pengelompokan data mahasiswa berdasarkan profil akademik, dimana hasil pengelompokan menunjukkan adanya variasi karakteristik mahasiswa yang dapat dibedakan ke dalam beberapa kelompok. Kesamaan tersebut menunjukkan bahwa metode K-Means mampu mengidentifikasi variasi karakteristik mahasiswa pada berbagai konteks pengamatan. Pada penelitian ini, pendekatan yang sama menghasilkan tiga kelompok keaktifan mahasiswa UINSU yang terdiri atas mahasiswa sangat aktif (C1), mahasiswa aktif dengan dominasi keterlibatan pada kepanitiaan (C0), dan mahasiswa kurang aktif (C2).

Perbedaan karakteristik antar cluster pada penelitian ini memberikan gambaran bahwa tingkat partisipasi mahasiswa tidak bersifat homogen. Cluster C1 merepresentasikan kelompok mahasiswa sangat aktif, Cluster C0 menunjukkan mahasiswa aktif pada kegiatan tertentu khususnya kepanitiaan, sedangkan Cluster C2 menggambarkan mahasiswa dengan keterlibatan yang relatif lebih rendah. Informasi tersebut dapat menjadi dasar bagi institusi dalam menyusun pendekatan pembinaan yang lebih sesuai dengan karakteristik masing-masing kelompok sehingga program kemahasiswaan dapat diarahkan secara lebih tepat sasaran.

## KESIMPULAN DAN SARAN

Penelitian ini menunjukkan bahwa keaktifan mahasiswa Universitas Islam Negeri Sumatera Utara (UINSU) dapat dikelompokkan menjadi tiga kategori, yaitu sangat aktif, aktif, dan kurang aktif, menggunakan metode K-Means Clustering dengan jumlah cluster optimal  $k = 3$  berdasarkan metode Elbow yang ditandai penurunan signifikan nilai WCSS. Setiap cluster memiliki karakteristik berbeda pada aspek organisasi, seminar, kepanitiaan, lomba, dan pengembangan diri, di mana kelompok sangat aktif memiliki keterlibatan merata, kelompok aktif dominan pada aspek tertentu, dan kelompok kurang aktif menunjukkan partisipasi rendah, yang juga didukung oleh visualisasi PCA dengan pemisahan cluster yang jelas. Hasil ini dapat menjadi dasar perancangan program kemahasiswaan yang lebih tepat sasaran serta acuan penelitian selanjutnya dengan penambahan responden dan variabel agar hasil lebih komprehensif.

