

PENGELOMPOKKAN DATA SISWA DI SEKOLAH MENENGAH KEJURUAN MENGUNAKAN ALGORITMA K-MEANS

Ayu Mutia Br Sbr¹, Akim Manaor Hara Pardede², Juliana Naftali Sitompul³
Sistem Informasi, STMIK Kaputama Binjai

E-mail: *ayumutya258@gmail.com¹, akimmhp@live.com², juliananaftali4424@gmail.com³

ABSTRAK

Sekolah Menengah Kejuruan (SMK) memainkan peran penting dalam menyiapkan lulusan siap kerja yang sesuai dengan kebutuhan industri. Namun, pemilihan jurusan oleh siswa seringkali hanya didasarkan pada keinginan pribadi, tanpa mempertimbangkan faktor-faktor lain seperti asal daerah dan kecenderungan minat. Penelitian ini bertujuan untuk menganalisis pola peminatan siswa berdasarkan asal kecamatan, jenis kelamin, dan jurusan menggunakan metode data mining, khususnya algoritma K-Means Clustering. Data yang digunakan adalah data siswa kelas XI SMK Esa Prakarsa tahun ajaran 2023. Proses pengelompokan dilakukan menggunakan perangkat lunak MATLAB, melalui tahapan transformasi data, inisialisasi centroid, perhitungan jarak dengan metode Euclidean Distance, dan iterasi hingga tercapai konvergensi. Hasil penelitian menunjukkan terbentuknya tiga kelompok utama peminatan siswa dengan karakteristik yang berbeda sesuai dengan asal kecamatan. Informasi ini memberikan gambaran mengenai kecenderungan jurusan yang paling diminati oleh siswa dari wilayah tertentu, sehingga dapat dimanfaatkan oleh pihak sekolah untuk merencanakan kapasitas jurusan, mengalokasikan sumber daya, dan menyusun strategi promosi yang lebih tepat sasaran. Penelitian ini juga membuktikan bahwa algoritma K-Means efektif digunakan dalam menganalisis pola peminatan siswa, sehingga dapat mendukung pengambilan keputusan yang lebih akurat dalam pengelolaan pendidikan kejuruan.

Kata kunci

Data Mining, K-Means Clustering, Siswa, Jurusan, Kecamatan

ABSTRACT

Vocational High Schools (SMK) play an important role in preparing work-ready graduates to meet industry needs. However, students' choice of major is often based solely on personal preferences, without considering other factors such as their region of origin and interests. This study aims to analyze students' major preferences based on district of origin, gender, and major using data mining methods, specifically the K-Means Clustering algorithm. The data used consists of class XI students of SMK Esa Prakarsa in the 2023 academic year. The clustering process was carried out using MATLAB software, through stages of data transformation, centroid initialization, distance calculation using the Euclidean Distance method, and iteration until convergence was reached. The results show that three main groups of student preferences were formed, each with distinct characteristics according to their district of origin. This information provides insights into the majors most preferred by students from specific areas, which can be used by schools to plan program capacities, allocate resources, and develop more targeted promotion strategies. This study also demonstrates that the K-Means algorithm is effective in analyzing students' preference patterns, thereby supporting more accurate decision-making in the management of vocational education.

Keywords

Data Mining, K-Means Clustering, Students, Majors, District

1. PENDAHULUAN

Pendidikan kejuruan merupakan jenjang pendidikan menengah yang strategis dalam menyiapkan lulusan yang memiliki kompetensi sesuai dengan kebutuhan industri. Salah satu tantangan yang dihadapi oleh Sekolah Menengah Kejuruan (SMK) adalah dalam mengelola peminatan jurusan siswa yang seringkali bervariasi. Peminatan jurusan ini dipengaruhi oleh berbagai faktor, salah satunya adalah asal kecamatan siswa, yang dapat menunjukkan pola kecenderungan minat terhadap program keahlian tertentu (Faisal, 2022). Analisis terhadap pola peminatan ini sangat penting sebagai dasar bagi pihak sekolah dalam merencanakan program keahlian yang relevan dan sesuai dengan kebutuhan masyarakat di berbagai wilayah.

