

Penerapan Metode K-Means untuk Klasterisasi Minat dan Bakat Siswa terhadap Ekstrakurikuler Sekolah

¹Ridho Taufiq Subagio, ²Viar Dwi Kartika, ³Muslimah
^{1,2,3}Universitas Catur Insan Cendekia
Cirebon, Indonesia

ridho.taufiq@cic.ac.id, viar.dk@cic.ac.id, muslimah.ti.20@cic.ac.id

*Penulis Korespondensi

Diajukan : 10/07/2024

Diterima : 14/08/2024

Dipublikasi : 15/08/2024

ABSTRAK

Kegiatan ekstrakurikuler adalah aktivitas yang dilakukan oleh siswa diluar jam pelajaran formal yang bertujuan mengembangkan minat, bakat dan keterampilan siswa di SMPN 2 Sedong. Dari beberapa ekstrakurikuler yang terdapat pada sekolah, siswa masih ragu-ragu untuk memilih ekstrakurikuler yang sesuai dengan minat dan bakat siswa sehingga dapat mengakibatkan kurangnya partisipasi siswa dalam mengikuti kegiatan ekstrakurikuler di sekolah. Penelitian ini bertujuan untuk menerapkan metode K-Means dalam mengklasterisasi minat dan bakat siswa terhadap ekstrakurikuler di SMPN 2 Sedong untuk membantu menghemat waktu dalam menyeleksi siswa berdasarkan minat dan bakatnya. Proses klasterisasi dilakukan menggunakan algoritma K-Means berdasarkan hasil kuesioner pada siswa kelas 7 tahun 2023/2024 untuk menentukan cluster berdasarkan data kuesioner masuk yang telah di-input oleh siswa serta sistem akan menghasilkan rekomendasi ekstrakurikuler. Hasil dari penelitian ini yaitu berupa aplikasi yang dapat digunakan oleh pembina ekstrakurikuler di SMPN 2 Sedong dengan menggunakan metode K-Means sehingga dapat mempercepat proses seleksi berdasarkan minat dan bakatnya. Dari hasil pengujian menggunakan white box diperoleh fungsionalitas dari aplikasi sesuai dengan kebutuhan pengguna sedangkan hasil pengujian white box pada aplikasi ini menunjukkan fitur dalam modul k-means sebagian besar fungsi code memiliki cyclomatic complexity (CC) yang rendah antara 1 hingga 3 sehingga aplikasi yang dibuat mengindikasikan risiko rendah terhadap cacat dalam kode, memastikan bahwa aplikasi yang dibuat cukup layak dan mudah dipahami, efisien dalam hal struktur dan aliran program serta mudah dipelihara.

Kata Kunci: Ekstrakurikuler, K-Means, Klasterisasi, Minat dan Bakat.

I. PENDAHULUAN

Pendidikan ekstrakurikuler merupakan bagian integral dari pengalaman belajar siswa di sekolah, yang bertujuan untuk mengembangkan keterampilan, minat, dan bakat diluar kurikulum akademik. Namun, dalam mengelola program ekstrakurikuler, sering kali terdapat tantangan dalam menemukan potensi siswa yang tepat dan sesuai dengan kegiatan ekstrakurikuler yang tersedia seperti memahami minat dan juga bakat dari siswa, karena tidak semua siswa memiliki kesadaran yang kuat tentang potensi mereka sendiri (Arifudin, 2022). Selain itu tantangan lainnya adalah faktor-faktor seperti keraguan dan ketidaknyamanan siswa dalam lingkungan tertentu juga dapat menjadi hambatan dalam mengeksplorasi dan mengembangkan potensi mereka melalui kegiatan ekstrakurikuler. Hal ini dapat mengakibatkan kurangnya partisipasi siswa dalam mengikuti kegiatan ekstrakurikuler yang paling sesuai dengan minat mereka dan bakat yang mereka miliki. Oleh karena itu, untuk dapat meningkatkan partisipasi dalam mengikuti ekstrakurikuler maka penulis mendapatkan solusi yaitu dengan mengelompokkan siswa tersebut untuk masuk ke ekstrakurikuler yang sesuai dengan minat dan bakatnya dengan menggunakan

metode k-means.

SMPN 2 Sedong beralamat di Jalan Irigasi, Nomor 02, Desa Sedong Lor, yang terletak di Kecamatan Sedong termasuk wilayah Kabupaten Cirebon Jawa Barat, sebagai lembaga pendidikan, memiliki kebutuhan untuk mendeteksi potensi siswa dalam berbagai kegiatan ekstrakurikuler yang ditawarkan. Namun pemilihan ekstrakurikuler di sekolah yang digunakan saat ini masih menggunakan dokumen formulir pendaftaran ekstrakurikuler serta tidak memadai dalam mengidentifikasi potensi siswa secara tepat dan juga sesuai dengan minat siswa dan bakat yang dimiliki siswa. Oleh karena itu, mengklasterisasi minat dan bakat siswa dapat membantu meningkatkan peminatan siswa dalam pemilihan ekstrakurikuler di sekolah. Dihadapkan beragam pilihan kegiatan ekstrakurikuler yaitu Paskibra, Pramuka, PMR (Palang Merah Remaja), Olahraga, Karawitan dan Marching Band, siswa siswi di SMPN 2 Sedong sering kali ragu-ragu didalam memilih ekstrakurikuler yang paling sesuai dengan minat sekaligus juga bakat mereka. Masalah yang dihadapi siswa yaitu tidak memiliki cukup informasi tentang berbagai macam kegiatan ekstrakurikuler yang tersedia di sekolah sehingga sulit bagi mereka untuk membuat pilihan yang tepat sesuai dengan minat yang dimiliki dan bakat siswa. Selain itu, beberapa siswa menghadapi kesulitan dalam menyesuaikan jadwal kegiatan ekstrakurikuler dengan kegiatan diluar sekolah seperti kursus tambahan.

