

KLASTERISASI MODEL PEMBELAJARAN DI UNIVERSITAS TIMOR MENGGUNAKAN METODE *K-MEANS*

Krisantus Jumarto Tey Seran¹, Glorita D.R.M Fallo², Dian Grace Ludji³, Debora chrisinta⁴

^{1,2,3,4} Program Studi Teknologi Informasi, Universitas Timor
 Jl. Jalan El Tari - Km. 09, Kefamenanu – Nusa Tenggara Timur, Indonesia
 Email: ¹krisantusteyseran@unimor.ac.id, ²merdinfallo@gmail.com, ³dianludji@unimor.ac.id,
⁴deborachrisinta@unimor.ac.id

ABSTRAK

Penelitian ini bertujuan untuk mengidentifikasi pola model pembelajaran di Universitas Timor menggunakan metode *K-Means*. Lima variabel utama yang digunakan dalam klusterisasi meliputi metode pembelajaran, jenis evaluasi, ketersediaan sumber daya, tingkat kepuasan, dan fleksibilitas waktu belajar. Data primer diperoleh dari 1.018 mahasiswa aktif melalui kuesioner skala Likert. Jumlah kluster optimal ditentukan menggunakan metode *Elbow*, yang menghasilkan dua kluster utama. Kluster 1 didominasi oleh mahasiswa yang memilih pembelajaran offline dengan pendekatan diskusi dan evaluasi berupa tugas individu. Ketersediaan sumber daya masih terbatas, tingkat kepuasan tinggi dan fleksibilitas waktu belajar sangat baik. Kluster 2 juga menunjukkan preferensi terhadap pembelajaran *offline* dengan metode pendekatan dosen berupa diskusi jenis evaluasinya tugas individu. Sumber daya baru sebagian yang terpenuhi, dengan tingkat kepuasan yang baik dan fleksibilitas waktu belajar yang cukup baik. Evaluasi kualitas hasil kluster menggunakan metode *Silhouette Coefficient* menghasilkan nilai 0,3377, yang termasuk dalam kategori *Weak Structure*. Hasil penelitian ini diharapkan dapat menjadi bahan pertimbangan dalam pengembangan kebijakan pembelajaran dan penyusunan kurikulum yang lebih sesuai dengan karakteristik mahasiswa.

Kata kunci: Jenis Pembelajaran, Universitas Timor, Klusterisasi, *K-Means*, *Silhouette Coefficient*

ABSTRACT

This study aims to identify learning pattern types at the University of Timor using the K-Means clustering method. Five main variables were used in the clustering process, including learning method, type of evaluation, availability of resources, student satisfaction, and time flexibility. Primary data were obtained from 1,018 active students through a Likert-scale questionnaire. The optimal number of clusters was determined using the Elbow method, which resulted in two main clusters. Cluster 1 is dominated by students who prefer offline learning with a discussion-based approach and individual task evaluation. Resource availability is still limited, but the level of satisfaction is high, and learning time flexibility is considered very good. Cluster 2 also shows a preference for offline learning with a discussion-based teaching approach and individual task evaluation. However, resource availability is only partially met, with good satisfaction levels and fairly good time flexibility. The quality of the clustering results was evaluated using the Silhouette Coefficient method, which produced a score of 0.3377, categorized as a Weak Structure. The results of this study are expected to serve as a consideration in developing learning policies and curriculum planning that better align with student characteristics.

Keywords: Learning Types, University of Timor, Clustering, *K-Means*, *Silhouette Coefficient*

1. PENDAHULUAN

Perkembangan teknologi informasi memberikan pengaruh besar dalam dunia pendidikan, termasuk pada model pembelajaran yang diterapkan dalam satu perguruan tinggi [1]. Jenis pembelajaran selalu merujuk pada pendekatan, metode, dan strategi yang digunakan untuk mencapai tujuan pendidikan di satu Universitas [2]. Namun, pendekatan dan metode yang diterapkan harus disesuaikan dengan karakteristik mahasiswa seperti: gaya belajar, minat, dan latar belakang akademik. Penerapan ini juga dilakukan di salah satu perguruan tinggi negeri di Pulau Timor yang terletak di Kota Kefamenanu yakni Universitas Timor (Unimor). Sebagai salah satu institusi pendidikan tinggi, dalam melakukan dharma pengajaran juga menerapkan beberapa jenis model yakni: *offline*, *online*, dan *hybrid* (gabungan) [3], [4]. Dengan keberagaman karakteristik mahasiswa dan metode pengajaran menciptakan tantangan tersendiri dalam pengelolaan pembelajaran yang efektif di kampus yang berada di wilayah perbatasan Indonesia dan Timor Leste (Distrik *Oecusse*).



Pada maret 2025, Unimor memiliki Empat Fakultas dengan Enam Belas Program Studi (Prodi). Jumlah mahasiswa aktif saat sebanyak 6.952 dan memiliki jumlah Dosen sebanyak 310 orang. Besarnya jumlah rombongan belajar tentunya mempengaruhi kebijakan Unimor dalam menyusun strategi khusus untuk metode yang tepat sasaran. Civitas akademika (Dosen dan Mahasiswa) Unimor berasal dari berbagai wilayah yang ada di Indonesia. Sebagian besar didominasi oleh mahasiswa dari tanah Timor seperti Kabupaten Malaka, Belu, TTU, TTS, dan Kupang, tidak sedikit mahasiswa dari pulau lain dalam wilayah Provinsi NTT bahkan ada yang datang dari Pulau Jawa, Sumatera, Sulawesi, Papua, dan dari negara tetangga yakni Timor Leste. Hal yang sama juga berlaku untuk tenaga pengajar (Dosen) Unimor. Banyak dosen yang datang dari berbagai wilayah yang ada di Indonesia. Mulai dari Sumatera (Aceh), Jawa, Bali, hingga Sulawesi, dan Pulau Timor tentunya. Jumlah mahasiswa yang banyak dipengaruhi oleh latar belakang budaya yang berbeda serta karakteristik setiap pengajar tentunya mempengaruhi cara pengelolaan kegiatan belajar mengajar (KBM) di kampus ini. Tiga model pembelajaran (Luring, Daring, Gabungan) sudah dijalankan selama ini di Unimor perlu dievaluasi kembali untuk mengetahui model mana yang paling tepat untuk mendukung proses KBM.

