

Identifikasi Kelompok Prestasi Mahasiswa Berdasarkan Kombinasi Nilai Tugas, MID, dan UAS Menggunakan *K-Means Clustering*

Nur Azizah Ayu Safanah[#]

Pendidikan Teknik Informatika dan Komputer, Universitas Negeri Makassar
Jl. Mallengkeri Raya, Parang Tambung, Kec. Tamalate, Kota Makassar, Sulawesi Selatan, 90224
nurazizahayusafanah14@gmail.com

Abstrak

Penelitian ini bertujuan untuk mengidentifikasi kelompok prestasi mahasiswa berdasarkan kombinasi nilai Ulangan Harian, Ujian Tengah Semester (UTS), dan Ujian Akhir Semester (UAS) menggunakan metode K-Means Clustering. Data yang digunakan adalah nilai mahasiswa dari mata kuliah Ilmu Budaya Dasar di Fakultas Ilmu Sosial dan Politik, Universitas Maritim Raja Ali Haji. Sampel terdiri dari 24 mahasiswa dengan atribut yang dianalisis meliputi Tugas Mandiri (TM), Tugas, UTS, dan UAS. Metode K-Means Clustering digunakan untuk mengelompokkan data ke dalam tiga kategori prestasi: tinggi, sedang, dan cukup. Proses clustering dilakukan melalui beberapa iterasi hingga tidak ada lagi perubahan anggota cluster. Hasil pengelompokan kemudian divalidasi menggunakan software KNIME *Analytic Platform* untuk memastikan akurasi. Penelitian ini menunjukkan bahwa metode K-Means Clustering dapat digunakan secara efektif untuk menganalisis dan mengelompokkan prestasi mahasiswa berdasarkan kombinasi nilai akademik mereka, memberikan wawasan yang bermanfaat bagi institusi pendidikan tinggi dalam memahami pola prestasi mahasiswa.

Kata kunci: K-Means Clustering, Prestasi Akademik, Analisis Akademik, Pengelompokan Data

Abstract

This research aims to identify student performance groups based on a combination of Daily Test scores, Midterm Exam, and Final Exam scores using the K-Means Clustering method. The data used are the scores of students from the Basic Cultural Studies course at the Faculty of Social and Political Sciences, Universitas Maritim Raja Ali Haji. The sample consists of 24 students with attributes analyzed including Independent Assignment (TM), Assignment, UTS, and UAS. The K-Means Clustering method is used to group the data into three performance categories: high, medium, and sufficient. The clustering process is carried out through several iterations until there are no more changes in cluster membership. The clustering results are then validated using the KNIME Analytic Platform software to ensure accuracy. This research shows that the K-Means Clustering method can be effectively used to analyze and group student performance based on their academic scores, providing valuable insights for higher education institutions in understanding student performance patterns.

Keywords: K-Means Clustering, Student Performance, Academic Analysis, Data Clustering.

I. PENDAHULUAN

Prestasi mahasiswa merupakan salah satu indikator keberhasilan sistem pendidikan tinggi dalam mencapai tujuannya, dan pendidikan tinggi memegang peranan penting dalam pertumbuhan masyarakat dan bangsa. Pendidikan tinggi sering kali menggunakan sistem penilaian Ujian Harian (UH), Ujian Tengah Semester (UTS/MID), dan Ujian Akhir Semester (UAS) [1]. Dalam konteks

khusus ini, prestasi siswa memegang peranan penting sebab prestasi dapat menjadi indikator luasnya perolehan pembelajaran yang diperoleh siswa sepanjang perjalanan akademis mereka [2]

Teknik pengelompokan K-means merupakan salah satu metode yang dapat digunakan untuk menganalisis prestasi belajar siswa. Dengan metode ini, kelompok prestasi siswa dapat diidentifikasi dengan lebih mudah, yaitu dengan mengelompokkan siswa berdasarkan gabungan nilai Ulangan Harian,

Nilai Tengah Semester (UTS), dan Nilai Akhir Semester (UAS) [3]. Hasilnya, metode ini memberikan kontribusi yang signifikan terhadap pemahaman yang lebih menyeluruh dan metodis tentang pola prestasi siswa.