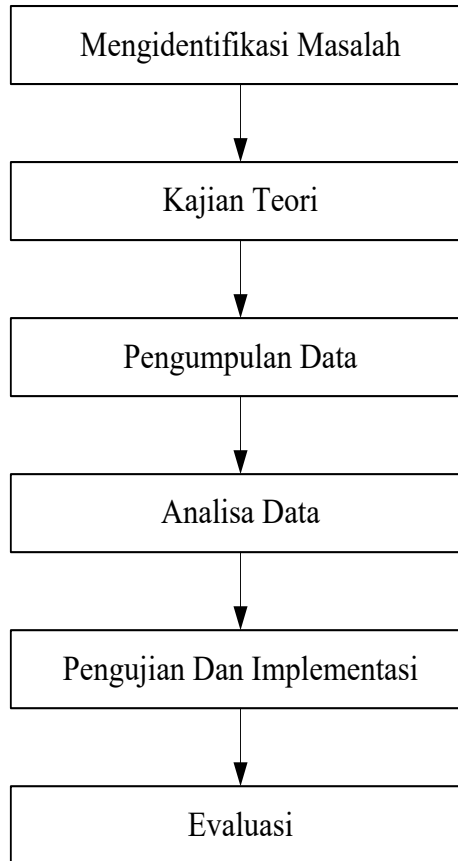
Data mining merupakan teknik yang dapat digunakan untuk menganalisis data dalam jumlah besar dan menemukan pola tersembunyi di dalamnya (Delima, 2023). Salah satu metode yang paling umum digunakan dalam data mining untuk tujuan pengelompokan data adalah clustering. Dengan menerapkan teknik clustering, khususnya dengan algoritma K-Means, pihak sekolah dapat mengidentifikasi jurusan-jurusan yang paling diminati oleh siswa dari kecamatan tertentu. Informasi ini menjadi acuan berharga dalam merencanakan kapasitas jurusan, mengalokasikan sumber daya, serta mengembangkan strategi promosi yang lebih tepat sasaran.

Berdasarkan penelitian terdahulu, penerapan algoritma K-Means telah terbukti efektif dalam berbagai kasus yang berkaitan dengan pendidikan kejuruan. Penelitian oleh Harianja & Fahmi (2022) dan Purnamasari et al. (2023) menunjukkan efektivitas algoritma ini dalam mengelompokkan data siswa untuk evaluasi hasil belajar dan penentuan jurusan. Syahra (2018) juga membuktikan bahwa K-Means dapat mempercepat dan meningkatkan akurasi dalam penentuan jurusan siswa.

Berangkat dari urgensi tersebut, penelitian ini bertujuan untuk mengidentifikasi pola pengelompokan peminatan jurusan di SMK Esa Prakarsa berdasarkan data siswa baru tahun ajaran 2023 dengan memanfaatkan algoritma K-Means Clustering.

2. METODE PENELITIAN

Metodologi penelitian adalah proses atau cara ilmiah untuk mendapatkan data yang akan digunakan untuk keperluan penelitian. Dalam melakukan penelitian pada skripsi ini, penulis mengikuti tahapan metodologi dalam penelitian ini yaitu sebagai berikut:



Gambar 1. Metode Penelitian

- a. Menentukan jumlah *cluster* (k).
- b. Menentukan nilai pusat (*Centroid*) awal secara acak.
- c. Menghitung jarak antara setiap data dengan centroid terdekat menggunakan rumus Euclidean Distance: $d\text{-}euclidean(X,Y) = \sqrt{(x_i - s_i)^2 + (y_i + t_i)^2}$
- d. Mengelompokkan objek berdasarkan jarak terdekatnya ke *centroid*.
- e. Melakukan iterasi hingga *centroid* mencapai posisi optimal (Aprilianda et al., 2023).

3. HASIL DAN PEMBAHASAN

3.1 Data Input

Dalam penelitian ini data yang digunakan adalah data siswa yang diperoleh dari SMK Esa Prakarsa. Rincian data yang digunakan dapat dijelaskan sebagai berikut :

- a. Data Masukan
 - 1) Nama file: data bab 4
 - 2) Jumlah data: 207 data
 - 3) Variabel penelitian
 - X = Kecamatan
 - Y = Jenis Kelamin
 - Z = Jurusan
- b. Cluster pengelompokkan : 3 Cluster
- c. Untuk mempermudah pengolahan, setiap variabel diberikan kode numerik sebagai berikut:

Tabel 1. Inisialisasi Kriteria Kecamatan

No	Kecamatan	Nilai Transformasi
1	Selesai	1
2	Kuala	2
3	Bahorok	3
4	Salapian	4
5	Hinai	5
6	Stabat	6
7	Kutambaru	7
8	Wampu	8
9	Sei Lapan	9
10	Gebang	10
11	Babalan	11
12	Padang Tualang	12
13	Sawit Sebrang	13
14	Tanjung Pura	14
15	Batang Serangan	15
16	Secanggang	16
17	Sibiru-biru	17
18	Sei Bingai	18
19	Serapit	19
20	Brandan Barat	20
21	Pangkalan Susu	21
22	Besitang	22
23	Pangkalan Jaya	23

Tabel 2. Inisialisasi Kriteria Pidana

No	Jenis Kelamin	Nilai Transformasi
1	Laki-laki	1
2	Perempuan	2

Tabel 3. Inisialisasi Kriteria Pendidikan

No	Jurusan	Nilai Transformasi
1	AK	1
2	AP	2
3	RPL	3
4	TB	4
5	TKJ	5
6	TKR	6
7	TBSM	7

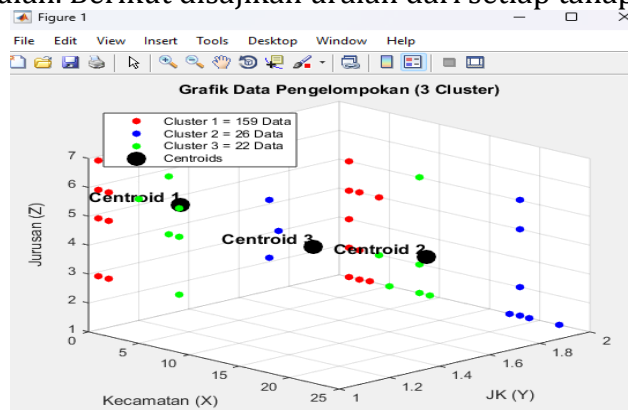
3.2 Implementasi

Implementasi merupakan tahap penerapan sekaligus pengujian bagi sistem baru serta merupakan tahap dimana aplikasi akan dioperasikan pada keadaan yang sebenarnya. Implementasi pada tahap ini merupakan proses tahap akhir dari penerapan metode Clustering pada sistem pengelompokan data sesuai dengan perancangan yang telah dilakukan. Implementasi dilakukan pada aplikasi pemrograman MATLAB dengan menerapkan metode Clustering pada coding program sehingga sistem dapat melakukan proses pengelompokan data sesuai dengan tahap-tahap pengelompokan data. Dari hasil perancangan yang telah dilakukan, proses berhasil menerapkan metode Clustering dengan algoritma K-Means.

3.3 Hasil

a. Hasil Pengujian 3 Cluster

Hasil pengujian tersebut menjadi dasar analisis yang akan digunakan dalam penyusunan kesimpulan. Berikut disajikan uraian dari setiap tahap pengujian.



Gambar 2. Hasil Pengujian 3 Cluster

C1 : 1,25 1,32 4,81

Dari gambar diatas dapat dianalisis pada hasil cluster pertama dapat diketahui hasil dari grup Kecamatan (X), Jenis Kelamin (Y), Jurusan (Z) jumlah data yang dimiliki sebanyak 159 data. Dari keterangan data menunjukkan bahwa kecamatan yang banyak dengan titik pusat centroid pada variable Kecamatan selesai adalah 1,25 dengan variabel Jenis Kelamin laki-laki dengan nilai 1,32 dan variabel Jurusannya adalah 4,81 yaitu TKJ.