Upaya evaluasi model pembelajaran dapat dilakukan melalui pendekatan berbasis data, salah satunya dengan menggunakan metode K-Means klustering. Klusterisasi merupakan teknik pengelompokan data tanpa label yang didasarkan pada tingkat kesamaan karakteristik antarobjek dalam satu kelompok. Beberapa penelitian terdahulu, seperti yang dilakukan di Universitas Pasir Pengaraian, menunjukkan bahwa metode K-Means mampu mengelompokkan dosen ke dalam tiga jenis pembelajaran, yaitu *offline*, *online*, dan *blended learning*, dengan hasil bahwa *blended learning* menjadi strategi paling efektif selama pandemi Covid-19. Adapun penelitian "Klusterisasi Jenis Pembelajaran di Universitas Timor Menggunakan Metode K-Means" memiliki perbedaan pada konteks lokasi, sumber data, dan variabel penelitian. Variabel yang dianalisis juga lebih komprehensif, meliputi metode pembelajaran, jenis evaluasi, tingkat kepuasan mahasiswa, serta ketersediaan sumber daya. Penelitian ini bertujuan untuk mengidentifikasi jenis pembelajaran yang paling efektif serta memberikan rekomendasi bagi pengembangan kurikulum dan metode pengajaran yang lebih optimal [5].

Dalam penelitian ini, diambil sampel sebanyak 1.018 mahasiswa aktif dari empat fakultas yang ada di Unimor sebagai responden untuk data primer. Instrumen kuisioner penelitian menggunakan Skala Likert [6]. Untuk analisis hasil digunakan pendekatan ilmu teknik data mining dengan metode *K-Means*. [7], [8]. Tujuannya agar dapat mengelompokkan jawaban dari semua responden kedalam tiga kategori metode pembelajaran. Algoritma *K-Means* dipakai karena terbukti efektif dalam mengkluster data berdasarkan kemiripan karakteristiknya [9], [10]. Atribut seperti Metode Pembelajaran, Jenis Evaluasi, Tingkat Kepuasan, Ketersediaan Sumber Daya, dan Fleksibilitas Waktu dipakai sebagai variabel dalam kuisioner penelitian. Cara kerja Algoritma *Machine Learning* ini adalah dengan memilih *centroid* awal secara acak [11], [12]. Kemudian data akan dikelompokkan berdasarkan jarak terdekat dengan menerapkan perhitungan Euclidean [13], [14]. Proses ini berulang hingga mendapatkan hasil yang stabil (tetap). Untuk mengevaluasi hasil pengelompokkan digunakan Metode *Silhouette Coefficient* dalam rentang -1 hingga 1, yang mengindikasikan seberapa kuat posisi suatu data berada pada kelompok yang tepat [15], [16].

Penelitian ini bertujuan untuk mengidentifikasi model pembelajaran yang paling banyak diterapkan pada kampus Unimor. Hasil pengelompokkan model pembelajaran memberikan gambaran yang jelas tentang dukungan atau kecenderungan (preferensi) mahasiswa Unimor dalam perbaikan proses KBM di Unimor. Luaran dari penelitian ini, diharapkan menjadi salah satu masukan untuk perbaikan model pembelajaran dalam implementasi kurikulum terbaru yang akan diterapkan di masa yang akan datang.

2. METODE PENELITIAN

Pendekatan kuantitatif [17] digunakan dalam penelitian ini dengan menggunakan data primer dari kuisioner *online* yang dibagikan melalui *google form* [18], [19]. Dalam penelitian ini dilakukan empat tahapan (langkah) mulai dari pengumpulan data sampai evaluasi hasil penelitian. Berikut detail penjelasan mengenai langkah dalam penelitian ini.

Pengumpulan Data

Proses penelitian diawali dengan uji validitas dan reliabilitas instrumen. Hasil uji validitas menunjukkan bahwa seluruh item memiliki nilai r hitung lebih besar dari r tabel, sehingga dinyatakan valid. Selanjutnya, uji reliabilitas menggunakan metode Cronbach's Alpha memperoleh nilai sebesar 0,8401, lebih besar dari nilai acuan 0,6. Dengan demikian, instrumen penelitian dinyatakan reliabel dan layak digunakan untuk analisis lebih lanjut. Tahap berikutnya adalah penyebaran alamat (*link*) *Elektronik* Kuisioner dari *google form* kepada mahasiswa Unimor sebagai responden. Hasilnya, sebanyak 1.018 mahasiswa dari enam belas Prodi dalam Empat Fakultas terlibat dalam memberikan jawaban kuisioner tersebut. Data yang tersedia dalam *google form* kemudian diubah kedalam bentuk tabulasi excel [20], [21]. Data mentah ini kemudian disiapkan untuk dikelola dan dianalisis lebih lanjut pada tahapan selanjutnya.

Preprocessing Data

Data mentah dalam bentuk Excel kemudian dilakukan proses Pembersihan Data (*Data Cleaning*). Proses ini dilakukan untuk menghilangkan data yang tidak lengkap, menghapus *outlier*, dan memperbaiki data yang tidak konsisten agar data siap untuk dianalisis. Total data setelah proses pembersihan adalah 1.018 data. Transformasi



Data (*Data Transformation*) yaitu proses mengubah atau menyesuaikan bentuk/skala data sebelum dilakukan proses pengelompokan agar perhitungan jarak menjadi lebih akurat. Validasi dan Penyusunan Ulang (*Validation and Restructuring*) yaitu tahap evaluasi dan perbaikan hasil pengolahan data agar model yang dihasilkan benar-benar valid dan representatif [22]. Semua tahapan ini tidak dilakukan secara manual dalam Excel, melainkan menggunakan bantuan komputasi berbasis Python [23]. Penggunaan lingkungan komputasi berbasis cloud sangat bermanfaat karena dapat mengintegrasikan penyimpanan data, penulisan kode program, dan mengeksekusi analisis. Diawali dengan mengimport (*upload*) data mentah Excel ke dalam sistem penyimpanan. Kemudian, menerapkan *library* milik Python agar dapat membaca *file* dan kemudian bisa dimanipulasi sesuai dengan kebutuhan [24]. Penggunaan lingkungan komputasi berbasis Python sangat membantu dalam alur kerja analisis data karena setiap tahapan dapat terdokumentasi, dilacak kembali, direplikasi, dan dikembangkan lebih lanjut secara sistematis.