Ekstrakurikuler telah menjadi sarana untuk siswa mengembangkan minat dan juga bakatnya, akan tetapi juga menjadi wahana bagi pengembangan karakter, keterampilan sosial, kepemimpinan dan nilai-nilai positif lainnya. Dalam konteks ini, peran pembina ekstrakurikuler menjadi sangat penting. Pembina ekstrakurikuler mendapatkan tantangan seperti menyusun program kegiatan ekstrakurikuler, mengarahkan dan melatih (dibantu oleh pelatih) siswa secara berkala, mengadakan pelatihan ekstrakurikuler guna menunjang tercapainya tujuan kegiatan ekstrakurikuler di sekolah, mengadakan evaluasi kegiatan ekstrakurikuler untuk mengetahui apa yang belum tercapai dan yang harus diperbaiki dalam kegiatan ekstrakurikuler serta membuat laporan pelaksanaan kegiatan ekstrakurikuler secara berkala kepada kepala sekolah. Masalah yang dihadapi oleh pembina ekstrakurikuler adalah membutuhkan waktu untuk menentukan kecocokkan antara minat dan bakat siswa dengan kegiatan ekstrakurikuler yang ada di sekolah berdasarkan formulir pendaftaran yang masih belum jelas tentang bentuk disetiap kegiatan ekstrakurikuler. Oleh karena itu, penulis menerapkan metode k-means untuk menjadikan sebuah solusi dalam mengelompokkan siswa pada ekstrakurikuler di sekolah sesuai dengan minat dan bakat mereka.

Partisipasi dalam kegiatan ekstrakurikuler di sekolah memiliki dampak positif terhadap perkembangan siswa, baik secara akademik maupun sosial. Mulanya pada saat MPLS (Masa Pengenalan Lingkungan Sekolah), siswa diperkenalkan dengan ekstrakurikuler yang ada di sekolah melalui agenda demonstrasi ekstrakurikuler. Setelah itu, siswa diajak untuk bergabung ekstrakurikuler menggunakan dokumen formulir pemilihan ekstrakurikuler. Setelah siswa kelas 7 mulai menjadi calon anggota ekstrakurikuler, siswa akan melalui beberapa tahap hingga menjadi anggota seperti pelantikan, pengukuhan, dan lain-lain sampai kelas 9 masa jabatan berakhir (purna). Oleh karena itu, pemahaman lebih mendalam tentang alur mengikuti ekstrakurikuler menjadi penting untuk membantu sekolah dan stakeholder terkait dalam meningkatkan partisipasi siswa dalam kegiatan ekstrakurikuler sehingga diharapkan akan dapat memberi manfaat yang besar untuk perkembangan siswa.

II. STUDI LITERATUR

Penelitian Terdahulu

Penelitian terkait klasterisasi menggunakan metode K-Means telah dilakukan sebelumnya. Beberapa penelitian terdahulu yang berkaitan dengan klasterisasi adalah penelitian yang dilakukan oleh (Iqbal, 2022) dengan mengembangkan metode MFEP untuk meningkatkan kualitas penentuan rekomendasi ekstrakurikuler bagi siswa di Sekolah Yayasan Pendidikan Aras Kabu Tanjungbalai. Masalah yang dihadapi adalah proses manual yang lambat dan kurang efisien dalam pengambilan keputusan ekstrakurikuler. Dengan menerapkan metode MFEP yang mencakup kriteria Minat, Bakat, Prestasi, Pengetahuan, dan Kreatifitas, penelitian ini berhasil memberikan rekomendasi yang lebih akurat, dengan kecenderungan memprioritaskan

ekstrakurikuler panduan suara daripada pramuka, olahraga, dan kesenian. Akhirnya, penelitian ini memberikan solusi efektif dalam meningkatkan efisiensi dan ketepatan rekomendasi ekstrakurikuler bagi siswa, serta memberikan kontribusi positif dalam pengembangan sistem pendukung keputusan di institusi pendidikan.

Penelitian yang dilakukan oleh (T. Magrisa et al., 2018) menghadirkan kontribusi yang berarti didalam memberi bantuan kepada siswa SMA di dalam memilih salah satu kegiatan ekstrakurikuler yang paling sesuai dengan potensi dan minat yang dimiliki. Dengan memanfaatkan Metode *Simple Multi Attribute Rating Technique* (SMART) didalam membuat sistem guna mendukung keputusan, penelitian ini berhasil mengatasi masalah kesulitan siswa dalam menentukan pilihan ekstrakurikuler. Dukungan kuat dari responden sebesar 83,415% yang menyatakan sangat setuju terhadap keefektifan sistem ini sebagai alat bantu dalam memilih kegiatan ekstrakurikuler menegaskan keberhasilan implementasi metode SMART dalam konteks ini, memberikan alternatif yang berharga bagi pengelolaan kegiatan ekstrakurikuler di lingkungan sekolah.

Penelitian yang dilakukan oleh (Rivaldi et al., 2022) merupakan upaya untuk mengetahui minat siswa terhadap kegiatan ekstrakurikuler futsal di SMA Negeri se-Kota Sukabumi. Dengan menggunakan metode survei deskriptif kuantitatif dan instrumen kuesioner melalui google form, penelitian ini mengungkapkan bahwa sebanyak 71% dari 114 siswa yang aktif mengikuti kegiatan futsal memiliki minat tinggi, sementara 20% memiliki minat rendah, dan 9% memiliki minat sangat rendah. Hasil ini menunjukkan bahwa secara keseluruhan, minat siswa terhadap kegiatan ekstrakurikuler khususnya futsal di SMAN di kota Sukabumi berada dalam kategori tinggi, memberikan gambaran yang relevan bagi pengelolaan kegiatan ekstrakurikuler dan peningkatan partisipasi siswa dalam aktivitas olahraga di sekolah tersebut.

Dalam konteks ini, penggunaan metode clustering seperti K-Means, dapat menjadi solusi yang kuat untuk mendeteksi potensi jenis ekstrakurikuler siswa. Metode ini memungkinkan identifikasi kelompok siswa berdasarkan karakteristik tertentu, seperti hobi, minat, bakat, atau nilai akademik, yang dapat membantu dalam penempatan siswa ke dalam kegiatan ekstrakurikuler yang sesuai (Nalendra et al., 2020). Namun, belum ada penelitian yang secara khusus mengaplikasikan metode K-Means untuk mengklasterisasi jenis ekstrakurikuler yang cocok untuk siswa dalam konteks lembaga pendidikan seperti SMPN 2 Sedong. Oleh karena itu, penelitian ini dianggap penting untuk dilakukan guna mengisi kesenjangan pengetahuan tersebut dan memberikan kontribusi dalam pengembangan sistem pendukung keputusan yang efektif dalam mengelola program ekstrakurikuler di lingkungan sekolah. Penulis menerapkan metode K-Means guna melakukan clustering, yang akan dituangkan dalam penelitian berjudul "Penerapan Metode K-Means untuk Klasterisasi Minat dan Bakat Siswa terhadap Ekstrakurikuler Sekolah" sehingga penelitian ini bertujuan untuk memfasilitasi proses pemilihan ekstrakurikuler bagi siswa dengan cara yang lebih tepat dan memuaskan sesuai dengan minat yang dimiliki dan bakat siswa, memberikan rekomendasi ekstrakurikuler kepada pihak sekolah SMPN 2 Sedong, menerapkan metode K-Means untuk mengklasterisasi minat dan bakat siswa di SMPN 2 Sedong terhadap ekstrakurikuler sekolah agar menyederhanakan dan memperjelas proses pemilihan kegiatan ekstrakurikuler bagi siswa.