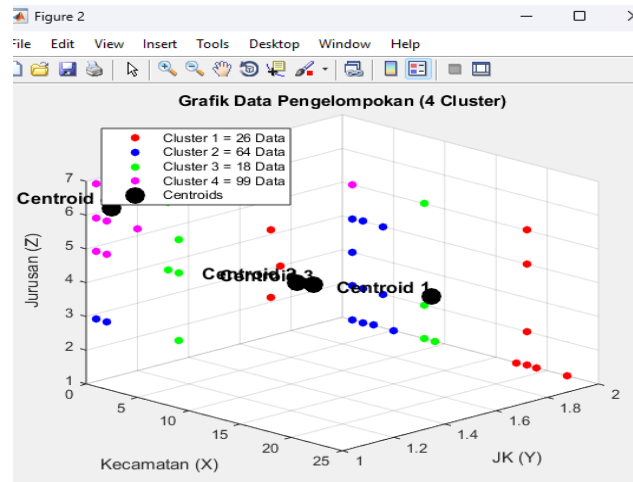
C2 : 18,26 1,61 3,80

Dari gambar diatas dapat dianalisis pada hasil cluster kedua dapat diketahui hasil dari grup Kecamatan (X), Jenis Kelamin (Y), Jurusan (Z) jumlah data yang dimiliki sebanyak 26 data. Dari keterangan data menunjukkan bahwa kecamatan yang banyak dengan titik pusat centroid pada variable Kecamatan Sei Bingai adalah 18,26 dengan variabel Jenis Kelamin Perempuan dengan nilai 1,61 dan variabel Jurusannya adalah 3,80 yaitu TB.

C3 : 8,15 1,57 3,36

Dari gambar diatas dapat dianalisis pada hasil cluster ketiga dapat diketahui hasil dari grup Kecamatan (X), Jenis Kelamin (Y), Jurusan (Z) jumlah data yang dimiliki sebanyak 22 data. Dari keterangan data menunjukkan bahwa kecamatan yang banyak dengan titik pusat centroid pada variable Kecamatan Wampu adalah 8,15 dengan variabel Jenis Kelamin Perempuan dengan nilai 1,57 dan variabel Jurusannya adalah 3,36 yaitu RPL.

b. Hasil Pengujian 4 Cluster



Gambar 3. Hasil Pengujian 4 Cluster

- 1) C1 : 18,26 1,61 3,80

Dari gambar diatas dapat dianalisis pada hasil cluster pertama dapat diketahui hasil dari grup Kecamatan (X), Jenis Kelamin (Y), Jurusan (Z) jumlah data yang dimiliki sebanyak 26 data. Dari keterangan data menunjukkan bahwa kecamatan yang banyak dengan titik pusat centroid pada variable Kecamatan Sei Bingai adalah 18,26 dengan variabel Jenis Kelamin Perempuan dengan nilai 1,61 dan variabel Jurusanya adalah 3,80 yaitu TB.

- 2) C2 : 8,33 1,55 3,5

Dari gambar diatas dapat dianalisis pada hasil cluster kedua dapat diketahui hasil dari grup Kecamatan (X), Jenis Kelamin (Y), Jurusan (Z) jumlah data yang dimiliki sebanyak 64 data. Dari keterangan data menunjukkan bahwa kecamatan yang banyak dengan titik pusat centroid pada variable Kecamatan Wampu adalah 8,33 dengan variabel Jenis Kelamin Perempuan dengan nilai 1,55 dan variabel Jurusanya adalah 3,5 yaitu TB.

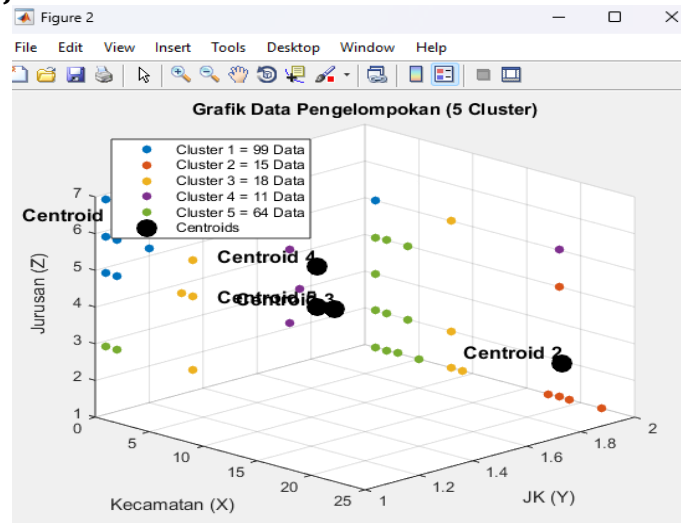
- 3) C3 : 1,39 1,76 2,59

Dari gambar diatas dapat dianalisis pada hasil cluster ketiga dapat diketahui hasil dari grup Kecamatan (X), Jenis Kelamin (Y), Jurusan (Z) jumlah data yang dimiliki sebanyak 18 data. Dari keterangan data menunjukkan bahwa kecamatan yang banyak dengan titik pusat centroid pada variable Kecamatan Selesai adalah 1,39 dengan variabel Jenis Kelamin Perempuan dengan nilai 1,76 dan variabel Jurusanya adalah 2,59 yaitu RPL.