Implementasi *K-Means*

Perhitungan awal Algoritma ini dengan menentukan jumlah kluster optimal. Metode yang digunakan adalah *Elbow*. Metode *Elbow* merupakan salah satu pendekatan yang digunakan untuk menentukan jumlah kluster optimal dalam analisis kluster, khususnya pada algoritma *K-Means*. Tujuan utama dari metode ini adalah untuk memilih nilai k (jumlah kluster) yang paling efisien, yakni nilai k terkecil yang masih menghasilkan tingkat kesalahan (*error*) yang rendah [25]. Setelah mendapatkan jumlah kluster, perhitungan jarak antara setiap data dan *centroid* menggunakan rumus *Euclidean* [26]. Pada awal perhitungan *centroid* ditentukan secara acak dari nilai yang ada dalam data. Hasil perhitungan menjadi dasar dalam pengelompokan (kluster) berdasarkan jarak terdekat. Selanjutnya, posisi *centroid* diperbaharui dengan menghitung kembali rata-rata dari nilai data yang ada dalam setiap kluster. Proses ini dilakukan secara berulang sampai pada tahap dimana setiap posisi kluster tidak berubah (tetap). Hal ini ditandai dengan kesamaan pengelompokan data pada 2 iterasi terakhir. Perhitungan menggunakan Metode *Elbow*, *Euclidean Distance* dapat dilihat dalam penjelasan berikut.

$$wcss = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2 \quad (1)$$

Keterangan:

- $wcss$: Jumlah kuadrat jarak dalam kluster
- k : Jumlah kluster
- C_i : Kluster ke- i
- $x \in C_i$: Data ke- x dalam kluster ke- i
- μ_i : *Centroid* kluster ke- i
- $\|x - \mu_i\|^2$: Jarak kuadrat antara data dan *centroid*

Rumus *Euclidean Distance*

$$d(x_i, \mu_i) = \sqrt{\sum (x_i - y_i)^2} \quad (2)$$

Keterangan:

- x_i : Data kriteria
- y_i : *centroid* pada kluster ke- j
- $d(x_i, \mu_i)$: jarak antara titik data x_i dan *centroid* μ_i

Rumus Pengelompokan data

$$a_{ji} = \begin{cases} 1, & d = \min(D(x_j, C_i)) \\ 0, & \text{lainnya} \end{cases} \quad (3)$$

Keterangan:

- a_{ji} : Indikator keanggotaan data x_j dalam kluster C_i
- x_j : Titik data ke- j dalam dataset
- C_i : *Centroid* pusat dari kluster ke- i
- $D(x_i, C_i)$: jarak antara titik data X_j dengan *centroid* C_i

Evaluasi Hasil Klustering

Untuk mengukur hasil klustering (pengelompokan), digunakan metode *Silhouette Coefficient* [27]. Metode ini menilai hasil pengelompokan dengan membandingkan kedekatan data dalam satu kluster dan jaraknya dengan kluster lain. Nilai yang mendekati +1 menandakan pengelompokan yang akurat. Berdasarkan hasil ini, setiap kluster kemudian ditafsirkan dengan mengidentifikasi ciri khas utamanya untuk menemukan pola yang ada, sehingga dapat disusun rekomendasi yang tepat guna meningkatkan efektivitas pembelajaran. Penjelasan mengenai *Silhouette Coefficient* dapat dilihat pada rumus di bawah ini.

$$s = \frac{b-a}{\max(a,b)} \quad (4)$$

Keterangan:

- s : Koefisien *Silhouette*
- a : Rata-rata jarak sampel ke seluruh titik data lain dalam kluster yang sama
- b : Rata-rata jarak sampel ke seluruh titik data lain dalam kluster terdekat



Rumus untuk mencari nilai a dan b dari *silhouette coefficient*

$$a(x_i) = \frac{1}{|C_i|-1} \sum_{x_j \in C_i, j \neq i} D(x_i, x_j) \quad (5)$$

Keterangan:

$a(x_i)$: Rata-rata jarak antara x_i dan semua titik lain dalam Klaster yang sama C_i (intra-klaster distance)

C_i : Jumlah total anggota dalam klaster C_i

$D(x_i, x_j)$: Jarak antara titik x_i dan titik x_j dalam klaster yang sama

$\sum_{x_j \in C_i, j \neq i} D(x_i, x_j)$: Penjumlahan semua jarak dari x_i ke setiap titik lain dalam klaster yang sama

$$b(x_i) = \min_{C_k \neq C_i} \frac{1}{|C_k|} \sum_{x_j \in C_k} D(x_i, x_j) \quad (6)$$

Keterangan:

$b(x_i)$: Rata-rata jarak antara x_i dengan Klaster terdekat yang berbeda dari klaster yang sama

$C_k \neq C_i$: Klaster lain yang bukan tempat x_i berada

$|C_k|$: Jumlah anggota dalam klaster lain C_k

$D(x_i, x_j)$: Jarak antara x_i dan titik x_j yang ada di klaster lain C_k

$\sum_{x_j \in C_k} D(x_i, x_j)$: Total jarak antara x_i dan semua anggota klaster lain C_k

3. HASIL DAN PEMBAHASAN

Uji Validitas

Sebelum proses pengumpulan data lewat kuesioner *online (Google form)*, dilakukan dulu uji validitas dan reabilitas dari instrumen kuesioner yang akan digunakan. Dalam pengujian ini digunakan data dari 26 responden awal yang mengisi kuisisioner. Kriteria uji validitas menggunakan Pearson Correlation (r hitung) [28] yaitu dengan membandingkannya dengan nilai r tabel. Hasilnya nilai **r Hitung** lebih besar dari **r Tabel** yang dapat dilihat pada Tabel 1 di bawah ini.

