

Evaluasi Performa Mahasiswa pada Pembelajaran Mata Kuliah Data Analitik Menggunakan K-Means Clustering: Studi Kasus di Telkom University

Luh Putri Kirana Maharani¹, Adinda Nathania Perangin Angin¹,
Pramasta Rashad Wardhana¹, Fandi Achmad^{1*}

¹ Program Studi Teknik Industri, Telkom University, Bandung, 40257, Indonesia

*Penulis Koresponden: fandiac.achmad@gmail.com

Abstrak

Pendidikan tinggi mengalami transformasi signifikan guna mempersiapkan mahasiswa menghadapi tuntutan dunia kerja yang semakin kompleks. Telkom University sebagai perguruan tinggi telah membuktikan dedikasinya dalam menyelaraskan pengetahuan akademis dengan keterampilan praktis. Dalam konteks meningkatnya permintaan ahli teknologi dan data science, mata kuliah Data Analitik di Telkom University menjadi krusial. Oleh karena itu, tujuan penelitian ini untuk menganalisis dan mengevaluasi performa mahasiswa pada mata kuliah Data Analitik di Telkom University menggunakan metode K-Means clustering. Identifikasi pola kinerja mahasiswa berdasarkan aspek-aspek kritis seperti critical thinking, collaborative skill, creativity and innovation, technology application, participation level, dan overall achievement. Hasil penelitian menunjukkan dua klaster yakni high performance (klaster 0) dan low performance (klaster 1). Cluster high performance menunjukkan rendahnya nilai pada X5 (pemecahan masalah kompleks) dan X11 (kemampuan berinovasi), sementara cluster low performance menunjukkan rendahnya nilai pada X7 (keaktifan kolaborasi) dan X10 (perolehan motivasi). Temuan ini dapat membantu pengembangan strategi optimal dalam meningkatkan performa pada mahasiswa.

Kata Kunci: Performa, K-Means, Clustering, Data Analitik.

Abstract

Higher education is undergoing a significant transformation to prepare students to face the demands of an increasingly complex world of work. As a tertiary institution, Telkom University has proven its dedication to aligning academic knowledge with practical skills. In the increasing demand for technology and data science experts, the Data Analytics course at Telkom University is crucial. Therefore, this research aims to analyze and evaluate student performance in the Data Analytics course at Telkom University using the K-Means clustering method. Identify student performance patterns based on critical aspects such as critical thinking, collaborative skill, creativity and innovation, technology application, participation level, and overall achievement. The research results show two clusters: high performance (cluster 0) and low performance (cluster 1). The high-performance cluster shows low scores on X5 (complex problem solving) and X11 (ability to innovate),

while the low-performance cluster shows low scores on X7 (active collaboration). These findings can help develop optimal strategies for improving student performance.

Keywords: *Performance, K-Means, Clustering, Data Analytics.*

A. PENDAHULUAN

Pendidikan tinggi saat ini mengalami transformasi signifikan untuk mempersiapkan mahasiswa dalam menghadapi tuntutan dunia kerja yang semakin kompleks (Akour *et al.*, 2022; Abu Talib *et al.*, 2021). Pendidikan bukan hanya pengetahuan melainkan proses bagaimana seseorang membangun karakter yang dapat bermakna bagi dirinya sendiri, orang lain, dan juga negaranya (Galuh *et al.*, 2021). Konsep ini sejalan Herlambang (2021), yang menekankan bahwa pendidikan bukanlah sekadar akuisisi pengetahuan, melainkan proses dinamis bagaimana seseorang membentuk karakter yang memiliki makna dalam konteks pribadi, interaksi sosial, dan kepentingan nasional. Oleh karena itu, pendidikan tinggi harus mampu memberikan pengalaman pembelajaran yang lebih holistik dan relevan dengan kebutuhan pasar kerja.

Salah satu program yang mendukung untuk menyiapkan kebutuhan kerja adalah program Kurikulum Merdeka Belajar Kampus Merdeka (MBKM). MBKM memiliki sistem pendidikan yang lebih ke arah student-centered learning. Program MBKM diharapkan dapat menjawab tantangan Perguruan Tinggi untuk menghasilkan lulusan yang sesuai perkembangan zaman, kemajuan IPTEK, tuntutan dunia usaha dan dunia industri, maupun dinamika masyarakat (Kampus Merdeka, 2023). Salah satu perguruan tinggi yang sudah menerapkan MBKM adalah Telkom University. Telkom University merupakan salah satu perguruan tinggi swasta terbaik yang secara konsisten menunjukkan dedikasi tinggi dalam menyiapkan lulusannya untuk mencapai kesuksesan di dunia kerja. Dengan berbagai program akademis yang komprehensif dan berkualitas, Telkom University telah membuktikan diri sebagai lembaga pendidikan tinggi yang mampu memadukan pengetahuan akademis dengan keterampilan praktis yang dibutuhkan oleh industri.

Menurut Data science Indonesia (2023), demand ahli di bidang teknologi dan data science di dunia kerja sangat meningkat. Hal tersebut mengartikan bahwa pentingnya skills di bidang teknologi dan data science dapat meningkatkan kesempatan karir. Dengan meningkatnya permintaan terhadap ahli di bidang ini, pentingnya mengembangkan keahlian dalam teknologi dan data science menjadi semakin nyata (Aurachman, 2018). Salah

satu mata kuliah di Telkom University yang menunjang di bidang teknologi dan data science adalah mata kuliah Data Analitik. Mata kuliah Data Analitik memiliki peran krusial dalam pembekalan keterampilan esensial. Evaluasi performa mahasiswa dalam mata kuliah ini menjadi sangat penting untuk menilai efektivitas pembelajaran dan memastikan bahwa lulusan memiliki keterampilan yang relevan dengan kebutuhan pasar kerja saat ini. Evaluasi performa mahasiswa tidak hanya sebatas pengukuran penguasaan konsep-konsep Data Analitik, tetapi juga perlu menyoroti keterampilan kritis yang menjadi landasan bagi kesuksesan di dunia kerja (AlMahdawi *et al.*, 2021). Dalam konteks ini, variabel seperti *critical thinking*, *collaborative skill*, *creativity and innovation*, *technology application*, *participation level*, dan *overall achievement* menjadi krusial (Achmad *et al.*, 2023; AlMahdawi *et al.*, 2021).