- 4) C4 : 1,20 1,05 6,21

Dari gambar diatas dapat dianalisis pada hasil cluster keempat dapat diketahui hasil dari grup Kecamatan (X), Jenis Kelamin (Y), Jurusan (Z) jumlah data yang dimiliki sebanyak 99 data. Dari keterangan data menunjukkan bahwa kecamatan yang banyak dengan titik pusat centroid pada variable Kecamatan Selesai adalah 1,20 dengan variabel Jenis Kelamin Laki-laki dengan nilai 1,05 dan variabel Jurusanya adalah 6,21 yaitu TKR.

c. Hasil Pengujian 5 Cluster



Gambar 4. Hasil Pengujian 5 Cluster

- 1) C1 : 1,20 1,05 6,21

Dari gambar diatas dapat dianalisis pada hasil cluster pertama dapat diketahui hasil dari grup Kecamatan (X), Jenis Kelamin (Y), Jurusan (Z) jumlah data yang dimiliki sebanyak 99 data. Dari keterangan data menunjukkan bahwa kecamatan yang banyak dengan titik pusat centroid pada variable Kecamatan Selesai adalah 1,20 dengan variabel Jenis Kelamin Laki-laki dengan nilai 1,05 dan variabel Jurusanya adalah 6,21 yaitu TKR.

- 2) C2 : 18,26 2 1,93

Dari gambar diatas dapat dianalisis pada hasil cluster kedua dapat diketahui hasil dari grup Kecamatan (X), Jenis Kelamin (Y), Jurusan (Z) jumlah data yang dimiliki sebanyak 15 data. Dari keterangan data menunjukkan bahwa kecamatan yang banyak dengan titik pusat centroid pada variable Kecamatan Sei Bingai adalah 18,26 dengan variabel Jenis Kelamin Perempuan dengan nilai 2 dan variabel Jurusanya adalah 1,93 yaitu AP.

- 3) C3 : 8,33 1,55 3,5

Dari gambar diatas dapat dianalisis pada hasil cluster ketiga dapat diketahui hasil dari grup Kecamatan (X), Jenis Kelamin (Y), Jurusan (Z) jumlah data yang dimiliki sebanyak 18 data. Dari keterangan data menunjukkan bahwa kecamatan yang banyak dengan titik pusat centroid pada variable Kecamatan Wampu adalah 8,33 dengan variabel Jenis Kelamin Perempuan dengan nilai 1,55 dan variabel Jurusanya adalah 3,5 yaitu TB.

- 4) C4 : 18,27 1,09 6,36

Dari gambar diatas dapat dianalisis pada hasil cluster keempat dapat diketahui hasil dari grup Kecamatan (X), Jenis Kelamin (Y), Jurusan (Z) jumlah data yang dimiliki sebanyak 11 data. Dari keterangan data menunjukkan bahwa kecamatan yang banyak dengan titik pusat centroid pada variable Kecamatan Sei Bingai adalah 18,27 dengan variabel Jenis Kelamin Laki-laki dengan nilai 1,09 dan variabel Jurusanya adalah 6,36 yaitu TKR.

- 5) C5 : 18,27 1,09 6,36

Dari gambar diatas dapat dianalisis pada hasil cluster keempat dapat diketahui hasil dari grup Kecamatan (X), Jenis Kelamin (Y), Jurusan (Z) jumlah data yang dimiliki sebanyak 64 data. Dari keterangan data menunjukkan bahwa kecamatan yang banyak dengan titik pusat centroid pada variable Kecamatan Selesai adalah 1,39 dengan variabel

Jenis Kelamin Perempuan dengan nilai 1,76 dan variabel Jurusan adalah 2,59 yaitu RPL.