III. METODE K-MEANS

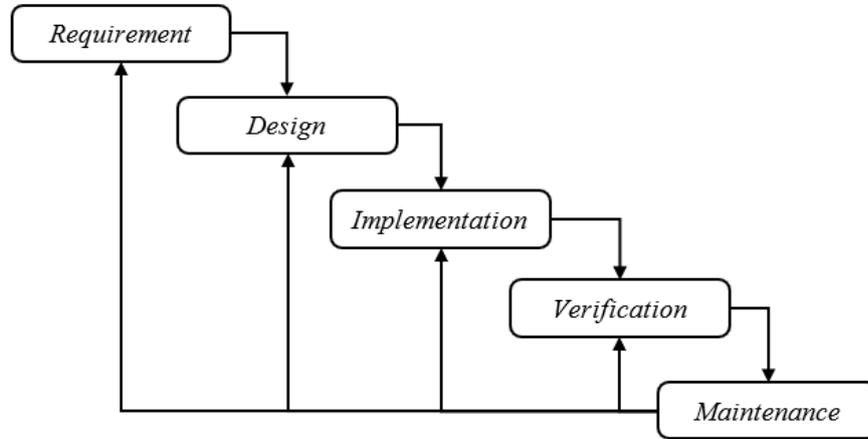
K-Means clustering adalah salah satu teknik pengelompokan data non-hierarchical yang membagi data ke dalam satu atau lebih kelompok. Data yang memiliki sifat yang serupa dikelompokkan bersama dalam satu kelompok, sementara data dengan sifat yang berbeda dikelompokkan ke kelompok lainnya. Hal ini bertujuan untuk menciptakan kelompok-kelompok di mana data yang berada dalam satu kelompok memiliki tingkat variasi yang rendah.

Penelitian ini menggunakan data penelitian primer dan sekunder. Data primer terdiri dari hasil observasi dan wawancara, sedangkan data sekunder berupa analisis literatur dari jurnal ilmiah, artikel, dan repositori dari perpustakaan. Alat penelitian yang digunakan terdiri dari perangkat keras yaitu laptop dan perangkat lunak menggunakan bahasa pemrograman Python dengan library Flask. (Alfarizi et al., 2023)

Prosedur penelitian terdiri dari metode pengumpulan data, metode pengembangan perangkat

lunak, metode pengolahan dan analisis data. Metode pengumpulan data yang digunakan adalah kuesioner, observasi, wawancara dan studi literatur.

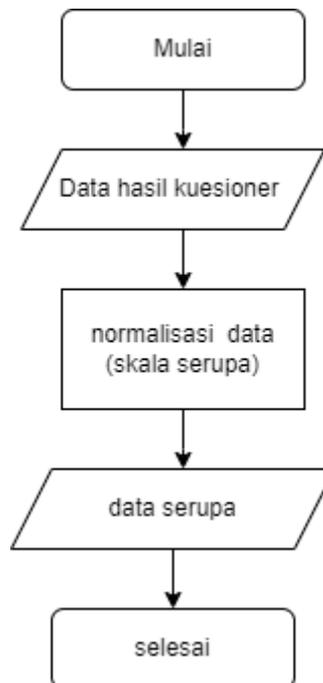
Metode pengembangan perangkat lunak menggunakan metode *waterfall* (Voutama & Novalia, 2022) yang memiliki 5 tahap berkaitan seperti pada Gambar 1.



Gambar 1. Metode waterfall

Tahap Requirement melibatkan identifikasi kebutuhan fungsional dan non-fungsional aplikasi berdasarkan pada tujuan penelitian dan kebutuhan pengguna. Analisis ini mencakup pemahaman pengguna tentang fitur-fitur yang diperlukan, fungsionalitas aplikasi, antarmuka pengguna, dan spesifikasi teknis lainnya.

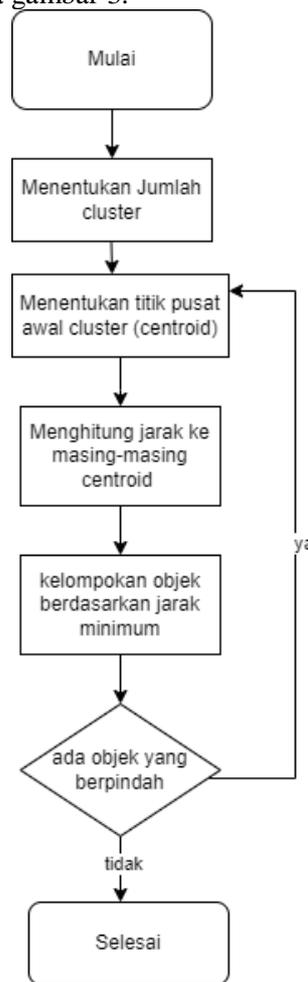
Tahap design dilakukan proses merancang struktur dan arsitektur aplikasi. Ini mencakup perancangan antarmuka pengguna (UI/UX), perencanaan database, serta pemodelan sistem secara keseluruhan. Rancangan ini akan menjadi panduan bagi pengembang dalam tahap implementasi. Berikut ini merupakan data hasil kuesioner yang dinormalisasikan datanya agar fiturnya memiliki skala yang serupa disajikan pada gambar 2.



Gambar 2. Pre-processing Data

Kemudian setelah dilakukan konversi dari categorical ke numeric, maka untuk selanjutnya data siap untuk dilanjutkan ke tahapan penghitungan K-Means. Berikut ini adalah flowchart dari

perhitungan K-Means disajikan pada gambar 3.