Oleh karena itu, penelitian ini bertujuan untuk menyelidiki dan menganalisis evaluasi performa mahasiswa pada mata kuliah Data Analitik di Telkom University menggunakan metode *K-Means clustering*. Identifikasi pola kinerja mahasiswa dalam aspek-aspek kritis seperti *critical thinking*, *collaborative skill*, *creativity and innovation*, *technology application*, *participation level*, dan *overall achievement* diharapkan dapat memberikan wawasan mendalam terkait dengan sejauh mana mahasiswa dapat menguasai materi dan menerapkan keterampilan kritis yang relevan dalam konteks mata kuliah Data Analitik.

Hasil penelitian ini dapat memberikan implikasi yang berharga dalam pengembangan kurikulum mata kuliah, serta melakukan penyesuaian dan meningkatkan materi ajar serta metode pembelajaran yang digunakan. Evaluasi yang lebih terfokus pada aspek-aspek ini dapat memberikan gambaran yang lebih holistik mengenai kemampuan mahasiswa. Dengan menilai keterampilan kritis ini, mahasiswa dapat lebih baik dipersiapkan untuk menghadapi kompleksitas dunia kerja yang membutuhkan tidak hanya pemahaman konsep, tetapi juga penerapan keterampilan praktis dalam mengatasi tantangan di bidang teknologi dan data science.

B. METODE PENELITIAN

1. Pengumpulan Data

Tujuan penelitian ini untuk menganalisis dan mengevaluasi performa mahasiswa pada mata kuliah Data Analitik di Telkom University. Pengumpulan data dilakukan dengan metode studi literatur melalui data primer yaitu data hasil kuesioner yang berisi pertanyaan-pertanyaan yang menjadi tolak ukur dari performansi mahasiswa di mata kuliah Data Analitik. Metode survei lapangan dilakukan melalui observasi terstruktur terhadap 90

mahasiswa yang mengambil mata kuliah tersebut sebagai sampel pada bulan Agustus - September 2023, yang ditunjukkan pada Tabel 1. Observasi dilakukan melalui pengisian survei oleh mahasiswa yang diberikan oleh dosen pengampu.

Tabel 1. Sampel Penelitian

No	Nama Mahasiswa	Kelas
1	<i>Student 1</i> <i>Student 2</i> ... <i>Student 15</i>	TI-45-03
2	<i>Student 16</i> <i>Student 17</i> ... <i>Student 30</i>	TI-45-04
3	<i>Student 31</i> <i>Student 32</i> ... <i>Student 45</i>	TI-45-05
4	<i>Student 46</i> <i>Student 47</i> ... <i>Student 60</i>	TI-45-06
5	<i>Student 61</i> <i>Student 62</i> ... <i>Student 75</i>	TI-45-07
6	<i>Student 76</i> <i>Student 77</i> ... <i>Student 90</i>	TI-45-08

Selanjutnya Tabel 2 merupakan definisi atribut yang akan digunakan dalam penelitian ini (AlMahdawi *et al.*, 2021).

Tabel 2. Definisi Atribut

Atribut	Definisi	Kode	Indikator
<i>Critical thinking</i>	Suatu kemampuan untuk mengevaluasi secara kritis suatu kepercayaan atau keyakinan.	X1	Pemahaman konsep baru
		X2	Project/tugas/quiz
		X3	Pemahaman pengaplikasian data analitik
		X4	<i>Challenging</i>
		X5	Pemecahan masalah kompleks
<i>Collaborative skills</i>	kemampuan untuk bekerja secara efektif dalam tim dan memiliki kemauan untuk membuat keputusan yang diperlukan demi mencapai tujuan bersama.	X6	Bekerja dalam kelompok
		X7	Keaktifan kolaborasi
		X8	Presentasi tugas di depan kelas
<i>Creativity and innovation</i>	Kemampuan untuk menghasilkan ide-ide baru, cara pandang terhadap masalah dan menerapkan solusi-solusi kreatif terhadap masalah dan peluang.	X9	Ide dan gagasan
		X10	Motivasi
		X11	Kemampuan berinovasi
<i>Technology application</i>	Penerapan teknologi dalam konteks praktis atau aplikatif untuk memecahkan masalah, meningkatkan efisiensi, atau mencapai tujuan tertentu dalam berbagai bidang dan kegiatan.	X12	Penggunaan teknologi
		X13	Kemampuan menggunakan teknologi
		X14	Pemahaman materi jika menggunakan teknologi
<i>Participation level</i>	Suatu kegiatan belajar mengajar yang melibatkan interaksi antara pendidik dan peserta didik.	X15	Keterlibatan dalam kelas
		X16	Ketepatan waktu/ kehadiran
<i>Overall achievement</i>	Pencapaian keseluruhan atau prestasi secara umum	X17	Nilai yang didapatkan

Sumber: AlMahdawi *et al.*, 2021.

Keenam atribut ini memiliki peran yang penting di dalam penelitian, ini dikarenakan keenam atribut seperti *critical thinking, collaborative skill, creativity and innovation, technology application, participation level*, dan *overall achievement* memiliki peran sebagai alat ukur dari performansi mahasiswa. Dalam atribut terdapat pengukuran yang memiliki *rate 1-5*, di mana setiap mahasiswa yang diberikan kuesioner dapat menjawab dengan memilih *rate 1-5* sesuai indikatornya, sebagai contoh atribut *critical thinking* dengan indikator pemahaman konsep baru di pilih dengan *rate 5* yaitu sangat mudah.