## REFERENSI

- [1] M. Baihaqi and S. Satriyadi, "Evaluasi Kegiatan Kemahasiswaan untuk Meningkatkan Kompetensi Mahasiswa Program Studi Pendidikan Agama Islam di INSAN Binjai," *EDU Soc. J. PENDIDIKAN, ILMU Sos. DAN Pengabd. Kpd. Masy.*, vol. 5, no. 1, pp. 1373–1386, 2025, doi: 10.56832/edu.v5i1.1029.
- [2] S. Yolanda, S. M. Sari, and I. Ismail, "Peran organisasi mahasiswa dalam membangun karakter kepemimpinan dan peningkatan soft skill," *Pendas J. Ilm. Pendidik. Dasar*, vol. 9, no. 04, pp. 361–373, 2024, doi: 10.23969/jp.v9i04.21514.
- [3] R. S. H. Mahiza and A. Nurhidayati, "Pengaruh lingkungan kampus, disiplin belajar, keaktifan berorganisasi terhadap prestasi belajar akademik mahasiswa," *JHIP-Jurnal Ilm. Ilmu Pendidik.*, vol. 8, no. 1, pp. 747–752, 2025, doi: 10.54371/jhip.v8i1.6415.
- [4] S. Y. A. Juana, P. Abdullatip, D. A. Fachrezi, A. Wahyudi, D. S. Permana, and B. Tarigan, "Dampak Perkembangan Teknologi Informasi Terhadap Aktivitas Fisik Mahasiswa," *J. Ilm. Spirit*, vol. 24, no. 2, pp. 80–91, 2024, doi: 10.36728/jis.v24i2.3652.
- [5] T. B. Pamungkas, S. Maesroh, and P. Ardiansyah, "Implementasi Data Mining Pada Stok Penggunaan Barang Di Gmf Aeroasia Menggunakan Algoritma K-Means Clustering," *J. Ilm. Sains dan Teknol.*, vol. 7, no. 2, pp. 112–123, 2023, doi: 10.47080/saintek.v7i2.2697.
- [6] I. A. Rosyada and D. T. Utari, "Penerapan principal component analysis untuk reduksi variabel pada algoritma K-Means clustering," *Jambura J. Probab. Stat.*, vol. 5, no. 1, pp. 6–13, 2024, doi: 10.37905/jips.v5i1.18733.
- [7] Deti Karmanita and Billy Hendrik, "Penerapan Metode Clustering dengan Algoritma K-Means pada Pengelompokan Peminatan Mata Kuliah," *J. Ilm. Dan Karya Mhs.*, vol. 1, no. 6, pp. 01–10, 2023, doi: 10.54066/jikma.v1i6.1028.
- [8] Ermatita M. Hafizh, "Cv infinite corporation jurnal pendidikan dan teknologi indonesia," vol. 5, no. 915, p. 2025, 2025.
- [9] H. Ermewaningsih *et al.*, "Klasterisasi Kinerja Akademik Mahasiswa Berdasarkan IPK Menggunakan K-Means dan Metode Elbow Clustering of Students' Academic Performance Based on GPA Using K -Means and the Elbow Method," vol. 6, no. 3, pp. 619–625, 2025.
- [10] N. Hendrastuty, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa," *J. Ilm. Inform. dan Ilmu Komput.*, vol. 3, no. 1, pp. 46–56, 2024, doi: 10.58602/jima-ilkom.v3i1.26.
- [11] R. Maulana and F. Fathoni, "Analisis Clustering Rekomendasi Mata Kuliah Peminatan Berdasarkan Karir Alumni Menggunakan Machine Learning," *JOISIE (Journal Inf. Syst. Informatics Eng.)*, vol. 9, no. 1, pp. 109–122, 2025, doi: 10.35145/joisie.v9i1.4954.
- [12] A. P. Putra, J. Tshivana, and E. Rilvani, "Perbandingan Teoritis dan Eksperimen Algoritma K-Means dan K-Medoids Dalam Klasterisasi Data," *Kohesi J. Sains dan Teknol.*, vol. 10, no. 2, pp. 61–70, 2025, doi: 10.2238/0b2z5035.
- [13] S. F. Mosse and I. Indrianto, "Penerapan Metode K-Means Clustering dalam Pengelompokan Data Penerima Beasiswa di Kabupaten Konawe Selatan," 2026, *Institut Teknologi PLN*.
- [14] A. Rajsya, P. Purnawansyah, and A. R. Manga, "Rancang Bangun Penerapan Metode Elbow Pada K-Means Untuk Clustering Data Persediaan Barang," *LINIER Lit. Inform. dan Komput.*, vol. 1, no. 4, pp. 395–403, 2024, doi: 10.33096/linier.v1i4.2539.
- [15] R. Sudrajat, A. I. Hadiana, and M. Melina, "Evaluasi Kualitas Klaster Wilayah Rawan Bencana Menggunakan K-Means dengan Silhouette dan Elbow Method," *J. Algoritm.*, vol. 22, no. 2, pp. 127–139, 2025, doi: 10.33364/algoritma/v.22-2.2379.
- [16] H. Amani, "Karakteristik tingkat kerusakan bangunan pascabencana alam menggunakan ekstraksi fitur Googlenet dan principal Component Analysis (PCA)-Clustering," 2025, *Universitas Islam Negeri Maulana Malik Ibrahim*.
- [17] N. Kholidin, N. Suarna, and W. Prihartono, "Segmentasi Gaya Hidup Mahasiswa Menggunakan Algoritma K-Means Clustering," *JELIKU (Jurnal Elektron. Ilmu Komput. Udayana)*, vol. 12, no. 3, 2024, doi: 10.24843/JLK.2023.v12.i03.p26.
- [18] A. F. Firdaus, A. F. Fitriani, and E. P. Nugroho, "Parallel Computing pada Clustering K-Means untuk Analisis Keketatan Program Studi SNBT 2023," *Komputika J. Sist. Komput.*, vol. 14, no. 1, p. 7, 2025, doi: 10.34010/komputika.v14i1.14983.
- [19] M. T. Wardani, V. O. Ramadhani, N. Anggreani, and A. D. Kuswanto, "Pengelompokan Provinsi Berdasarkan Kelayakan Ruang Kelas dan Tenaga Kependidikan Sekolah Dasar Menggunakan Algoritma K-Means: Analisis Data Periode 2023-2024," *Riau J. Tek. Inform.*, vol. 4, no. 2, pp. 163–172, 2025, doi: 10.30606/rjti.v4i2.3389.
- [20] J. Parhusip, J. Feranita, M. P. Mentari, T. Apriliani, and A. M. Jessend, "Implementasi Algoritma K-Means Clustering Untuk Segmentasi Tingkat Kesehatan Mental Mahasiswa Berdasarkan Faktor Stress Dan Akademik," *J. Media Inform.*, vol. 7, no. 1, pp. 449–458, 2026, doi: 10.55338/jumin.v7i1.7856.
- [21] L. Harmaji, A. S. Asaury, M. Wijana, and M. E. Habiby, "Segmentasi Kinerja Akademik Mahasiswa," vol. 7, no. 2, pp. 73–81, 2024.
- [22] Ahmad Zaki, Irwan, and I. A. Sembe, "Penerapan K-Means Clustering dalam Pengelompokan Data (Studi Kasus Profil Mahasiswa Matematika FMIPA UNM)," *J. Math. Comput. Stat.*, vol. 5, no. 2, pp. 163–176, 2006, [Online]. Available: <http://www.ojs.unm.ac.id/jmathcos>