3.4 Evaluasi Hasil

Untuk menentukan jumlah Cluster yang paling optimal, dilakukan evaluasi dengan mengukur nilai varians (Variance), varians minimum (Vmin), varians maksimum (Vmax), dan Cluster Variance.

a. Pengujian Hasil 3 Cluster

- 1) Cluster 1 (1,25; 1,32; 4,81) Dengan 159 Data

$$\text{Variance } C1 = 1/(159) (3,890 + 3,890 + 3,890 + \dots + 2,223) = 0,0262$$

$$V_{\min} = \min (3,8819 + 3,8819 + 3,8819 + \dots + 2,2239) = 4,2493$$

$$V_{\max} = \max (3,8819 + 3,8819 + 3,8819 + \dots + 2,2239) = 0,4531$$
- 2) Cluster 2 (18,26; 1,61; 3,80) Dengan 26 Data

$$\text{Variance } C1 = 1/26 (3,105 + 2,846 + 2,926 + \dots + 3,270) = 1,3014$$

$$V_{\min} = \min (3,105 + 2,846 + 2,926 + \dots + 3,262) = 16,2163$$

$$V_{\max} = \max ((3,105 + 2,846 + 2,926 + \dots + 3,262) = 21,1054$$
- 3) Cluster 3 (8,15; 1,57; 3,36) Dengan 22 Data

$$\text{Variance } C1 = 1/22 (2,548 + 3,970 + 3,7998 + \dots + 2,410) = 0,3937$$

$$V_{\min} = \min (2,548 + 3,969 + 3,7998 + \dots + 2,410) = 5,3893$$

$$V_{\max} = \max (2,548 + 3,969 + 3,7998 + \dots + 2,410) = 8,6614$$

b. Pengujian Hasil 4 Cluster

- 1) Cluster 1 (18,26; 1,61; 3,80) Dengan 26 Data

$$\text{Variance } C1 = 1/(26) (3,105 + 2,846 + 2,926 + \dots + 3,270) = 0,1801$$

$$V_{\min} = \min (3,105 + 2,846 + 2,926 + \dots + 3,262) = 0,5073$$

$$V_{\max} = \max (3,105 + 2,846 + 2,926 + \dots + 3,262) = 4,6850$$
- 2) Cluster 2 (8,33; 1,51; 3,5) Dengan 64 Data

$$\text{Variance } C1 = 1/64 (9,692 + 10,653 + 2,926 + \dots + 2,536) = 0,1683$$

$$V_{\min} = \min (9,692 + 10,653 + 10,434 + \dots + 2,536) = 2,415$$

$$V_{\max} = \max ((9,692 + 10,653 + 10,434 + \dots + 2,536) = 10,7715$$
- 3) Cluster 3 (1,39; 1,76; 2,59) Dengan 18 Data

$$\text{Variance } C1 = 1/18 (17,500 + 17,500 + 17,500 + \dots + 17,367) = 17,3678$$

$$V_{\min} = \min (17,500 + 17,500 + 17,500 + \dots + 17,367) = 13,5684$$

$$V_{\max} = \max (17,500 + 17,500 + 17,500 + \dots + 17,367) = 17,5002$$
4. Cluster 4 (1,20; 1,05; 6,21) Dengan 99 Data

$$\text{Variance } C1 = 1/99 (17,314 + 16,324 + 17,321 + \dots + 17,580) = 0,1775$$

$$V_{\min} = \min (17,314 + 16,324 + 17,321 + \dots + 17,572) = 13,4631$$

$$V_{\max} = \max (17,314 + 16,324 + 17,321 + \dots + 17,572) = 17,5725$$

c. Pengujian Hasil 5 Cluster

- 1) Cluster 1 (3,78; 1,01; 3,20) Dengan 153 Data

$$\text{Variance } C1 = \frac{1}{153} (1,217 + 0,297 + 0,834 + \dots + 0,834) = 0,0176$$

$$V_{\min} = \min (1,217 + 0,297 + 0,834 + \dots + 0,834) = 0,2966$$

$$V_{\max} = \max (1,217 + 0,297 + 0,834 + \dots + 0,834) = 2,6928$$
- 2) Cluster 2 (3,29; 4,54; 2,62) Dengan 116 Data

$$\text{Variance } C1 = \frac{1}{116} (2,673 + 3,001 + 3,714 + \dots + 4,168) = 0,0462$$