Tabel 1. Bobot Kriteria

Item	r Hitung	r Tabel	Keterangan
1	0,893731	0,388	Valid
2	0,848086148	0,388	Valid
3	0,832795218	0,388	Valid
4	0,708685398	0,388	Valid

Karena semua item (1-4) memiliki r hitung (0,893731; 0,848086148; 0,832795218; 0,708685398) yang lebih besar dari r tabel (0,388), maka seluruh item valid dan layak. Selanjutnya dilakukan uji reabilitas (Cronbach's Alpha) [29]. Berdasarkan kriteria pengujian, suatu instrumen dinyatakan reliabel jika nilai Cronbach's Alpha melebihi batas minimal 0,7. Dari tabel hasil uji, nilai 0,8401 jauh melampaui batas tersebut, sehingga dapat disimpulkan bahwa instrumen ini memiliki konsistensi internal yang **sangat baik**..

Tabel 2. Uji Reabilitas

Nilai Acuan	Cronbach's Alpha	Keterangan
0,7	0,8401	Reliabel

Dengan demikian, instrumen penelitian tidak hanya dinyatakan valid berdasarkan uji sebelumnya, tetapi juga reliabel. Artinya alat ukur ini dapat dipercaya untuk menghasilkan data yang stabil dan konsisten apabila digunakan berulang kali dalam mengukur variabel yang sama. Langkah berikut adalah proses pengumpulan (penyebaran) data kuisisioner dan didapatkan sebanyak 1.018 respon jawaban dari mahasiswa Unimor. Gambaran data yang sudah dikumpulkan lebih jelasnya pada Tabel 3.

Tabel 3. Dataset

No	Metode Pembelajaran	Jenis evaluasi	Ketersediaan Sumber Daya	Tingkat Kepuasan	fleksibilitas
1	3	2	2	7	2
2	4	2	1	10	4



No	Metode Pembelajaran	Jenis evaluasi	Ketersediaan Sumber Daya	Tingkat Kepuasan	fleksibilitas
3	2	2	2	10	2
4	4	2	1	9	3
5	2	1	2	12	4
6	3	3	2	9	2
7	2	2	2	9	4
8	4	2	2	9	3
9	2	3	2	7	3
10	3	3	2	9	3
...
1017	2	1	0	6	2
1018	3	1	1	9	3

Tahap Preprocessing

Dalam tahapan *preprocessing* dilakukan persiapan dan pengolahan awal terhadap data mentah untuk menjamin kualitas, konsistensi, dan kesiapannya sebagai bahan analisis dalam proses data mining. Tahap ini menjadi dasar yang penting, karena hasil akhir penambangan data sangat bergantung pada kondisi data kuisioner yang diolah. Dalam *preprocessing* penelitian ini dilakukan pembersihan data, menghapus duplikasi data, data kotor (tidak terisi dengan baik). Normalisasi data tidak dilakukan pada tahap ini dikarenakan Seluruh variabel dalam penelitian telah berada dalam skala numerik yang relatif sebanding dan tidak memiliki perbedaan rentang yang sangat ekstrem antar atribut. Meskipun atribut Tingkat Kepuasan memiliki rentang nilai 1–10, nilai tersebut masih berada dalam skala yang tidak terlalu jauh dibandingkan dengan atribut lainnya yang juga telah dikonversi ke dalam bentuk numerik diskrit. Dari hasil *preprocessing* tidak ditemukan nilai yang hilang seperti yang ada dalam Gambar 1.

```

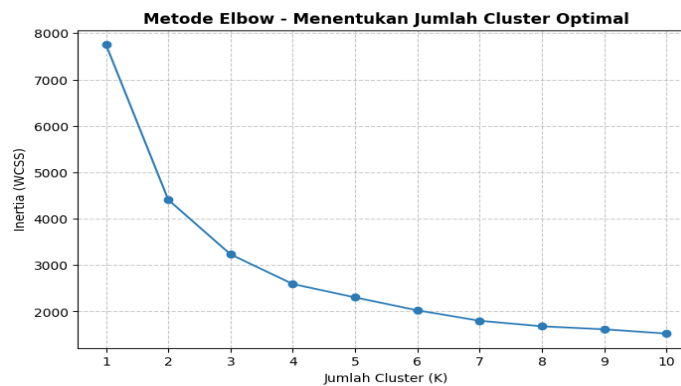
→ Nilai hilang per kolom:
Metode Pembelajaran           0
Jenis evaluasi                 0
Ketersediaan Sumber Daya      0
Tingkat Kepuasan              0
Seberapa fleksibel waktu belajar dalam model pembelajaran yang Anda ikuti? 0
dtype: int64

```

Gambar 1. Hasil *preprocessing* data

Implementasi Algoritma K-Means

Seperti yang sudah dijelaskan dalam tahap penelitian, langkah pertama dalam algoritma ini adalah penentuan jumlah kluster optimal dengan Metode *Elbow*. Perhitungan Algoritma *K-Means* menggunakan *platform* Google Colab. Titik optimal dalam metode ini ditandai dengan penurunan nilai SSE yang mulai melambat (lengkung) secara signifikan sehingga membentuk sudut siku. Hasil analisis mendapatkan jumlah kluster optimal (K) sebanyak 2.



Gambar 2. Penentuan Kluster Metode *Elbow*

Setelah mendapatkan $K=2$, dilanjutkan dengan menentukan titik *centroid* awal untuk proses pengelompokan. langkah selanjutnya adalah menentukan titik awal *centroid* untuk memulai proses pengelompokan. Fungsi dari penentuan titik awal ini sangat penting, karena *centroid* awal bertindak sebagai titik

acuan pertama yang akan digunakan sebagai dasar dalam mengelompokkan setiap data ke kluster terdekat. *Centroid* awal data dilihat pada Tabel 4.