2. K-means Clustering

K-Means merupakan algoritma *clustering* yang membagi data menjadi kelompok sehingga setiap data masuk ke dalam kelompok yang memiliki *centroid* terdekat (Achmad *et al.*, 2023). Menurut Ikotun *et al.*, (2023) *K-Means clustering* merupakan suatu metode dalam pengelompokan data yang termasuk dalam kategori *unsupervised learning*, yakni data dibagi menjadi beberapa kelompok berdasarkan jarak antar data. Algoritma ini efektif digunakan ketika data yang dihadapi bersifat numerik (Fa’rifah *et al.*, 2023). Dengan kata lain, tidak ada informasi awal mengenai klasifikasi atau kelompok mana yang dimiliki oleh setiap data.

Dalam implementasinya, algoritma *K-Means* bekerja dengan cara mengelompokkan data ke dalam kluster-kluster berdasarkan kedekatan jarak antar data numerik. Setiap *cluster* yang dihasilkan akan mencerminkan kesamaan karakteristik atau pola tertentu dalam data. Algoritma ini memiliki keterbatasan, yaitu hanya dapat digunakan pada atribut atau variabel yang berupa angka atau numerik (Septiningrum *et al.*, 2022). Oleh karena itu, *K-Means* menjadi pilihan yang tepat ketika kita berurusan dengan *dataset* yang bersifat kuantitatif dan ingin mengidentifikasi struktur atau pola bantuan label atau kelas yang sudah ditetapkan sebelumnya (Fa’rifah *et al.*, 2023). Rumus umum dalam metode klusterisasi *K-Means* ialah sebagai berikut:

- a. Menentukan nilai *centroid* (titik pusat) dari setiap kelompok, nilai ini diambil dari nilai rata-rata (*means*) dari keseluruhan data. Rumus umum dalam *K-Means* ditunjukkan pada Gambar 1 yang dimana *M* menyatakan jumlah data dalam suatu kelompok dan *I* menyatakan fitur ke-*i* dalam sebuah kelompok.

$$C_i = \frac{1}{M} \sum_{j=1}^M x_j$$

Gambar 1. Rumus *K-Means*

- b. Menghitung jarak titik terdekat (*Euclidean Distance*) dengan menggunakan rumus pada Gambar 2.

$$D(x_2, x_1) = \sqrt{\sum_{j=1}^p |x_{2j} - x_{1j}|^2}$$

Gambar 2. *Euclidian Distance*

- c. Mengalokasikan keanggotaan dari setiap titik menggunakan rumus pada Gambar 3.

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

Gambar 3. Rumus Alokasi Titik

- d. Menentukan fungsi objektif suatu data melalui formula yang terlampir pada Gambar 4.

$$F = \sum_{j=1}^N \sum_{i=1}^K a_{ji} D(x_j, C_i)$$

Gambar 4. Fungsi Objektif

3. Tahapan Kmeans Clustering

a. *Eksplorasi Data Frame*

Pengolahan data pada penelitian ini menggunakan google colab. Hasil data kuesioner yang diperoleh dilakukan eksplorasi data dengan cara melakukan input data untuk menampilkan hasil rekap data yang ada. Eksplorasi data dilakukan untuk menganalisa dan menginvestigasi kumpulan data, meringkas karakteristik utama dari data ke metode visualisasi. Hal ini dapat membantu menemukan pola data, melihat anomaly, pengujian hipotesis, dan asumsi.

b. *Data Preprocessing: Normalization*

Tahap ke dua ialah data *pre-processing* dengan melakukan normalisasi *z-score* dengan menggunakan fungsi *standard scaler* yang kemudian hasilnya disimpan sebagai sebuah data frame dengan nama *standardized attributes*.

c. *Model Turning*

Pada algoritma *k-means clustering* terdapat nilai parameter k, sehingga sebelum diputuskan subjektif nilai k, dapat dilakukan model tuning, misalnya dengan metode *Elbow*. Untuk melakukan analisis nilai k optimal

dengan metode Elbow, dapat dibuat program sederhana plot antara nilai *within-cluster some of squares* atau inerti model untuk setiap nilai k yang di setting pada fungsi k -means. Hal ini dilakukan secara *looping* menggunakan fungsi `for...in`. Nilai inerti akan semakin kecil dengan bertambahnya jumlah *cluster* dengan nilai k optimal diperoleh dari titik balik, yang mana penurunan inerti sudah relatif “kecil”, dicirikan dengan kurva mulai landai. Selanjutnya, untuk lebih meyakinkan berapa nilai k optimal, dilakukan analisa *silhouette coefficient* berdasarkan konfigurasi setiap anggota *cluster*.

d. *Model Building*

Setelah nilai k terpilih, selanjutnya dilakukan pembangunan model *clustering* dengan menggunakan fungsi *k-means* dan *input* atribut-atribut yang sudah dinormalisasi. Hasil dari *clustering* disajikan dalam sebuah *dataframe* dengan nama “hasil”.

e. *Model Visualization*

Hasil dari *clustering* selanjutnya dapat dilakukan visualisasi ke dalam plot 2 dimensi. Dengan jumlah atribut sebanyak “angka”, untuk membuat plot 2 dimensi dapat dilakukan teknik reduksi data dengan PCA.

C. HASIL DAN PEMBAHASAN

Tujuan penelitian ini untuk menyelidiki dan menganalisis evaluasi performa mahasiswa pada mata kuliah Data Analitik di Telkom University menggunakan metode *K-Means clustering*. Identifikasi pola kinerja mahasiswa dalam aspek-aspek kritis seperti *critical thinking, collaborative skill, creativity and innovation, technology application, participation level*, dan *overall achievement* diharapkan dapat memberikan wawasan mendalam terkait dengan sejauh mana mahasiswa dapat menguasai materi dan menerapkan keterampilan kritis yang relevan dalam konteks mata kuliah Data Analitik. Berikut merupakan langkah-langkah dalam pengolahan data.