Gambar 3. Flowchart K-means

Tahap implementasi melibatkan penerjemahan rancangan aplikasi menjadi kode-kode yang dapat dieksekusi oleh komputer. Pengembang akan menulis kode sesuai dengan spesifikasi yang telah ditetapkan dalam tahap perancangan. Dalam konteks penggunaan Python dengan Flask dan template AdminLTE, pengembang akan menggunakan bahasa pemrograman Python untuk mengimplementasikan logika aplikasi dan template AdminLTE untuk mendesain antarmuka pengguna.

Tahap verification pada aplikasi dilakukan serangkaian pengujian untuk memverifikasi kesesuaian dengan kebutuhan dan spesifikasi yang telah ditetapkan. Pengujian ini mencakup pengujian fungsionalitas, keamanan, dan kinerja. Hasil dari pengujian tersebut akan digunakan sebagai dasar untuk melakukan perbaikan atau penyesuaian jika diperlukan.

Tahap maintenance, pada tahap ini aplikasi akan dikirimkan kepada pengguna atau disiapkan untuk digunakan dalam lingkungan produksi. Pemeliharaan juga diperlukan untuk memastikan bahwa aplikasi tetap berjalan dengan baik dan memenuhi kebutuhan pengguna seiring waktu.

Metode pengolahan data dalam penelitian ini diantaranya 1) Menentukan jumlah cluster yang diperlukan, 2) Mengalokasikan data secara acak ke dalam cluster, 3) Menghitung centroid atau rata-rata dari data yang terdapat di setiap cluster, 4) Menetapkan setiap data ke centroid atau rata-rata terdekat, 5) Kembali kelangkah 3 apabila terdapat data yang berubah cluster, perubahan nilai centroid melebihi nilai threshold yang telah ditentukan, atau perubahan nilai pada fungsi objektif melampaui nilai threshold yang telah ditentukan. Teknik analisis data yang dilakukan adalah pengumpulan data dari berbagai sumber, seperti survei, wawancara, atau observasi, akan dimasukkan ke dalam format yang sesuai dan dimanipulasi jika diperlukan menggunakan pustaka seperti Pandas dan NumPy dalam bahasa pemrograman Python. Langkah berikutnya adalah

melakukan pra-pemrosesan data, yang mencakup pembersihan data untuk menghapus entri yang tidak lengkap atau tidak relevan, penanganan nilai yang hilang, dan normalisasi data untuk memastikan bahwa setiap atribut memiliki skala yang serupa. (Ariawan, 2019)

Setelah data diproses, analisis eksploratif akan dilakukan untuk memahami karakteristik dasar dari data, seperti distribusi, korelasi antar atribut, dan pola-pola yang mungkin tersembunyi. Visualisasi data akan menjadi bagian integral dari analisis ini, dimana Matplotlib atau Plotly dapat digunakan untuk membuat grafik dan plot yang menggambarkan data dengan jelas (Hidayat et al., 2023). Selanjutnya, data akan siap untuk penggunaan algoritma K-Means, dimana pengelompokan data akan dilakukan berdasarkan pada kemiripan antar data. Scikit-learn akan menjadi pustaka yang digunakan untuk mengimplementasikan algoritma K-Means ini. Scikit-Learn adalah library Python yang dikembangkan untuk memberikan kemudahan dalam pengkodean machine learning dalam Python. (Fahmi, 2023)

Setelah pengelompokan selesai, hasilnya akan dievaluasi dan diinterpretasikan. Ini dapat melibatkan penggunaan metrik evaluasi. Selain itu, analisis lebih lanjut juga akan dilakukan untuk memahami karakteristik setiap kelompok dan mengidentifikasi pola-pola yang muncul. Hasil analisis ini akan diinterpretasikan secara menyeluruh untuk mendapatkan pemahaman yang lebih dalam tentang potensi ekstrakurikuler siswa. Selama seluruh proses ini, dokumentasi yang baik akan sangat penting, termasuk pencatatan setiap langkah yang diambil, parameter yang digunakan, serta hasil-hasil yang diperoleh. Hal ini akan membantu memvalidasi temuan dan memungkinkan replikasi atau pengujian ulang oleh peneliti lain di masa depan. Dengan mengikuti metodologi pengolahan dan analisis data ini, diharapkan penelitian ini akan menghasilkan hasil yang akurat dan bermakna tentang potensi ekstrakurikuler siswa.

IV. HASIL DAN PEMBAHASAN

Penerapan metode K-means dilakukan dengan tahap pre-processing dengan normalisasi data minat ekstrakurikuler untuk memastikan semua fitur memiliki skala yang serupa kemudian penentuan jumlah cluster. Clustering adalah metode pembelajaran mesin yang digunakan untuk mengelompokkan data menjadi beberapa kelompok atau klaster yang sesuai (Nahjan et al., 2023), selanjutnya pelatihan Model dengan menerapkan algoritma K-Means pada data untuk mengklasifikasikan siswa ke dalam kelompok-kelompok berdasarkan minat mereka terhadap ekstrakurikuler, dan evaluasi hasil kecocokan dan homogenitas klaster yang dihasilkan untuk memastikan kualitas rekomendasi ekstrakurikuler.

Tabel 1 menyajikan data siswa berdasarkan hasil kuesioner siswa SMPN 2 Sedong kelas 7B dan 7C tahun ajaran 2023/2024, yang menjadi acuan untuk klaster siswa baru berikutnya.

Tabel 1. Tabel berdasarkan Konversi Variabel pada Siswa

No	Nama Siswa	L/P	Minat	Bakat	Hobi	Ekstrakurikuler
1	Siti Nur Mutiara	2	1	1	1	5
2	Mega Ayu Lestari	2	2	2	2	3
3	Bayu Saputra	1	1	3	3	3
4	Aulia Rahma Putri	2	3	3	3	4
5	Neli Nadziroh Zulfa	2	2	2	4	5
6	Fanessa Lipatunnissa	2	2	2	3	3
7	Shinta Dewi Lestari	2	1	4	4	6
8	Reva Setiana Dewi	2	1	1	3	5
9	Wisnu Aditia	1	1	2	3	5
10	Shafa Aulia Azzahra	2	3	3	5	4
11	Titik Werdhi Cahya Ningslam	2	4	5	6	2
12	Andra Jaka Dhiva Ain Dian	1	2	1	7	5

	Syah					
13	Muhammad Arshavin	1	2	2	3	3
14	Serli Agustina	2	5	3	8	4
15	Savira Maharani	2	1	4	3	6
16	Revani Dian Nuril Aini	2	5	4	2	6
17	Amira Wahyuni	2	3	3	7	4
18	Yeyen Yasinta	2	6	3	9	3