Selain mencerminkan bakat individu, prestasi siswa juga dapat berfungsi sebagai tolak ukur keberhasilan strategi pengajaran dan pembelajaran yang digunakan dalam pendidikan tinggi. Oleh karena itu, lembaga pendidikan tinggi dapat memperoleh manfaat besar dari pemahaman menyeluruh tentang prestasi siswa melalui analisis menggunakan metode K-means Clustering, yang dapat membantu meningkatkan kualitas pembelajaran dan mencapai tujuan pendidikan dengan lebih sukses [2]. Selain itu, analisis prestasi mahasiswa juga memberikan informasi berharga bagi pihak eksternal, seperti calon mahasiswa, orang tua, dan pemangku kepentingan lainnya. Analisis yang sistematis menggunakan K-means Clustering memberikan informasi transparan dan akurat terkait kualitas pendidikan yang ditawarkan, meningkatkan kepercayaan masyarakat terhadap lembaga pendidikan tinggi dan membantu calon mahasiswa dalam memilih program studi yang sesuai dengan minat dan potensi mereka.

II. METODE PENELITIAN

Metode Penelitian yang dilakukan menjelaskan tahapan penelitian dan objek penelitian yang akan dijelaskan sebagai berikut.

A. Tahapan Penelitian

Untuk melakukan penelitian secara lebih metodis, cermat, dan terarah, Algoritma K-Means digunakan untuk mengelompokkan data akademik mahasiswa melalui sejumlah langkah yang lebih terstruktur. Berikut ini adalah langkah-langkah yang akan diambil untuk mengimplementasikan algoritma K-Means:

1. Pengumpulan dan penentuan data adalah tahap yang dilakukan untuk mengumpulkan data. Pada penelitian ini penulis mengambil data yang dikumpulkan dari Google Search dengan data lengkap dan valid. Data yang dikumpulkan ini akan disamakan agar tidak melanggar privasi individu.
2. Data yang dikumpulkan berupa data nilai Mahasiswa dengan beberapa atribut yang akan dijadikan sebagai penelitian berupa nilai Tugas Mandiri (TM), Tugas, Ujian Tengah Semester (UT) dan Ujian Akhir Semester (UAS).
3. Menentukan jumlah kluster (K) dari dataset yang dikumpulkan sebelumnya, diikuti dengan

penentuan centroid yang biasanya diturunkan dari baris pertama atau secara acak.

4. Perhitungan algoritma K-Means terhadap data nilai mahasiswa adalah perhitungan dengan menggunakan Microsoft Excel kemudian dilakukan validasi dengan menggunakan aplikasi KNIME Analytic Platform.

B. Objek Penelitian

Objek penelitian ini adalah data nilai mahasiswa dari Universitas Maritim Raja Ali Haji, Fakultas Ilmu Sosial dan Politik pada mata kuliah Ilmu

Budaya Dasar. Data tersebut diambil dari pencarian Google Search dan mencakup dataset sebanyak 48 mahasiswa. Atribut yang digunakan sebagai variabel dalam algoritma ini telah disederhanakan menjadi beberapa atribut utama, yaitu NIM, Nama Mahasiswa, Tugas Mandiri (TM), Tugas, Ujian Tengah Semester (UTS), dan Ujian Akhir Semester (UAS). Dari dataset tersebut, diambil sampel sebanyak 24 mahasiswa untuk diuji coba dalam penelitian ini. Data yang dikumpulkan ini akan dianalisis menggunakan teknik yang relevan untuk mendapatkan wawasan lebih mendalam mengenai prestasi akademik mahasiswa dan faktor-faktor yang mempengaruhinya. Proses penyamaran data juga dilakukan untuk memastikan tidak ada pelanggaran privasi individu.

III. HASIL DAN PEMBAHASAN

Sampel hasil ujian mahasiswa dari mata kuliah Ilmu Budaya Dasar tahun ajaran 2014–2015 di Fakultas Ilmu Sosial dan Ilmu Politik Universitas Maritim Raja Ali Haji digunakan untuk mengolah data dalam penelitian ini. Atribut yang digunakan sebagai variabel dalam algoritma ini telah direduksi menjadi beberapa atribut utama: NIM, Nama Mahasiswa, Tugas Mandiri (TM), Tugas, Ujian Tengah Semester (UTS), dan Ujian Akhir Semester (UAS). Ada dua puluh empat data mahasiswa dalam sampel yang akan diuji.