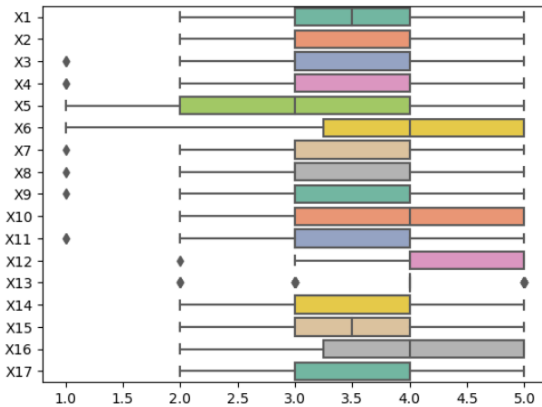
1. *Exploratory Data Frame*

Pengolahan data pada penelitian ini menggunakan google colab. Hasil data kuesioner yang diperoleh dilakukan terlebih dahulu eksplorasi data yang ada dengan cara melakukan input data untuk menampilkan hasil rekap data yang ada. *Exploratory data frame* yang dilakukan ditampilkan pada Gambar 5.

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16	X17
count	90.000000	90.000000	90.000000	90.000000	90.000000	90.000000	90.000000	90.000000	90.000000	90.000000	90.000000	90.000000	90.000000	90.000000	90.000000	90.000000	90.000000
mean	3.522222	3.555556	3.355556	3.611111	2.944444	4.022222	3.733333	3.433333	3.211111	4.022222	3.300000	4.344444	3.988889	3.900000	3.577778	4.177778	3.833333
std	0.657332	0.582731	0.890791	1.002182	0.878726	0.834344	0.871522	0.887491	0.867574	0.847704	0.866998	0.689587	0.756907	0.719707	0.752915	0.943074	0.722822
min	2.000000	2.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	2.000000	1.000000	2.000000	2.000000	2.000000	2.000000	2.000000	2.000000
25%	3.000000	3.000000	3.000000	3.000000	2.000000	3.250000	3.000000	3.000000	3.000000	3.000000	3.000000	4.000000	4.000000	3.000000	3.000000	3.250000	3.000000
50%	3.500000	4.000000	3.000000	4.000000	3.000000	4.000000	4.000000	3.000000	3.000000	4.000000	3.000000	4.000000	4.000000	4.000000	4.000000	3.500000	4.000000
75%	4.000000	4.000000	4.000000	4.000000	4.000000	5.000000	4.000000	4.000000	4.000000	5.000000	4.000000	5.000000	4.000000	4.000000	4.000000	4.000000	4.000000
max	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000

Gambar 5. Definisi *Data Frame*

Langkah selanjutnya adalah memeriksa *data outlier* menggunakan *Box Plot* yang terlampir pada gambar 6. *Box plot* atau diagram kotak merupakan metode dalam statistik deskriptif yang digunakan untuk secara visual mengilustrasikan distribusi data numeris (Abbas *et al.*, 2020). Dan hasil *box plot* dalam *exploratory data frame* dapat dilihat pada Gambar 6.



Gambar 6. *Box plot* Sebelum Normalisasi

2. *Data Pre-processing: Normalization*

Data preprocessing merupakan tahapan perubahan data mentah menjadi data yang dapat mudah dimengerti. Normalisasi merupakan suatu kegiatan menormalisasi data dengan tujuan agar data memiliki nilai nol dan standar deviasi yang berjumlah satu. Gambar 7 menunjukkan bahwa penggunaan kode '*StandardScaler*' digunakan untuk menstandarisasikan data dalam yaitu *Data Frame* ('df').

```
[8] from sklearn.preprocessing import StandardScaler
scale=StandardScaler()
std_atr=scale.fit_transform(df)
std_atr=pd.DataFrame(std_atr,columns=df.columns)
print(std_atr)

      X1      X2      X3      X4      X5      X6      X7 \
0 -2.328733  0.766965 -0.401382 -0.613196  0.063577 -0.026784  0.307692
1  0.730916  0.766965  1.856392  1.393628  1.207963  1.178477  1.461538
2  0.730916 -0.958786  0.727505 -0.613196  0.063577 -0.026784  0.307692
3 -0.798909 -0.958786 -0.401382 -1.616609 -1.080809 -0.026784 -0.846154
4 -0.798909  0.766965  0.727505  1.393628  1.207963 -1.232045 -0.846154
..      ..      ..      ..      ..      ..      ..
85  0.730916  0.766965 -0.401382  0.390216  1.207963  1.178477  1.461538
86  0.730916 -0.958786 -2.659157 -0.613196  1.207963  1.178477  1.461538
87  0.730916  0.766965 -0.401382  1.393628  0.063577 -0.026784  1.461538
88 -0.798909  0.766965 -0.401382  0.390216  1.207963 -1.232045  0.307692
89  0.730916  0.766965  0.727505  1.393628 -1.080809  1.178477  0.307692

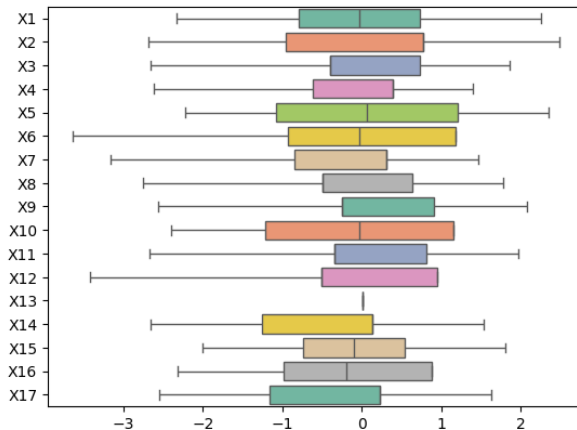
      X8      X9      X10      X11      X12      X13      X14 \
0 -0.491003 -0.244698 -0.026361 -0.347960 -0.502292  0.014762 -1.257515
1 -0.491003  0.914399  1.159905  1.971774  0.955975  0.014762  1.536963
2  0.642081 -0.244698 -0.026361  0.811907 -0.502292 -1.313805 -1.257515
3  0.642081 -0.244698 -0.026361  0.811907 -0.502292  0.014762 -1.257515
4 -0.491003  0.914399 -0.026361  0.811907 -1.960559 -1.313805  0.139724
..      ..      ..      ..      ..      ..      ..
85 -0.491003  0.914399  1.159905 -1.507827 -1.960559 -1.313805 -1.257515
86  1.775166  0.914399  1.159905  0.811907  0.955975  1.343329  1.536963
87  0.642081 -0.244698 -0.026361 -0.347960  0.955975  0.014762  0.139724
88  0.642081  2.073495 -0.026361  0.811907  0.955975  0.014762  0.139724
89  0.642081 -0.244698  1.159905 -0.347960  0.955975  1.343329  0.139724

      X15      X16      X17 \
0 -0.732758 -1.255868  0.231869
1  1.803712  0.876738  1.623086
2  0.535477 -1.255868  1.623086
3 -0.732758 -0.189565 -1.159347
4 -0.732758  0.876738  0.231869
..      ..      ..
85 -0.732758  0.876738  0.231869
86  1.803712  0.876738  1.623086
87  0.535477  0.876738  0.231869
88 -0.732758  0.876738  0.231869
89 -0.732758  0.876738  0.231869

[90 rows x 17 columns]
```