Keterangan :

Jenis kelamin

1. Laki-laki
2. Perempuan

Minat

1. Kepemimpinan
2. Aktifitas Fisik dan Mengikuti Kompetisi
3. Pelayanan Kemanusiaan
4. Bermusik dan Tampil dipublik
5. Belajar Budaya Lokal
6. Petualangan dan Kerja Tim

Bakat

1. Pengaturan Formasi di Paskibra
2. Berbagai Macam Olahraga
3. Merawat dan Memberikan Pertolongan kepada Orang yang Membutuhkan
4. Bangun Perkemahan dan Inisiatif dalam Kegiatan Outdoor
5. Memaikan Instrumen Musik
6. Memaikan Instrumen Tradisional

Hobi

1. Menonton film
2. Menyanyi atau bermain musik
3. Bermain game
4. Olahraga
5. Menulis
6. Melukis dan menggambar
7. Memasak

Berikut ini contoh perhitungan manual, misalkan kita memiliki 18 data siswa dengan dua fitur (x dan y) dan ingin mengklasterkannya menjadi 6 cluster menggunakan metode K-Means. Langkah pertama melakukan inialisasi centroid yang hasilnya disajikan dalam tabel 2.

Tabel 2. Inialisasi Centroid Awal

Cluster	Centroid Awal (jenis kelamin, minat, bakat, hobi)
1	(2 ; 1,5 ; 1,5 ; 1,5)
2	(2 ; 4 ; 3,5 ; 2,5)
3	(1,5 ; 1 ; 3,3 ; 3,3)
4	(2 ; 1 ; 1 ; 3)
5	(1,7 ; 2 ; 2 ; 3,3)
6	(1,8 ; 3,8 ; 3 ; 7)

Pada tabel 2 pemilihan centroid awal dilakukan secara acak, contoh centroid awal

menunjukkan beberapa titik data yang dipilih sebagai titik pusat awal untuk masing-masing cluster. Langkah kedua melakukan perhitungan jarak menggunakan rumus euclidean distance :

$$D = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 + (w_1 - w_2)^2} \quad (1)$$

Berikut ini data mengenai perhitungan jarak disajikan pada tabel 3.

Tabel 3. Perhitungan Jarak

Siswa	Jarak					
	C1	C2	C3	C4	C5	C6
1	0,866	4,183	3,559	2,000	2,622	6,932
2	0,866	2,550	2,380	1,732	1,369	5,421
3	2,398	3,240	1,000	2,236	1,369	4,972
4	2,598	1,225	2,160	2,828	1,696	4,089
5	2,598	2,915	2,082	1,732	0,935	3,659
6	1,658	2,550	2,000	1,414	0,612	4,515
7	3,571	3,391	0,816	3,162	2,318	4,249
8	1,658	3,937	2,708	0,000	1,369	5,297
9	1,936	3,536	1,826	1,414	0,935	5,072
10	4,093	2,739	2,708	3,464	2,424	2,173
11	6,225	3,808	4,243	5,831	4,677	2,248
12	5,635	5,612	4,690	4,243	3,921	2,838
13	1,936	2,739	2,082	1,732	0,612	4,589
14	7,533	5,612	6,191	6,708	5,863	1,546
15	2,958	3,082	0,577	3,000	2,208	5,006
16	4,330	1,225	4,243	5,099	4,047	5,233
17	5,895	4,637	4,243	4,899	4,108	0,850
18	8,874	6,819	7,594	8,062	7,237	2,953

Pada tabel 3 merupakan hasil perhitungan jarak antara setiap data dengan centroid untuk masing-masing cluster. Dengan data ini, kita dapat melanjutkan proses K-Means untuk mengelompokkan data ke dalam cluster yang sesuai. Langkah ketiga dilakukan pengelompokkan data ke dalam klaster dengan jarak terdekat yang disajikan dalam tabel 4.

Langkah keempat melakukan pembaruan centroid lanjutan dengan menghitung rata-rata dari setiap cluster untuk mendapatkan posisi centroid yang baru. (Lubis & Hendrik, 2023)

Hasil yang didapatkan disajikan pada tabel 5. Langkah kelima melakukan iterasi yaitu mengulangi langkah-langkah sebelumnya (2-4) hingga tidak ada perubahan signifikan dalam posisi centroid atau hingga mencapai iterasi maksimum yang ditentukan pada tahap sebelumnya. Kemudian proses berlanjut sampai tidak ditemukan perubahan yang signifikan dalam posisi centroid atau sampai mencapai batas iterasi maksimum yang ditentukan sebelumnya. Setelah itu, kita dapat menggunakan centroid terakhir sebagai representasi dari setiap cluster untuk membuat rekomendasi ekstrakurikuler kepada siswa berdasarkan cluster yang mereka miliki. Ini adalah gambaran lengkap tentang bagaimana proses K-Means bekerja, dari inisialisasi centroid hingga konvergensi, dengan fitur jenis kelamin, minat, bakat, dan hobi, serta 6 cluster baru yang telah ditentukan.

Tabel 4. Pengelompokan Data

Siswa	Jarak				I.	C5	C6
	C1	C2	C3	C4			
1	1	0	0	0	0	0	
2	1	0	0	0	0	0	
3	0	0	1	0	0	0	
4	0	1	0	0	0	0	
5	0	0	0	0	1	0	
6	0	0	0	0	1	0	
7	0	0	1	0	0	0	
8	0	0	0	1	0	0	
9	0	0	0	0	1	0	
10	0	0	0	0	0	1	
11	0	0	0	0	0	1	
12	0	0	0	0	0	1	
13	0	0	0	0	1	0	
14	0	0	0	0	0	1	
15	0	0	1	0	0	0	
16	0	1	0	0	0	0	
17	0	0	0	0	0	1	
18	0	0	0	0	0	1	

Pada tabel 4. pengelompokan data ini adalah keanggotaan cluster untuk setiap siswa berdasarkan nilai jarak terkecil.

Tabel 5. Pembaruan Centroid Lanjutan

Cluster	Centroid Baru (jenis_kelamin, minat, bakat, hobi)
C1	(2 ; 1,5 ; 1,5 ; 1,5)
C2	(2 ; 4 ; 3,5 ; 2,5)
C3	(1,6 ; 1 ; 3,6 ; 3)
C4	(2 ; 1 ; 2 ; 3,2)
C5	(1,5 ; 1,7 ; 2 ; 3,2)
C6	(1,8 ; 3,8 ; 3 ; 7)

Pada tabel 5. merupakan centroid baru berdasarkan hasil dari clustering yang menjadi centroid untuk iterasi selanjutnya. Langkah keenam menentukan hasil dan rekomendasi, Setelah ditemukan cluster masing-masing, maka akan dilakukan pengujian homogenisasi (proses pencampuran menjadi seragam) dari masing-masing siswa pada setiap cluster yang disajikan pada tabel 6.