Selain itu, algoritma K-Means akan digunakan untuk membagi data menjadi tiga kelompok. Langkah-langkah penyelesaiannya adalah sebagai berikut:

1. Menghitung Jumlah Cluster: Jumlah kelompok dengan kategori prestasi siswa cukup, sedang, dan tinggi yang akan dihasilkan adalah jumlah cluster.
2. Mencari titik pusat awal: Titik pusat awal dipilih secara acak. Titik pusat cluster pertama dikenal sebagai titik pusat awal. Titik pusat awal penelitian adalah

Tabel 1. Centroid Awal

Centroid Awal	TM	Tugas	UTS	UAS
C0	95	80	80	80
C1	100	60	70	80
C2	100	75	80	70

Berikut di bawah ini merupakan tabel data mahasiswa yang akan digunakan:

Tabel 2. Data Nilai Mahasiswa Universitas Maritim Raja Ali Haji

No	Mahasiswa	TM	Tugas	UTS	UAS
1	A	95	80	80	80
2	B	100	80	100	70
3	C	100	70	60	80
4	D	100	60	100	80
5	E	100	65	100	80
6	F	100	65	60	80
7	G	95	70	60	80
8	H	100	75	75	70
9	I	100	80	80	80
10	J	100	80	80	80
11	K	100	80	85	80
12	L	100	75	100	70
13	M	100	60	70	80
14	N	95	70	80	80
15	O	100	65	100	80
16	P	100	80	80	80
17	Q	100	80	80	80
18	R	100	80	80	80
19	S	100	60	60	80
20	T	100	80	80	80
21	U	100	60	100	80
22	V	100	80	70	80
23	W	100	65	100	80
24	X	100	75	80	70

3. Menghitung jarak setiap data yang ada terhadap setiap pusat cluster. Berikut perhitungannya dengan menggunakan persamaan Euclidean Distance Space:

- a. Jarak antara data siswa pertama dengan pusat cluster pertama.

$$c_1 \sqrt{(95 - 95)^2 + (80 - 80)^2 + (80 - 80)^2 + (80 - 80)^2} = 0.00$$

- b. Jarak antara data siswa pertama dengan pusat cluster ke dua.

$$c_1 \sqrt{(95 - 100)^2 + (80 - 60)^2 + (80 - 70)^2 + (80 - 80)^2} = 22.91$$

- c. Jarak antara data siswa pertama dengan pusat cluster ke tiga.

$$c_1 \sqrt{(95 - 100)^2 + (80 - 75)^2 + (80 - 80)^2 + (80 - 70)^2} = 12.25$$

Adapun hasil perhitungan dari keseluruhan data terhadap tiap pusat cluster awal disajikan pada table 3

Tabel 3. Hasil Perhitungan Jarak Data Siswa pada Masing-Masing Centroid

	C0	C1	C2
0.00	22.91	12.25	
22.91	37.42	20.62	
22.91	14.14	22.91	
28.72	30.00	26.93	
25.50	30.41	24.49	
25.50	11.18	24.49	
22.36	15.00	23.45	
13.23	18.71	5.00	
5.00	22.36	11.18	
5.00	22.36	11.18	
7.07	25.00	12.25	
23.45	35.00	20.00	
22.91	0.00	20.62	
10.00	15.00	12.25	
25.50	30.41	24.49	
5.00	22.36	11.18	
5.00	22.36	11.18	
5.00	22.36	11.18	
28.72	10.00	26.93	
5.00	22.36	11.18	
28.72	30.00	26.93	
11.18	20.00	15.00	
25.50	30.41	24.49	
12.25	20.62	0.00	

4. Pada table 3 menunjukkan pengelompokan cluster setelah perhitungan jarak data pada centroid. Pada penentuan ini dilakukan dengan cara melihat nilai yang paling kecil atau nilai minimum setelah dilakukan perhitungan jarak. Berdasarkan hasil yang diperoleh dari iterasi 1 maka diperoleh pengelompokan data anggota C0, C1, dan C2 seperti berikut:

- a. Anggota Cluster 0 terdiri dari 10 data Mahasiswa
 - b. Anggota Cluster 1 terdiri dari 5 data Mahasiswa
 - c. Anggota Cluster 2 terdiri dari 9 data mahasiswa
5. Setelah semua data ditempatkan ke dalam cluster yang terdekat, hitung Kembali pusat cluster yang baru berdasarkan rata-rata anggota ada pada cluster tersebut.