Gambar 7. Standard Scaler

Pada fase ini dilakukan penskalaan standar data numerik ke dalam *data frame* dengan output yang dihasilkan berupa kolom berdasarkan *'numeric_df.columns'*. setelah melakukan normalisasi data, maka diperoleh hasil *boxplot* yang terlampir pada Gambar 8.



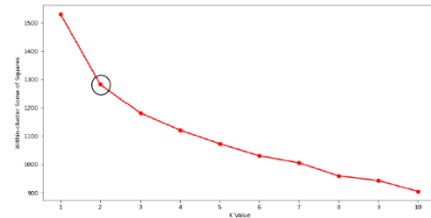
Gambar 8. Boxplot Setelah Normalisasi

3. Model Turning

Proses model *turning* merupakan proses yang dapat dilakukan setelah mendapatkan *dataset* yang dimulai dengan menentukan tujuan yang ingin dicapai dengan memilih algoritma, dalam kasus ini adalah *clustering*, yang ditunjukkan pada Gambar 9. Dalam algoritma *k-means clustering*, terdapat nilai paramter *k* sehingga sebelum keputusan subjektif dibuat pada nilai *k*, *model turning* dapat dilakukan dengan metode Elbow yang dapat di lihat pada Gambar 10.

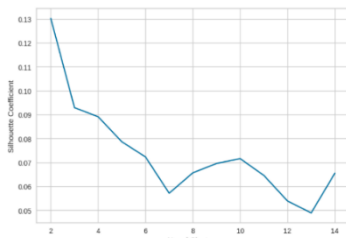
```
[ ] from sklearn.cluster import KMeans
wcss=[]
for i in range(1,11):
    km=KMeans(n_clusters=i)
    km.fit(std_atr)
    wcss.append(km.inertia_)
#The elbow curve
plt.figure(figsize=(12,6))
plt.plot(range(1,11),wcss)
plt.plot(range(1,11),wcss,linewidth=2,color="red",marker="8")
plt.xlabel("K Value")
plt.xticks(np.arange(1,11,1))
plt.ylabel("Within-cluster Some of Squares")
plt.show()
```

Gambar 9. Model Tuning



Gambar 10. Grafik Kluster

Metode *elbow* adalah metode untuk menentukan jumlah *cluster* yang benar melalui persentase hasil perbandingan antara jumlah *cluster* yang membentuk sudut pada suatu titik. Hasil analisis menggunakan metode *elbow* dari Gambar 10 adalah ($k=2$). Dan untuk memperkuat gagasan bahwa nilai ($k=2$) maka dapat melakukan perbandingan kualitas pengelompokan dengan metode *silhouette coefficient* pada Gambar 11. Hal ini bertujuan untuk mengukur hubungan antara objek dalam *cluster* dan metode pemisahan yang mengukur sejauh mana *cluster* dapat dipisah.



Gambar 11. Grafik silhouette

```
For n_clusters=2, Silhouette Coefficient = 0.1302917363598488
For n_clusters=3, Silhouette Coefficient = 0.09296025526612069
For n_clusters=4, Silhouette Coefficient = 0.08915680264168283
For n_clusters=5, Silhouette Coefficient = 0.0786911392492921
For n_clusters=6, Silhouette Coefficient = 0.07234611522630954
For n_clusters=7, Silhouette Coefficient = 0.057163584444635145
For n_clusters=8, Silhouette Coefficient = 0.06569759504022138
For n_clusters=9, Silhouette Coefficient = 0.06960222076101652
For n_clusters=10, Silhouette Coefficient = 0.07162378687273976
For n_clusters=11, Silhouette Coefficient = 0.06460959199502425
For n_clusters=12, Silhouette Coefficient = 0.05390318521273501
For n_clusters=13, Silhouette Coefficient = 0.04892849624717842
For n_clusters=14, Silhouette Coefficient = 0.06552893620340311
```

Gambar 12. Output silhouette coefficient

Gambar 11 menunjukkan hasil dari grafik koefisien siluet yang dimana setiap jumlah *cluster* memiliki nilai koefisien siluet yang berbeda. Koefisien siluet ini digunakan untuk memilih jumlah *cluster* (*n-clusters*). Nilai yang lebih tepat untuk masing-masing (*n-clusters*) dapat dilihat pada gambar 12. Nilai koefisien siluet yang sesuai dengan skor siluet adalah (*n-cluster*) = 2 karena memiliki nilai skor siluet 0.130.