Dari tabel 6. dapat diketahui bahwa siswa 1 mendapat rekomendasi cluster 1 yaitu Karawitan, siswa 2 mendapat rekomendasi cluster 2 yaitu Karawitan, siswa 3 mendapat rekomendasi cluster 3 Olahraga, dan seterusnya hingga siswa 18 dengan melalui proses perhitungan manual menggunakan algoritma K-means tersebut.

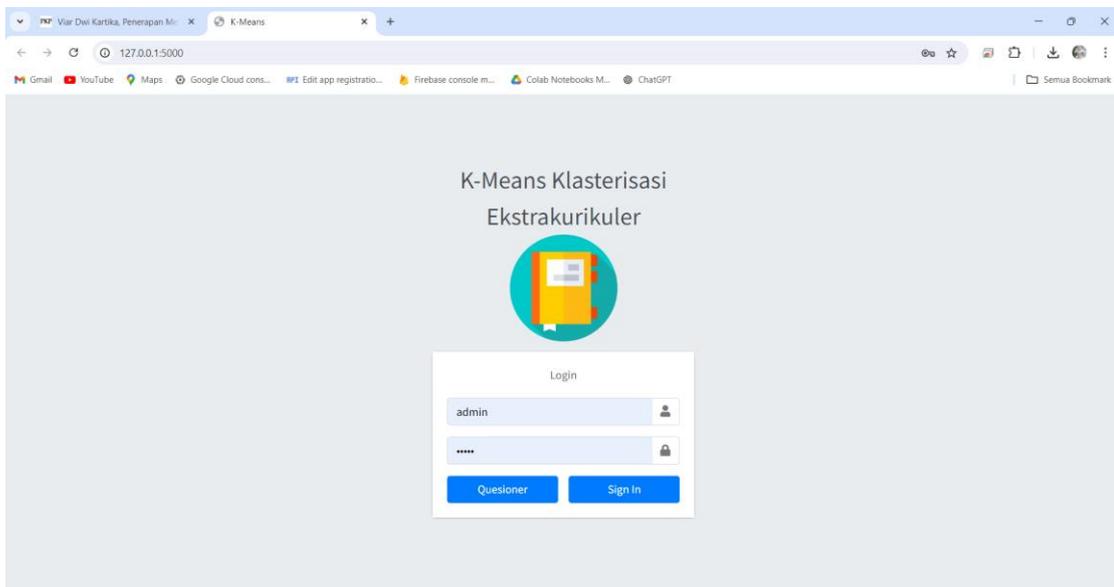
Tabel 6. Hasil dan Rekomendasi

Siswa	Cluster	Rekomendasi Ekstrakurikuler
1	1	Karawitan

Siswa	Cluster	Rekomendasi Ekstrakurikuler
2	1	Karawitan
3	3	Olahraga
4	2	Marching Band
5	5	Paskibra
6	5	Paskibra
7	3	Olahraga
8	4	PMR
9	5	Paskibra
10	6	Pramuka
11	6	Pramuka
12	6	Pramuka
13	5	Paskibra
14	6	Pramuka
15	3	Olahraga
16	2	Marching Band
17	6	Pramuka
18	6	Pramuka

Tampilan Program

Pada tampilan program yaitu menampilkan hasil akhir sistem beserta penjelasan fungsi menu dari program tersebut. Berikut ini merupakan tampilan program k-means klasterisasi ekstrakurikuler.



Gambar 4. Form Login

Form login bertujuan untuk memastikan bahwa aplikasi digunakan oleh orang yang memiliki hak akses admin (pembina ekstrakurikuler) dan siswa dengan verifikasi berupa username dan password untuk kemudian diharuskan meng-klik tombol sign in. Kemudian formulir kuesioner disajikan pada gambar 5.

Gambar 5. Formulir Kuesioner

Formulir kuesioner digunakan untuk mengisi data formulir ketika siswa ingin memilih ekstrakurikuler terdahulu sebelum dilakukan klusterisasi pada aplikasi yang digunakan oleh admin. Data kuesioner ini akan masuk kedalam file csv yang akan masuk kedalam sistem coding lalu data tersebut akan dikelola oleh admin untuk mencari rekomendasi ekstrakurikuler yang cocok untuk siswa tersebut. Catatan untuk mengisi kuesioner siswa tidak dapat mengisi lagi jika NISN yang digunakan sama. Berikut ini tampilan dashboard disajikan pada gambar 6.

No	Nama Siswa	L/P	Minat	Bakat	Hobi	Ekstrakurikuler	Tahun
1	Siti Nur Mutiara	P	Kepemimpinan	Pengaturan Formasi di Paskibra	Menonton film	Paskibra	2023
2	Mega Ayu Lestari	P	Aktivitas Fisik dan Mengikuti Kompetisi	Berbagai Macam Olahraga	Menyanyi atau Bermain Musik	Olahraga	2023
3	Bayu Saputra	L	Kepemimpinan	Merawat dan Memberikan Pertolongan kepada Orang yang membutuhkan	Bermain game	Olahraga	2023
4	Aulia Rahma Putri	P	Pelayanan Kemanusiaan	Merawat dan Memberikan Pertolongan kepada Orang yang membutuhkan	Bermain game	PMR	2023
5	Neli Nadzirah Zulfa	P	Aktivitas Fisik dan Mengikuti Kompetisi	Berbagai Macam Olahraga	Olahraga	Paskibra	2023
6	Fanessa Lipatunnissa	P	Aktivitas Fisik dan Mengikuti Kompetisi	Berbagai Macam Olahraga	Bermain game	Olahraga	2023

Gambar 6. Dashboard

Dashboard admin tampilan ini menunjukkan informasi data siswa yang menjadi data training pada kelas 7B dan 7C tahun ajaran 2023/2024 yang mencakup nama, jenis kelamin, minat, bakat, hobi, dan kegiatan ekstrakurikuler dengan tujuan menganalisis dan mengelompokkan siswa berdasarkan variabel-variabel tersebut. Berikut ini tampilan halaman dataset disajikan pada gambar 7.