Tabel 4. Titik *Centroid* Awal

Centroid	V1	V2	V3	V4	V5
C1	2	2	2	10	2
C2	2	2	2	9	4

Kemudian dilanjutkan dengan menghitung jarak setiap data dan *centroid* menggunakan rumus Eucliden Distance. Iterasi berlangsung sebanyak 3 kali hingga mencapai kondisi konvergen, yaitu saat posisi *centroid* dan keanggotaan data pada masing-masing kluster tidak lagi mengalami perubahan. Berdasarkan hasil perhitungan jarak, setiap data kemudian dikelompokkan ke dalam kluster yang memiliki jarak paling kecil (jarak minimum) terhadap *centroid*. Tahap selanjutnya adalah memperbarui posisi *centroid* untuk setiap kluster. Pada tahap ini, cara perhitungan *centroid* baru dilakukan dengan mengambil nilai rata-rata dari seluruh data yang berada dalam satu kluster, untuk setiap variabel (fitur) yang digunakan. Gambar di bawah ini merupakan representasi dari 5 data awal hasil perhitungan proses *K-Means* yang dilakukan sebanyak 3 iterasi dimana tidak lagi terjadi perubahan atau telah mencapai kondisi konvergen. Gambar 3, 4, dan 5 menunjukkan nilai perhitungan Eucliden dari Iterasi 1 sampai dengan 3.

```

===== Iterasi 1 =====
Jumlah data per cluster:
Cluster 1: 399 data
Cluster 2: 619 data

Nilai jarak Euclidean dan assignment cluster:
  Data Ke  Jarak ke C1  Jarak ke C2  Cluster Assignment
0         1         3.162278   3.000000           2
1         2         3.000000   2.449490           2
2         3         0.000000   2.236068           1
3         4         2.645751   2.449490           2
4         5         3.000000   3.162278           1
...
1013      1014       2.236068   2.449490           1
1014      1015       5.000000   5.099020           1
1015      1016       4.242641   3.872983           2
1016      1017       4.582576   4.242641           2
1017      1018       2.236068   2.000000           2

[1018 rows x 4 columns]

Nilai centroid terbaru:
Centroid C1: [ 2.99749373  1.90726817  1.10025063 10.64912281  3.00501253]
Centroid C2: [3.06462036  1.88045234  1.06138934  7.7641357  2.86429725]

```

Gambar 3. Iterasi 1

```

===== Iterasi 2 =====
Jumlah data per cluster:
Cluster 1: 368 data
Cluster 2: 650 data

Nilai jarak Euclidean dan assignment cluster:
  Data Ke  Jarak ke C1  Jarak ke C2  Cluster Assignment
0         1         3.891568   1.493443           2
1         2         1.560458   2.679908           1
2         3         1.801264   2.788332           1
3         4         1.934761   1.561657           2
4         5         2.332927   4.692757           1
...
1013      1014       1.623432   2.700326           1
1014      1015       3.442174   5.278909           1
1015      1016       4.944954   2.402408           2
1016      1017       5.064886   2.625708           2
1017      1018       1.884893   1.526079           2

[1018 rows x 4 columns]

Nilai centroid terbaru:
Centroid C1: [ 3.04076087  1.90217391  1.07880435 11.18478261  3.41576087]
Centroid C2: [3.03692308  1.88461538  1.07538462  7.59846154  2.63846154]

```

Gambar 4. Iterasi 2

```

===== Iterasi 3 =====
Jumlah data per cluster:
Cluster 1: 368 data
Cluster 2: 650 data

Nilai jarak Euclidean dan assignment cluster:
  Data Ke  Jarak ke C1  Jarak ke C2  Cluster Assignment
0         1         4.514047  1.278820         2
1         2         1.637365  2.927061         1
2         3         2.312886  2.849291         1
3         4         2.425294  1.743999         2
4         5         1.936909  4.892851         1
...       ...         ...         ...         ...
1013      1014      1.957840  2.853607         1
1014      1015      3.127106  5.479612         1
1015      1016      5.548843  2.196850         2
1016      1017      5.652192  2.444774         2
1017      1018      2.401650  1.698414         2

[1018 rows x 4 columns]

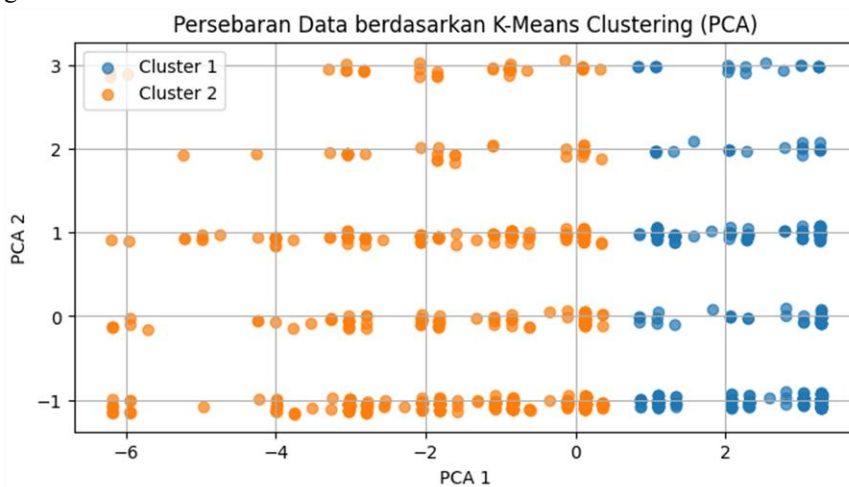
Nilai centroid terbaru:
Centroid C1: [ 3.04076087  1.90217391  1.07880435  11.18478261  3.41576087]
Centroid C2: [3.03692308  1.88461538  1.07538462  7.59846154  2.63846154]

 Konvergen pada iterasi ke-3

```

Gambar 5. Iterasi 3

Setelah proses klusterisasi selesai dilakukan, data yang telah dikelompokkan kemudian divisualisasikan dengan memanfaatkan *Principal Component Analysis* (PCA). Teknik ini digunakan untuk melakukan reduksi dimensi, yaitu mengurangi jumlah variabel atau fitur dalam data dengan tetap mempertahankan sebanyak mungkin informasi penting.