4. *Model Building*

Proses model building pada metode *k-means* clustering melibatkan beberapa tahapan kritis untuk membentuk kelompok atau kluster dari data yang ada. Tahap awal melibatkan inisialisasi centroids, di mana kita secara acak memilih K titik data sebagai pusat awal kluster. Setelah inisialisasi, data di atribusikan ke kluster berdasarkan jarak *euclidean* antara setiap data dan pusat kluster terdekat. Pusat kluster kemudian dihitung ulang dengan mengambil rata-rata dari data dalam kluster tersebut. Gambar 13 merupakan *output data frame model building* dan Gambar 14 merupakan *output hasil kluster*.

```
inertia: 1282.0469105883649
cluster_centroids: [[ 0.45894745  0.38348249  0.35120936  0.16723541  0.08900782  0.24105221
 0.66666667  0.54136253  0.3992446  0.60631378  0.09021188  0.21063855
 0.42809383  0.35707227  0.45092791  0.23695618  0.54102871]
[-0.45894745 -0.38348249 -0.35120936 -0.16723541 -0.08900782 -0.24105221
-0.66666667 -0.54136253 -0.3992446 -0.60631378 -0.09021188 -0.21063855
-0.42809383 -0.35707227 -0.45092791 -0.23695618 -0.54102871]]
```

Gambar 13. *Output Data Frame Model Building*

```
[ ] clt=model.labels_
hasil=pd.Series(clt,name="Clusters")
hasil=pd.DataFrame(hasil)
hasil
```

Clusters	
0	1
1	0
2	0
3	1
4	1
...	...
85	0
86	0
87	0
88	0
89	0

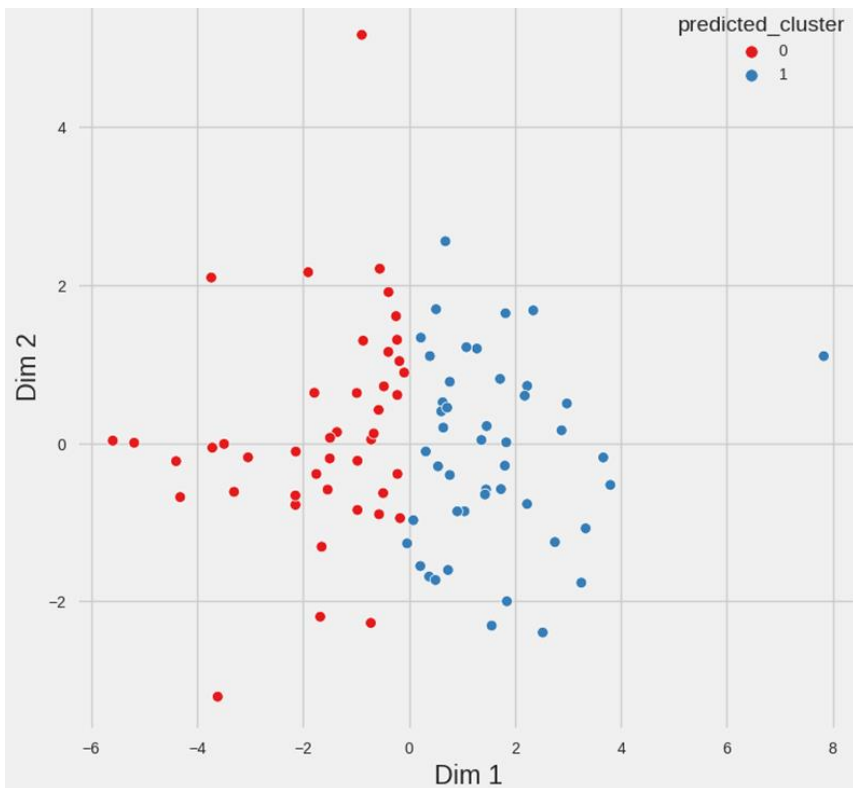
90 rows x 1 columns

Gambar 14. *Output Hasil Kluster*

Berdasarkan terpilihnya 2 kluster maka pada tahap pembangunan model ini menunjukkan urutan nomor Mahasiswa yang masuk ke masing-masing kategori *cluster* yakni 0 dan 1.

5. **Model Visualization**

Gambar 18 menunjukkan hasil dari performa mahasiswa mata kuliah data analitik. Berdasarkan hasil visualisasi segmentasi kluster, titik-titik objek termasuk dalam dua kluster yakni kluster 0 dengan warna merah dan kluster 1 memiliki warna biru. Sentroid merupakan nilai pusat dari kluster, dalam kasus ini $k=2$; maka *centroid* C1 dan C2 akan terbentuk secara acak. Tahap ini akan menghitung jarak masing-masing data ke sentroid yang telah dibuat menggunakan jarak *Euclidean*. Nilai sentroid untuk masing-masing *variable* mempengaruhi penilaian kluster yang terbentuk. Nilai sentroid untuk *variable* kluster dijelaskan pada Tabel 3.



Gambar 18. *Model Visualization 2D*

Tabel 3. Nilai Sentroid

Cluster	Centroids																
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16	X17
0	0,45895	0,38348	0,35121	0,16724	0,08901	0,24105	0,66667	0,54136	0,39924	0,60631	0,09021	0,21064	0,42809	0,35707227	0,45093	0,23696	0,54103
1	-0,4589	-0,3835	-0,3512	-0,1672	-0,089	-0,2411	-0,6667	-0,5414	-0,3992	-0,6063	-0,0902	-0,2106	-0,4281	-0,35707227	-0,4509	-0,237	-0,541

Berdasarkan Tabel 3 nilai *centroid* pada setiap *variable* mempengaruhi penelitian *cluster* yang terbentuk. Pada kluster 0, nilai *centroid* yang memiliki nilai paling signifikan adalah keaktifan kolaborasi dan motivasi dengan skor masing-masing 0.66667 dan 0.60631. Sedangkan pada kluster 1 nilai terbesar adalah pemecahan masalah kompleks dan kemampuan berinovasi dengan skor masing-masing adalah -0.089 dan -0.0902.