No	Nama Siswa	L/P	Minat	Bakat	Hobi	Ekstrakurikuler	Tahun
1	Siti Nur Mutiara	P	Kepemimpinan	Pengaturan Formasi di Paskibra	Menonton film	Paskibra	2023
2	Mega Ayu Lestari	P	Aktivitas Fisik dan Mengikuti Kompetisi	Berbagai Macam Olahraga	Menyanyi atau Bermain Musik	Olahraga	2023
3	Bayu Saputra	L	Kepemimpinan	Merawat dan Memberikan Pertolongan kepada Orang yang membutuhkan	Bermain game	Olahraga	2023
4	Aulia Rahma Putri	P	Pelayanan Kemanusiaan	Merawat dan Memberikan Pertolongan kepada Orang yang membutuhkan	Bermain game	PMR	2023
5	Neli Nadziroh Zulfa	P	Aktivitas Fisik dan Mengikuti Kompetisi	Berbagai Macam Olahraga	Olahraga	Paskibra	2023
6	Fanessa Lipatunnissa	P	Aktivitas Fisik dan Mengikuti Kompetisi	Berbagai Macam Olahraga	Bermain game	Olahraga	2023
7	Shinta Dewi Lestari	P	Kepemimpinan	Bangun Perkemahan dan Inisiatif dalam Kegiatan Outdoor	Olahraga	Pramuka'	2023

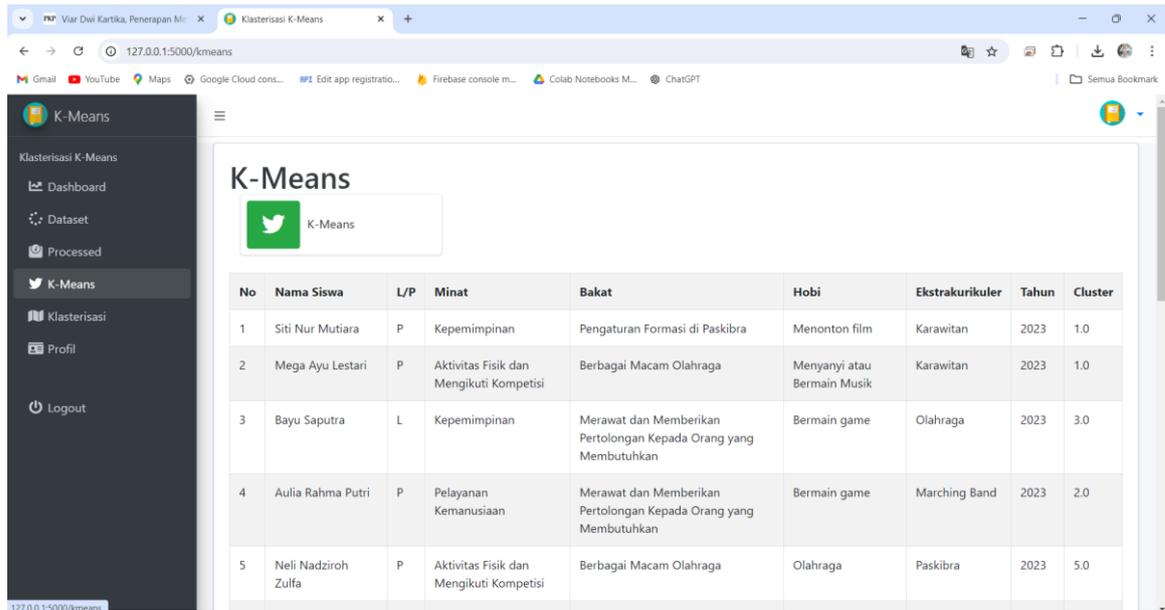
Gambar 7. Halaman Dataset

Halaman dataset ini berisikan tampilan data siswa yang mana menunjukkan nama siswa, jenis kelamin, minat, bakat dan hobi, ekstrakurikuler dan tahun ajaran siswa. Dataset ini digunakan untuk melakukan pre-processing untuk diubah data categorical menjadi numeric untuk melanjutkan proses klasterisasi pada halaman processed. Berikut ini tampilan halaman pre-processed disajikan pada gambar 8.

No	Nama Siswa	L/P	Minat	Bakat	Hobi	Ekstrakurikuler	Tahun
1	Siti Nur Mutiara	2	1	1	1	Paskibra	2023
2	Mega Ayu Lestari	2	2	2	2	Olahraga	2023
3	Bayu Saputra	1	1	3	3	Olahraga	2023
4	Aulia Rahma Putri	2	3	3	3	PMR	2023
5	Neli Nadziroh Zulfa	2	2	2	4	Paskibra	2023
6	Fanessa Lipatunnissa	2	2	2	3	Olahraga	2023
7	Shinta Dewi Lestari	2	1	4	4	Pramuka'	2023
8	Reva Setiana Dewi	2	1	1	3	Paskibra	2023
9	Wisnu Aditia	1	1	2	3	Paskibra	2023
10	Shafa Aulia Azzahra	2	3	3	5	PMR	2023