$$V_{\min} = \min (2,673 + 3,001 + 3,714 + \dots + 4,168) = 2,1427$$

$$V_{\max} = \max (2,673 + 3,001 + 3,714 + \dots + 4,168) = 5,3550$$
- 3) Cluster 3 (1,90; 10,10; 3,40) Dengan 10 Data

$$\text{Variance } C1 = \frac{1}{10} (7,171 + 7,562 + 7,562 + \dots + 48,030) = 1,3173$$

$$V_{\min} = \min (7,171 + 7,562 + 7,562 + \dots + 48,030) = 7,1709$$

- $V_{max} = \max (7,171 + 7,562 + 7,562 + + 48,030) = 13,1732$
- 4) *Cluster 3* (2,25; 1,13; 3,90) Dengan 138 Data
 $Variance\ C1 = \frac{1}{138} (2,785 + 4,313 + 3,508 + + 1,119) = 0,0349$
 $V_{min} = \min (2,785 + 4,313 + 3,508 + + 1,119) = 1,1186$
 $V_{max} = \max (2,785 + 4,313 + 3,508 + + 1,119) = 4,8125$
- 5) *Cluster 3* (2,58; 1,05; 1,71) Dengan 83 Data
 $Variance\ C1 = \frac{1}{83} (2,143 + 2,207 + 3,024 + + 1,427) = 0,0436$
 $V_{min} = \min (2,143 + 2,207 + 3,024 + + 1,427) = 1,4268$
 $V_{max} = \max (2,143 + 2,207 + 3,024 + + 1,427) = 3,6186$

Tabel 4. Pengujian Hasil Cluster

<i>Cluster</i>	<i>Centroid</i>	<i>Variance</i>	<i>V_{min}</i>	<i>V_{max}</i>	<i>Cluster Variance</i>
3	1,25; 1,32; 4,81	0,0262	4,2493	0,4531	10,943
	18,26; 1,61; 3,80	1,3014	16,2163	21,1054	
	8,15; 1,57; 3,36	0,3937	5,3893	8,6614	
4	18,26; 1,61; 3,80	0,1801	0,5073	4,6850	7,300
	8,33; 1,51; 3,5	0,1683	2,4150	10,7715	
	1,39; 1,76; 2,59	17,3678	13,5684	17,5002	
	1,20; 1,05; 6,21	0,1775	13,4631	17,5725	
5	3,78; 1,01; 3,20	0,0176	0,2966	2,6928	8,536
	3,29; 4,54; 2,62	0,0462	2,1427	5,355	
	1,90; 10,10; 3,40	1,3173	7,1709	13,1732	
	2,25; 1,13; 3,90	0,0349	1,1186	4,8125	
	2,58; 1,05; 1,71	0,0436	1,4268	3,6186	

4. KESIMPULAN

Penelitian ini menggunakan metode Data Mining dengan algoritma K-Means Clustering untuk mengelompokkan data siswa di Sekolah SMK Esa Prakarsa. Pengelompokan ini bertujuan untuk mengidentifikasi pola dari Pengelompokan Berdasarkan Jurusan Siswa, khususnya variabel Kecamatan, Jenis Kelamin Dan Jurusan.

5. DAFTAR PUSTAKA

- Faisal, M., Nurdin, N., Fajriana, F., & Fitri, Z. (2022). Clustering Kompetensi Teknologi Informasi dan Komunikasi Bagi Siswa SMK Menggunakan Algoritma Clustering K-Means. *Jurnal Internasional Teknik, Sains dan Teknologi Informasi* [https://ijesty.org/index.php/ijesty/artikel/v:2\(3\), 1-7](https://ijesty.org/index.php/ijesty/artikel/v:2(3), 1-7).
- Aprinia, P. R. M., & Sutanta, H. (2025). Survei dan Pembuatan Sistem Informasi Geografis Alamat Berkode Lokasi (Geocoded Address) untuk Wilayah Kalurahan Mantrijeron, Kota Yogyakarta. *JGISE: Journal of Geospatial Information Science and Engineering*, 7(2), 208. <https://doi.org/10.22146/jgise.102071>