Tabel 4. Centroid Baru

Centroid Awal	TM	Tugas	UTS	UAS
C0	99	79	80	80
C1	99	65	62	80
C2	100	69	95	76

Hitung ulang data menggunakan pusat kluster baru setelah diperoleh titik pusat baru dari setiap kluster untuk memastikan bahwa pola akhir tetap tidak berubah. Data dalam penelitian ini dilakukan hingga iterasi ketiga, di mana tidak ada lagi perpindahan data antar kluster dan tidak terjadi perubahan pada data dalam setiap kluster. Hasil pengelompokan berikut dihasilkan berdasarkan perhitungan data yang dilakukan pada iterasi 3.

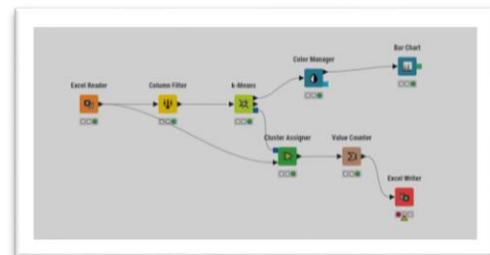
Tabel 5. Hasil dan Pola Terakhir Klusterisasi Nilai Cluster pada Iterasi 3

	C0	C1	C2
5.00	23.77	24.56	
22.58	42.07	14.76	
20.74	5.48	40.22	
27.75	38.34	7.62	
24.80	38.01	3.61	
23.13	2.24	40.16	
21.10	6.71	40.53	
9.49	19.24	27.17	
3.16	23.45	24.04	
3.16	23.45	24.04	
6.71	27.48	20.07	
22.69	40.56	10.63	
20.25	9.49	30.95	
9.22	19.10	21.05	
24.80	38.01	3.61	
3.16	23.45	24.04	
3.16	23.45	24.04	
3.16	23.45	24.04	
26.27	5.48	40.72	

3.16	23.45	24.04
27.75	38.34	7.62
9.49	17.03	32.83
24.80	38.01	3.61
8.66	22.91	22.65

Terdapat 12 mahasiswa tergabung dalam kluster 0, 5 mahasiswa tergabung dalam kluster 1, dan 7 mahasiswa tergabung dalam kluster 2. Penelitian mencapai konvergensi setelah jumlah iterasi tertentu, di mana tidak terjadi perubahan pada pengelompokan mahasiswa antar iterasi.

6. Pengujian dengan aplikasi yang telah ditentukan. Penuliss melakukan pengujian menggunakan program yang telah ditentukan untuk menilai akurasi hasil pengelompokan yang diperoleh. Setelah hasil perhitungan awal dari Microsoft Excel diperoleh, perhitungan diulangi menggunakan Perangkat Lunak Platform Analisis KNIME. Tujuannya adalah untuk memverifikasi akurasi hasil pengelompokan dan memastikan kesesuaiannya dengan harapan. Rangkaian simpul dalam perangkat lunak KNIME, yang menggunakan metode K-means Clustering untuk memproses data, diilustrasikan sebagai berikut.