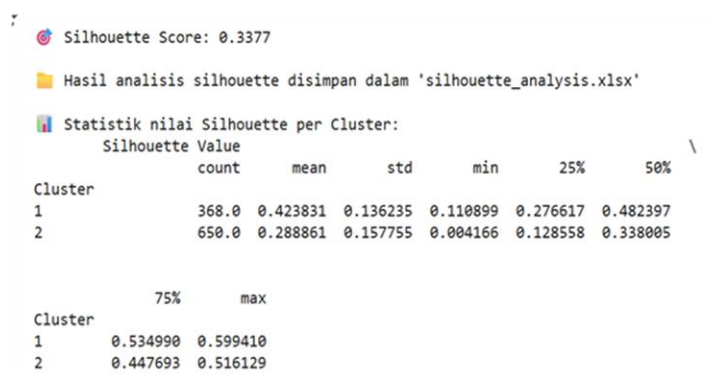
Gambar 6. Visualisasi Data Dengan PCA

Hasil visualisasi menggunakan PCA memperlihatkan bahwa data terbagi secara jelas ke dalam dua klaster dengan masing-masing karakteristik seperti di bawah ini:

- 1) Pada Klaster 1, sebagian besar mahasiswa (50,5%) memilih metode pembelajaran offline dengan pendekatan diskusi, menunjukkan preferensi terhadap pembelajaran tatap muka yang interaktif. Jenis evaluasi yang dominan adalah tugas individu (65,2%), disusul ujian tertulis (22,3%) dan kuis atau tes (12,5%). Ketersediaan sumber daya dinilai sebagian besar dalam kategori menengah (67,7%), yang berarti belum sepenuhnya terpenuhi. Meskipun demikian, tingkat kepuasan tergolong tinggi, dengan 50,3% responden menyatakan puas, dan mayoritas (51,4%) menilai waktu belajar sangat fleksibel.
- 2) Sementara itu, Klaster 2 juga menunjukkan kecenderungan serupa, di mana 49,2% responden memilih pembelajaran offline dengan pendekatan diskusi, serta tugas individu sebagai jenis evaluasi utama (63,2%). Namun, ketersediaan sumber daya dinilai masih terbatas, dengan 65,7% responden menyatakan hanya sebagian yang tersedia, bahkan 13,4% menilai tidak tersedia sama sekali. Tingkat kepuasan tetap tergolong baik, dengan 42,2% responden memberikan nilai tinggi terhadap kepuasan belajar. Selain itu, sebanyak 56,0% menyatakan bahwa waktu belajar dalam klaster ini cukup fleksibel, yang mendukung kenyamanan proses belajar.

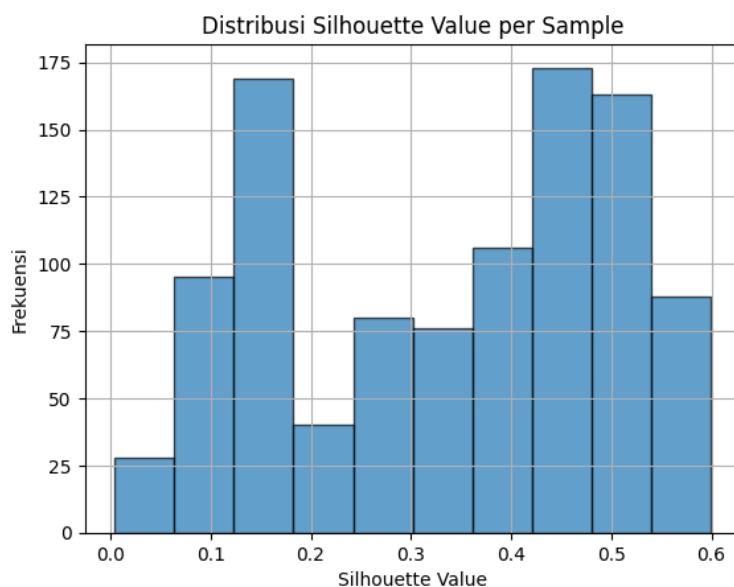
Evaluasi Silhouette Coefficient

Untuk mengevaluasi kualitas hasil klustering yang telah dilakukan, digunakan Metode Silhouette Coefficient sebagai alat ukur objektif. Metode ini bertujuan untuk menilai seberapa baik struktur pengelompokan yang terbentuk, dengan mengukur dua aspek utama: sejauh mana suatu objek mirip dengan objek lain dalam kluster yang sama (kohesi) dan seberapa jelas pemisahannya dari objek dalam kluster lain (separasi).



Gambar 7. Nilai Silhouette Score

Dari hasil perhitungan didapatkan nilai silhouette score sebesar 0,3377.



Gambar 8. Nilai Silhouette Score

Berdasarkan evaluasi menggunakan Silhouette Coefficient yang divisualisasikan pada Gambar 8, dapat disimpulkan bahwa hasil klustering menunjukkan kecenderungan positif meskipun belum optimal secara menyeluruh. Kluster 1 mencatat performa terbaik dengan nilai rata-rata tertinggi sebesar 0,4238, yang mengindikasikan bahwa objek-objek di dalamnya memiliki tingkat kohesi yang baik dan terpisah secara jelas dari kluster lain. Secara keseluruhan, distribusi nilai Silhouette dalam histogram menunjukkan bahwa sebagian besar data terkonsentrasi pada rentang 0,1 hingga 0,3, dengan puncak frekuensi pada kisaran 0,15 dan 0,45. Hal ini menggambarkan bahwa meskipun pengelompokan telah berjalan dengan cukup baik, masih ada ruang untuk peningkatan mengingat nilai-nilai tersebut belum mendekati ambang ideal di atas 0,5. Dari sisi yang lain, kehadiran sejumlah sampel dengan nilai mencapai 0,6 menjadi penanda bahwa sebagian data berhasil dikelompokkan secara sangat akurat, memperlihatkan pemisahan kluster yang sangat tajam.

4. SIMPULAN

Metode *K-Means* dalam evaluasi model pembelajaran berhasil diimplementasikan dalam penelitian ini. Berdasarkan hasil analisis data, penelitian ini menemukan bahwa sebagian besar (mayoritas) mahasiswa Unimor lebih memilih pembelajaran tatap muka yang berfokus pada diskusi dan tugas individu dibandingkan dengan model yang lain. Pola ini didukung oleh tingginya kepuasan mahasiswa dan fleksibilitas waktu yang mereka rasakan.

Namun, sumber daya pendukung seperti bahan ajar dan fasilitas belajar dinilai masih perlu ditingkatkan untuk mendukung pengalaman belajar yang lebih optimal.