6. *Evaluation Result*

Langkah selanjutnya adalah mengidentifikasi anggota dari masing-masing kluster performansi mahasiswa di mata kuliah data analitik. Hasil pengelompokan berdasarkan keenam atribut yang memiliki indikator yaitu pemahaman konsep baru, *project/tugas/quiz*, pemahaman pengaplikasian data analitik, *challenging*, pemecahan masalah kompleks, bekerja dalam kelompok, keaktifan kolaborasi, presentasi tugas di depan kelas, ide dan gagasan, motivasi, kemampuan berinovasi, penggunaan teknologi, kemampuan menggunakan teknologi, pemahaman materi jika pembelajaran menggunakan teknologi, keterlibatan dalam kelas, ketepatan waktu/kehadiran, dan nilai yang di dapatkan. Hal ini menunjukkan bahwa kluster terbaik adalah performansi mahasiswa yang penilaian indikator atributnya yang tinggi. Penjelasan lebih rinci mengenai kluster performansi mahasiswa di mata kuliah data analitik berdasarkan nilai *centroid* masing-masing variabel disajikan pada Tabel 4.

Tabel 4. *Cluster Value*

<i>Cluster</i>	<i>Cluster value</i>	<i>Cluster Name</i>
0	0.365850097	<i>High performance</i>
1	-0.365850097	<i>Low performance</i>

Tabel 3 menunjukkan bahwa setiap kluster memiliki nama. Kluster 0 memiliki nilai kluster 0,365850097 berada pada posisi tinggi, yang berarti

setiap mahasiswa yang berada pada kluster 0 disebut dengan mahasiswa dengan *High performance*. Kluster 1 memiliki nilai kluster sebesar -0.365850097 berada pada posisi rendah, yang berarti setiap mahasiswa yang berada pada kluster 1 disebut sebagai mahasiswa dengan *Low performance*. Tabel 5 merupakan hasil daftar mahasiswa beserta golongan klusternya.

Tabel 5. Hasil *Cluster* Performa Mahasiswa

Cluster Name	Nama Mahasiswa		Student
High Performance	Student 2	Low Performance	Student 1
	Student 3		Student 4
	Student 9		Student 5
	student 10		Student 6
	Student 11		Student 7
	Student 14		Student 8
	Student 15		Student 12
	Student 17		Student 13
	Student 18		Student 16
	Student 19		Student 21
	Student 20		Student 22
	Student 24		Student 23
	Student 26		Student 25
	Student 28		Student 27
	Student 29		Student 31
	Student 30		Student 33
	Student 32		Student 35
	Student 34		Student 36
	Student 39		Student 37
	Student 40		Student 38
	Student 43		Student 41
	Student 44		Student 42
	Student 45		Student 48
	Student 46		Student 49
	Student 47		Student 50
	Student 52		Student 51
	Student 53		Student 56
	Student 54		Student 57
	Student 55		Student 59
	Student 58		Student 60
Student 71	Student 61		
Student 73	Student 62		
Student 74	Student 63		
Student 76	Student 64		
Student 77	Student 65		
Student 79	Student 66		
Student 80	Student 67		
Student 82	Student 68		
Student 83	Student 69		
Student 84	Student 70		
Student 86	Student 72		
Student 87	Student 75		
Student 88	Student 78		
Student 89	Student 81		
Student 90	Student 85		

Berdasarkan Tabel 6 dapat dilihat jumlah dan daftar mahasiswa yang tergolong ke dalam *low performance* dan *high performance*.

Tabel 6. Nilai *Centroid Variable*

		Centroids																
Cluster	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16	X17	
0	0.45895	0.38348	0.35121	0.16724	0.08901	0.24105	0.66667	0.54136	0.39924	0.60631	0.09021	0.21064	0.42809	0.35707227	0.45093	0.23696	0.54103	
1	-0.4589	-0.3835	-0.3512	-0.1672	-0.089	-0.2411	-0.6667	-0.5414	-0.3992	-0.6063	-0.0902	-0.2106	-0.4281	-0.35707227	-0.4509	-0.237	-0.541	

Berdasarkan Tabel 5 dapat dilihat bahwa pada klaster 0 terdapat 2 indikator yang memiliki nilai rendah yaitu pada X5 (Pemecahan masalah kompleks di perkuliahan Data Analitik) dan X11 (Kemampuan berinovasi dalam perkuliahan Data Analitik), sedangkan pada klaster 1 terdapat 2 indikator yang memiliki nilai rendah yaitu X7 (Keaktifan Kolaborasi di perkuliahan Data Analitik) dan X10 (Mendapatkan Motivasi selama perkuliahan Data Analitik). Sehingga terdapat beberapa strategi yang dapat diterapkan oleh Dosen Pengampu berdasarkan masing-masing klaster yaitu:

- a. *High performance*: Jika dosen pengampu ingin meningkatkan performansi dari mahasiswa di klaster 0 maka dosen pengampu dapat meningkatkan kemampuan memecahkan masalah kompleks dan juga kemampuan berinovasi selama perkuliahan Data Analitik. Salah satu cara yang diterapkan adalah dengan memberikan beberapa *study case* kepada mahasiswa yang mengambil mata kuliah tersebut. Dengan di berikannya *study case* diharapkan mahasiswa dapat berlatih dalam memecahkan masalah kompleks dan juga meningkatkan kemampuan berinovasi.
- b. *Low performance*: Jika dosen pengampu ingin meningkatkan performansi dari mahasiswa di klaster 1 maka dosen pengampu dapat meningkatkan keaktifan kolaborasi dan juga meningkatkan motivasi di perkuliahan Data Analitik. Salah satu cara yang dapat di terapkan adalah dengan memberikan apresiasi lebih ketika mahasiswa bertanya dengan cara memberikan nilai tambahan setiap mahasiswa bertanya. Selain itu Dosen Pengampu dapat mengajak mahasiswa untuk berpartisipasi dalam tantangan atau kompetisi yang relevan dengan materi kuliah karena dapat memotivasi mereka untuk terlibat lebih aktif.