- Darma, Y. A., & Astuti, S. (2021). *Pemahaman Konsep Literasi Gender* (B. Riswandi, Ed.; 1st ed.). Langgam Pustaka.
- Delima, S. (2023). *Algoritma Apriori* (E. Indra, Ed.; 1st ed., Vol. 1). Universitas Prima Indonesia.
- Faizah, N. (2023). Pengelolaan Siswa Pada Sekolah Agama Islam. *Manajemen Pendidikan Islam*, 1(1), 461–474. <https://doi.org/10.30868/im.v4i02.4612>
- Fatwa, M., Ristu, R., Pandiangan, S., & Supriyadi, E. (2022). Pengaplikasian Matlab Pada Perhitungan Matriks. *Papanda Journal of Mathematics and Sciences Research*, 1(2), 2022. <https://ejournal.papanda.org/index.php/pjmsr/article/view/260>
- Harianja, T., & Fahmi, H. (2022). Implementasi Metode K-Means Clustering Dalam Mengklasterkan Buku Perpustakaan Di Smk Negeri 1 Pantai Labu. *Jurnal Sistem Informasi Kaputama (JSIK)*, 6(1), 39–47.
- Hendrastuty, N. (2024). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa. *Jurnal Ilmiah Informatika Dan Ilmu Komputer (JIMA-ILKOM)*, 3(1), 46–56. <https://doi.org/10.58602/jima-ilkom.v3i1.26>
- Ma'ady, M. N. P., Lidiawaty, B. R., Dzulkarnain, A., & Ramadan, A. (2024). *Data Mining Algoritma Dan Contoh Perhitungan Matematis* (1st ed., Vol. 1). Cv. Budi Utama. https://www.google.co.id/books/edition/Data_Mining_Algoritma_Dan_Contoh_Perhitu/xsQ0EQAAQBAJ?hl=id&gbpv=0
- Pitrawati, & Sanjaya, A. (2021). Rekayasa Perangkat Lunak Perhitungan Harga Pokok Produksi Metode Full Costing Pada Umkm Mitra Cake Di Bandar Lampung. *Jurnal Informasi Dan Komputer*, 11(2), 154–162.
- Rahman, A., Restuono, L., Maulidya, A., Siregar, A., Khairul, & Wijaya, R. F. (2023). Pengelompokan Mahasiswa Potensial Drop Out Menggunakan Metode Clustering (Studi Kasus : Stmik Kaputama). *Jurnal Informatika Kaputama (JIK)*, 7(1), 39–46.
- Syahabuddin, A., Agustang, A., & Idkhan, A. M. (2021). Fungsi Leadership dalam Peningkatan Pelayanan Publik di Kantor Kecamatan Makassar. *Jurnal Governance And Politics (JGP)*, 1(2), 118–126.
- Ndruru, E., & Limbong, R. (2018). *Implementasi Data Mining Dalam Pengelompokan Jurusan yang Diminati Siswa SMK Negeri 1 Lolowua menggunakan Metode Clustering. MEANS (Media Informasi Analisa dan Sistem)*, 3(2), 107–112. e-ISSN: 2599-3089, p-ISSN: 2548-6985.
- Syakra, Yohanni. (2018). Penerapan Data Mining Dalam Pengelompokan Data Nilai Siswa Untuk Penentuan Jurusan Siswa Pada SMA Tamora Menggunakan Algoritma K-Means Clustering. *Jurnal SAINTIKOM (Jurnal Sains Manajemen Informatika Dan Komputer)*, 17(2), 228-233.
- NM Purnamasari, Achmad Syarief, & Danar Ardian Purnama. (2023). Pengelompokan Data Calon Siswa Baru Di Sekolah Menengah Kejuruan menggunakan Algoritma K-Means. *Jurnal Sistem Informasi dan Teknologi Pendidikan (JSITP)*, 4(1), 24-
- Faiza Rini, Novhirtamely Kahar, & Juliana. (2016). Penerapan Algoritma K-Means pada Pengelompokan Data Siswa Baru Berdasarkan Jurusan di SMK Negeri 1 Kota Jambi Berbasis Web. Dalam *Seminar Nasional APTIKOM (SEMNASTIKOM)* (hlm. 94).

Muhammad Norshahlan, Hendra Jaya, & Rini Kosim. (2023). Penerapan Metode Clustering Dengan Algoritma K-means Pada Pengelompokan Data Calon Siswa Baru. *Jurnal Sistem Informasi TGD*, 2(6), 1042-1053.