Hasil pengelompokan mahasiswa menunjukkan pola yang bisa dikenali, namun batas antar kelompok tidak terlalu jelas. Ini terlihat dari nilai Silhouette Score 0.3377 yang menunjukkan pemisahan kelompok yang cukup saja. Hal ini menggambarkan bahwa karakteristik mahasiswa seperti gaya belajar, tingkat kepuasan, dan kebutuhan akan sumber daya saling tumpang tindih dan tidak terbagi secara jelas, melainkan membentuk suatu perubahan atau perbedaan yang terjadi secara bertahap.

Oleh karena itu, disarankan untuk tidak hanya memperbaiki ketersediaan sumber daya, tetapi juga melakukan pendekatan yang lebih personal dalam pembelajaran. Selanjutnya, penelitian dengan metode (algoritma) yang lebih mendalam serta penambahan variabel juga perlu untuk memahami kebutuhan setiap mahasiswa secara lebih rinci di masa mendatang.

DAFTAR PUSTAKA

- [1] M. W. Zafar, S. A. H. Zaidi, S. Mansoor, A. Sinha, and Q. Qin, "ICT and Education as Determinants of Environmental Quality: The Role of Financial Development in Selected Asian Countries," *Technological Forecasting and Social Change*, vol. 177, p. 121547, Apr. 2022, doi: 10.1016/J.TECHFORE.2022.121547.
- [2] M. T. Machmud, A. P. Widiyan, and N. R. Ramadhani, "The Development and Policies of ICT Supporting Educational Technology in Singapore, Thailand, Indonesia, and Myanmar," *International Journal of Evaluation and Research in Education*, vol. 10, no. 1, pp. 78–85, Mar. 2021, doi: 10.11591/ijere.v10i1.20786.
- [3] H. Mujianto and U. Suryadhianto, "Transformation of Learning Systems With A Hybrid Approach Based On Information Technology," *Pendidikan dan Humaniora*, vol. 9, no. 1, p. 2025, 2025, doi: 10.36526/js.v3i2.5007.
- [4] W. A. Blegur, K. J. Tey Seran, and F. M. A. Blegur, "Efektivitas Pembelajaran Daring Mahasiswa Prodi Biologi, Matematika dan Teknologi Informasi Unimor di Masa Pandemi Covid-19," *Jurnal Saintek Lahan Kering*, vol. 4, no. 2, pp. 19–21, Jan. 2022, doi: 10.32938/slk.v4i2.1534.
- [5] L. Fimawahib and E. Rouza, "Penerapan K-Means Clustering pada Penentuan Jenis Pembelajaran di Universitas Pasir Pengaraian," *INOVTEK Polbeng*, vol. 6, no. 2, pp. 234–247, 2021, doi: 10.35314/isi.v6i2.2096.
- [6] Moh. F. Erinsyah, G. W. Sasmito, D. S. Wibowo, and V. K. Bakti, "Sistem Evaluasi pada Aplikasi Akademik Menggunakan Metode Skala Likert dan Algoritma Naïve Bayes," *KOMPUTA*, vol. 13, no. 1, pp. 74–81, Apr. 2024, doi: <https://doi.org/10.34010/komputa.v13i1.10940>.
- [7] K. Annisa, B. S. Ginting, and M. A. Syari, "Penerapan Data Mining Pengelompokan Data Pengguna Air Bersih Berdasarkan Keluhannya Menggunakan Metode Clustering Pada PDAM Langkat," *ALGORITMA*, vol. 6, no. 1, pp. 112–127, Apr. 2022, doi: <https://doi.org/10.59697/jsik.v6i2.167>.
- [8] N. Karolina, "Data Mining Pengelompokan Pasien Rawat Inap Peserta BPJS Menggunakan Metode Clustering (Studi Kasus : RSUD.Bangkitan)," *JITU (Journal of Information and Technology Unimor)*, pp. 47–53, Sep. 2021, doi: <https://doi.org/10.32938/jitu.v1i2.1470>.
- [9] N. Hendrastuty, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa," *Jurnal Ilmiah Informatika dan Ilmu Komputer (JIMA-ILKOM)*, vol. 3, no. 1, pp. 46–56, Mar. 2024, doi: 10.58602/jima-ilkom.v3i1.26.
- [10] C. S. D. Sembiring, L. Hanum, and S. P. Tamba, "Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Menentukan Judul Skripsi dan Jurnal Penelitian (Studi Kasus FTIK UNPRI)," *Jurnal Sistem Informasi dan Ilmu Komputer Prima*, vol. 5, no. 2, pp. 80–85, Feb. 2022, doi: <https://doi.org/10.34012/jurnalsisteminformasidanilmukomputer.v5i2.2393>.
- [11] A. P. Putra, J. Tshivana, and E. Rilvani, "Perbandingan Teoritis dan Eksperimen Algoritma K-Means dan K-Medoids Dalam Klasterisasi Data," *Kohesi*, vol. 10, no. 2, pp. 1–24, 2025, doi: 10.8734/Kohesi.v1i2.365.
- [12] V. Rapika Sari, E. Buulolo, and K. Kunci ABSTRAK, "Implementasi Algoritma K-Means dengan Normalisasi Sigmoidal Untuk Klastering Data Ternak Sapi," *JIKTEKS*, vol. 2, no. 1, pp. 30–42, Dec. 2023, doi: <https://doi.org/10.70404/jikteks.v2i01.52>.
- [13] F. N. Cahya, Y. Mahatma, and S. R. Rohimah, "Perbandingan Metode Perhitungan Jarak Euclidean dengan Perhitungan Jarak Manhattan pada K-Means Clustering Dalam Menentukan Penyebaran Covid di Kota Bekasi," *JMT: Jurnal Matematika dan Terapan*, vol. 5, no. 1, pp. 43–55, Feb. 2023, doi: 10.21009/jmt.5.1.5.
- [14] W. W. Pribadi, A. Yunus, and A. Sartika Wiguna, "Perbandingan Metode K-Means Euclidean Distance dan Manhattan Distance Pada Penentuan Zonasi Covid-19 di Kabupaten Malang," *Jurnal Mahasiswa Teknik Informatika*, vol. 6, no. 2, pp. 439–500, Sep. 2022, doi: <https://doi.org/10.36040/jati.v6i2.4808>.
- [15] Febby Arisca Zurfani, Sawaluddin, Mardiningsih, and Muhammad Romi Syahputra, "Analisis Metode Clustering K-Means pada Zonasi Daerah Terdampak Banjir di Kota Medan dengan Evaluasi Silhouette Coefficient," *Algoritma : Jurnal Matematika, Ilmu pengetahuan Alam, Kebumihan dan Angkasa*, vol. 2, no. 6, pp. 170–181, Nov. 2024, doi: 10.62383/algoritma.v2i6.270.