D. KESIMPULAN

Strategi dapat dilakukan oleh Dosen Pengampu untuk meningkatkan *Student's Performance* pada mata kuliah Data Analitik. Pengembangan *Student's Performance* akan lebih efektif jika kita mengelompokkan mahasiswa berdasarkan performansi mereka sehingga dapat menghasilkan klasterisasi performa mahasiswa. Disimpulkan terdapat 2 klaster yaitu *High Performance* dan *low performance*. Hasil klasterisasi tersebut diperoleh dengan menggunakan enam atribut yaitu *critical thinking, collaborative skill, creativity and inovation,*



technology application, participation level, dan overall achievement. Atribut tersebut diolah menggunakan algoritma *K-Means* sehingga dapat menentukan strategi untuk mengoptimalkan performansi mahasiswa yang mengambil mata kuliah Data Analitik.

Penelitian ini memiliki keterbatasan yang dapat ditingkatkan pada penelitian selanjutnya. Jumlah mahasiswa yang cukup banyak dapat menyebabkan adanya data *outlier* dan juga perlu dilakukannya sampel per kelas. Jadi, pada penelitian selanjutnya diharapkan dapat menggunakan seluruh data mahasiswa tanpa menggunakan sampel per kelas agar mendapatkan variasi klaster dan juga menemukan strategi yang lebih akurat.

E. REFERENSI

- Abbas, S. A., Aslam, A., Rehman, A. U., Abbasi, W. A., Arif, S., & Kazmi, S. Z. H. (2020). K-means and k-medoids: Cluster analysis on birth data collected in city Muzaffarabad, Kashmir. *IEEE Access*, 8, 151847-151855. DOI: 10.1109/ACCESS.2020.3014021.
- Abu Talib, M., Bettayeb, A. M., & Omer, R. I. (2021). Analytical study on the impact of technology in higher education during the age of COVID-19: Systematic literature review. *Education and information technologies*, 1-28. DOI: <https://doi.org/10.1007/s10639-021-10507-1>
- Achmad, F., Abdillah, I. T., & Amani, H. (2023). Decision-Making Process for Tourism Potential Segmentation. *International Journal of Innovation in Enterprise System*, 7(01), 19-30. DOI: <https://doi.org/10.25124/ijies.v7i01.204>
- Achmad, F., Prambudia, Y., & Rumanti, A. A. (2023). Improving Tourism Industry Performance through Support System Facilities and Stakeholders: The Role of Environmental Dynamism. *Sustainability*, 15(5), 4103. DOI: <https://doi.org/10.3390/su15054103>
- Achmad, F., Prambudia, Y., & Rumanti, A. A. (2023). Sustainable Tourism Industry Development: A Collaborative Model of Open Innovation, Stakeholders, and Support System Facilities. *IEEE Access*, 11, 83343 - 83363. DOI: 10.1109/ACCESS.2023.3301574
- Akour, I. A., Al-Marouf, R. S., Alfaisal, R., & Salloum, S. A. (2022). A conceptual framework for determining metaverse adoption in higher institutions of gulf area: An empirical study using hybrid SEM-ANN approach. *Computers and Education: Artificial Intelligence*, 3, 100052. DOI: <https://doi.org/10.1016/j.caeai.2022.100052>

- AlMahdawi, M., Senghore, S., Ambrin, H., & Belbase, S. (2021). High school students' performance indicators in distance learning in chemistry during the COVID-19 pandemic. *Education Sciences*, 11(11), 672. DOI: <https://doi.org/10.3390/educsci11110672>
- Aurachman, R. (2018). Faktor Eksternal Dalam Penentuan Strategi Dasar Program Pendidikan Bidang Rekayasa Logistik dan Rantai Suplai. *Jurnal Manajemen Industri dan Logistik*, 2(2), 116-126. DOI: 10.30988/jmil.v2i2.32
- Fa'rifah, R.Y., Achmad, F., Septiningrum, L., Wiratmadja, I.I. (2023). Segmentation of Potential Sustainable Tourism Based on Support System Facility Perspective. In: Rosyidi, C.N., Laksono, P.W., Jauhari, W.A., Hisjam, M. (eds) Proceedings of the 6th Asia Pacific Conference on Manufacturing Systems and 4th International Manufacturing Engineering Conference. iMEC-APCOMS 2022. Lecture Notes in Mechanical Engineering. Springer, Singapore. https://doi.org/10.1007/978-981-99-1245-2_26
- Fiyanda, A. D. T. (2022). *Identifikasi Mikroplastik Udara dalam Ruangan Sekolah di Jalan Arteri Kota Makassar= Identification of Airbone Microplastics in School Rooms on Arterial Roads in Makassar City* (Doctoral dissertation, Universitas Hasanuddin).
- Galuh, A. D., Maharani, D., Meynawati, L., Anggraeni, D., & Furnamasari, Y. F. (2021). Urgensi nilai dan moral dalam upaya meningkatkan pendidikan karakter melalui pembelajaran pkn di sekolah dasar. *Jurnal Basicedu*, 5(6), 5169-5178. DOI: <https://doi.org/10.31004/basicedu.v5i6.1598>
- Herlambang, Y. T. (2021). *Pedagogik: Telaah Kritis Ilmu Pendidikan Dalam Multiperspektif*. Bumi Aksara.
- Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhajja, B., & Heming, J. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences*, 622, 178-210. DOI: <https://doi.org/10.1016/j.ins.2022.11.139>
- Kampus Merdeka. (2023). Latar Belakang. <https://kampusmerdeka.kemdikbud.go.id/web/about/latar-belakang#:~:text=Program%20Merdeka%20Belajar%20%2D%20Kampus%20Merdeka,dunia%20industri%2C%20maupun%20dinamika%20masyarakat>

Septiningrum, L., Achmad, F., Rumanti, A. A., & Wiratmadja, I. I. (2022, October). Analysis of Employment Increasing for Potential Tourism Industry in Indonesia: An Exploration Method Approach. In *Proceedings of the International Manufacturing Engineering Conference & The Asia Pacific Conference on Manufacturing Systems* (pp. 205-212). Singapore: Springer Nature Singapore. DOI: https://doi.org/10.1007/978-981-99-1245-2_19