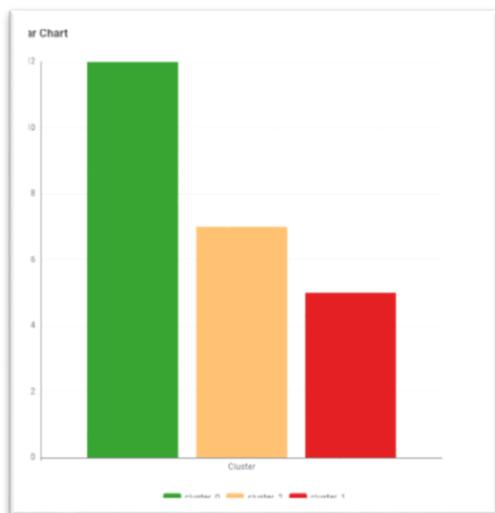
Gambar 1. Rangkaian Node menggunakan Software KNIME Analytic Platform

Setelah proses pengelompokan data selesai, frekuensi masing-masing cluster dapat diamati pada node Value Counter. Node ini menampilkan jumlah data yang tergolong dalam setiap cluster. Informasi ini membantu para peneliti dalam memahami distribusi data dan menilai efektivitas metode K-means Clustering dalam mengelompokkan mahasiswa berdasarkan prestasi mereka.

#	RowID	count Number (integer)
1	cluster_0	12
2	cluster_1	5
3	cluster_2	7

Gambar 2. Tampilan frekuensi dari Node Value Counter

Demi mempermudah interpretasi hasil pengolahan data, penulis memanfaatkan node Bar Chart dalam perangkat lunak KNIME untuk memvisualisasikan data yang telah diperoleh. Visualisasi ini disajikan dalam bentuk batang, sehingga memungkinkan pembaca untuk dengan mudah memahami pola dan tren yang terkandung dalam data.



Gambar 3. Visualisasi Hasil Perolehan Data

IV. KESIMPULAN

Penelitian ini bertujuan untuk mengategorikan prestasi siswa ke dalam tiga kelompok, yaitu tinggi, sedang, dan cukup, menggunakan metode algoritma K-Means Clustering. Algoritma ini memanfaatkan nilai Ujian Tengah Semester (UTS), Ujian Akhir Semester (UAS), Tugas Mandiri (TM), dan Tugas sebagai data dasar untuk mengelompokkan siswa. Proses pengelompokan data ini dilakukan dengan perangkat lunak KNIME, menghasilkan output berupa pengelompokan data prestasi siswa.

Penggunaan KNIME dalam penelitian ini meningkatkan keakuratan data prestasi siswa sehingga menghasilkan pengelompokan yang lebih tepat dan informatif.

REFERENSI

- [1] J. Baxter, G. Aranda-Mena, and K. Robson, "Chasing student satisfaction in the delivery of property higher education," in *24th Annual European Real Estate Society Conference*, Delft, Netherlands: European Real Estate Society, 2017. doi: 10.15396/eres2017_84.
- [2] I. Fialho, M. Cid, and M. Coppi, "Pedagogical Assessment in Higher Education: The Importance of Training," *Educ. Sci.*, vol. 13, no. 12, p. 1248, Dec. 2023, doi: 10.3390/educsci13121248.
- [3] A. A. Widiyanto *et al.*, "Practising Eco-Theology: Pesantren and Green Education in Narmada Lombok, Nusa Tenggara Barat (NTB), Indonesia," in *Proceedings of the 1st Lawang Sewu International Symposium on Humanities and Social Sciences 2022 (LEWIS 2022)*, vol. 758, D. Mulyadi, S. Aimah, and E. Winaryati, Eds., in *Advances in Social Science, Education and Humanities Research*, vol. 758. , Paris: Atlantis Press SARL, 2023, pp. 118–125. doi: 10.2991/978-2-38476-078-7_14.
- [4] Qorik IM, Eko SB, "Implementasi Algoritma K-Means Clustering Dalam Penilaian Kedisiplinan Siswa," in *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 3, 2022
- [5] Reynaldi SG, Erika R, "Implementasi Bcrypt dengan SHA-256 pada Password Pengguna Aplikasi Golek Kost," in *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 3, 2022
- [6] Gustientiedina dkk, "Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan," in *Jurnal Teknologi dan Sistem Informasi (TEKNOSI)*, vol. 5 No.1, Universitas Andalas, 2019
- [7] Tifani dkk, "Metode K-Means Clustering Dalam Pengelompokan Penjualan Frozen Food," in *Jurnal Ilmiah Wahan Pendidikan*, vol. 8, 2022