- [16] S. Paembonan and H. Abduh, "Penerapan Metode Silhouette Coeficient Untuk Evaluasi Clustering Obat Clustering; K-means; Silhouette coeficient," *PENA TEKNIK*, vol. 6, no. 2, pp. 48–54, Sep. 2021, doi: 10.51557/pt_jiit.v6i2.659.
- [17] Ardiansyah, Risnita, and M. S. Jailani, "Teknik Pengumpulan Data Dan Instrumen Penelitian Ilmiah Pendidikan Pada Pendekatan Kualitatif dan Kuantitatif," *IHSAN: Jurnal Pendidikan Islam*, vol. 1, no. 2, pp. 1–9, Apr. 2023, doi: <https://doi.org/10.61104/ihsan.v1i2.57>.
- [18] Suriadi, M. M. Maq, Randi, M. Karina, and R. Rais, "Pelatihan Teknis Pembuatan Kuosioner Penelitian Melalui Aplikasi Google Form Dalam Meningkatkan Kualitas Hasil Penelitian Bagi Dosen Pemula Di Era Digital," *Journal of Human And Education*, vol. 4, no. 3, pp. 538–544, 2024, doi: <https://doi.org/10.31004/jh.v4i3.1069>.
- [19] A. R. Pratama, W. Irsyad, R. H. Hassan, and M. Rawati, "Pemanfaatan Google Form Sebagai Alat Evaluasi Pada Pembelajaran Pendidikan Agama Islam," *Pendidikan Tunas Bangsa*, vol. 2, no. 1, pp. 19–30, Apr. 2024, doi: 10.58222/jptunasbangsa.v2i1.890.
- [20] M. I. Rahayu, R. Antika, S. Anggraeni, and R. Wijaya, "Sistem Kuesioner Evaluasi Layanan Perkuliahan dan Kegiatan Kampus Stmik Bandung Terintegrasi," *Jurnal Teknologi Informasi dan Komunikasi*, vol. 13, no. 2, pp. 17–25, Dec. 2024, doi: <https://doi.org/10.58761/jurtikstmikbandung.v13.i2.1725>.
- [21] Saman, "Persepsi Mahasiswa Universitas Muhammadiyah Palopo Terhadap Pelaksanaan Pembelajaran Daring Saat Pandemi COVID-19," *Jurnal Manajemen dan Ilmu Pendidikan*, vol. 3, no. 1, pp. 70–82, Feb. 2021, doi: 10.36088/manazhim.v3i1.1069.
- [22] L. Santoso and Priyadi, "Mengoptimalkan Proses Pembersihan Data dalam Analisis Big Data Menggunakan Pipeline Berbasis AI," *Jurnal Elektronika dan Komputer*, vol. 17, no. 2, pp. 657–666, Dec. 2024, doi: 10.51903/elkom.v17i2.2311.
- [23] Dr. S. R. Sukhdeve and S. S. Sukhdeve, *Google Colaboratory*. Apress, Berkeley, CA, 2023. doi: 10.1007/978-1-4842-9688-2_2.
- [24] Nabila Sava Nur Az-zahrani, Hana Kurnia Asti Eloji, Feilisia Salim, Az-Zahra Aqeela Ramadhani, Charel Meysyanti, and Laurensia Nadia Ayu Purwantiningsih, *Python untuk Analisis Data*. Semarang: SIEGA, 2025. Accessed: Jan. 30, 2026. [Online]. Available: https://books.google.co.id/books?hl=en&lr=&id=BthpEQAAQBAJ&oi=fnd&pg=PP8&dq=Google+Cola+b+sangat+membantu+dalam+alur+kerja+analisis+data+karena+bisa+dilacak+kembali,+diulang,+kemudi+an+dikembangkan+lebih+lanjut&ots=7a2ie60I0c&sig=P8b2ybLImZwtvOYPqfXIn4pQCY&redir_esc=y#v=onepage&q&f=false.
- [25] A. W. Fuadah, N. Arifin, and O. Juwita, "Optimasi K-Klasterisasi Ketahanan Pangan Kabupaten Jember Menggunakan Metode Elbow," *Informatics Journal*, vol. 6, no. 3, pp. 136–141, 2021, doi: 10.19184/isj.v6i3.28363.
- [26] D. Lestari, Abd. C. Fauzan, and Harliana, "Penerapan Algoritma Pillar Untuk Optimasi Penentuan Titik Awal Centroid Pada Algoritma K-Means Clustering," *JOISIE Journal of Information System and Informatics Engineering*, vol. 6, no. 1, pp. 15–24, Jun. 2022, doi: <https://doi.org/10.35145/joisie.v6i1.2053>.
- [27] A. Putra, D. Abdullah, and M. Daud, "Menggunakan Metode K-Harmonic Means dengan Validasi Silhouette Index dan C-index," *Jurnal Janitra Informatika dan Sistem Informasi*, vol. 4, no. 2, pp. 74–86, Oct. 2024, doi: 10.59395/janitra.v4i2.203.
- [28] R. Akbar, U. S. Sukmawati, and K. Katsirin, "Analisis Data Penelitian Kuantitatif," *Jurnal Pelita Nusantara*, vol. 1, no. 3, pp. 430–448, Jan. 2024, doi: 10.59996/jurnalpelitanusantara.v1i3.350.
- [29] B. J. Forester, A. Idris, A. Khater, M. W. Afgani, and M. Isnaini, "Penelitian Kuantitatif: Uji Reliabilitas," *Edu Society: Jurnal Pendidikan, Ilmu Sosial, dan Pengabdian Kepada Masyarakat*, vol. 4, no. 3, pp. 1812–1820, Dec. 2024, doi: <https://doi.org/10.56832/edu.v4i3.577>.