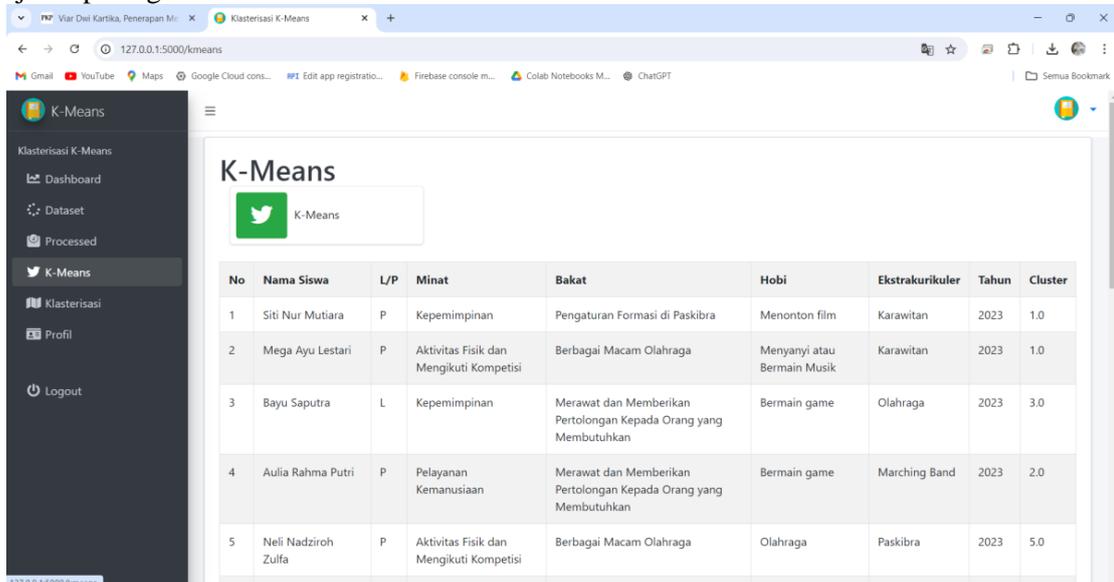
Gambar 8. Preprocessed

Halaman “Preprocessed” yang menampilkan tabel data siswa setelah melakukan preprocessing, data diubah dari categorical menjadi numeric untuk kolom jenis kelamin, minat, bakat, hobi dan kegiatan ekstrakurikuler sehingga siap untuk dilakukan penghitungan clustering menggunakan K-Means. Berikut ini tampilan halaman k-means disajikan pada gambar 9.



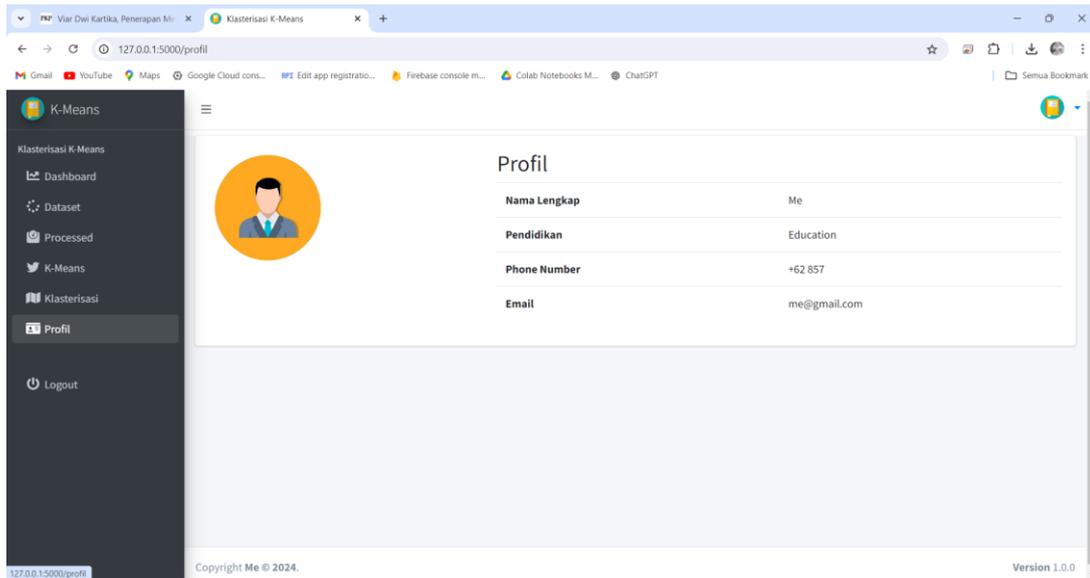
Gambar 9. K-means

Pada halaman K-Means sudah menghasilkan cluster sesuai dengan hasil perhitungan K-Means pada Microsoft Excel sebelum dijadikan sebuah tampilan sistem aplikasi seperti ini. Pada halaman ini menampilkan data siswa termasuk nama, jenis kelamin, minat, bakat, hobi, ekstrakurikuler, tahun ajaran dan hasil cluster setiap siswa. Berikut ini tampilan halaman predict disajikan pada gambar 10.



Gambar 10. Klasterisasi

Pada halaman ini admin dapat melakukan klasterisasi saran untuk ekstrakurikuler yang diambil oleh siswa berdasarkan tahun ajaran, hasil cluster dan saran akan ditampilkan ketika siswa telah menginput kuesioner pada aplikasi yang akan diolah oleh sistem untuk diklaster berdasarkan tahun yang diinput oleh siswa dengan mengisi NISN, nama siswa, jenis kelamin, minat, bakat dan hobi. Pada halaman ini akan menghasilkan rekomendasi ekstrakurikuler sesuai perhitungan. Berikut ini tampilan halaman profil disajikan pada gambar 11.



Gambar 11. Profil

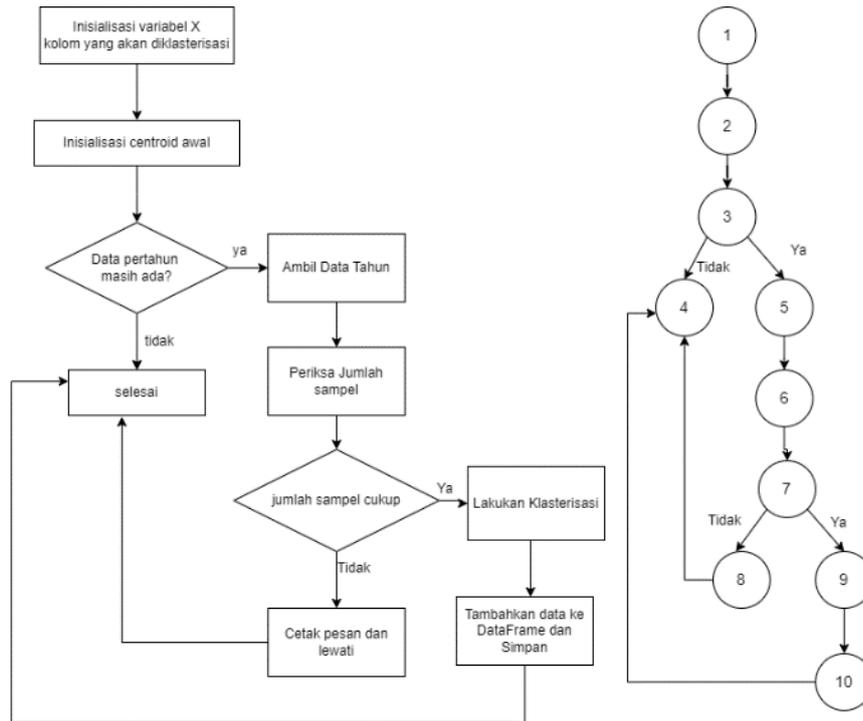
Halaman profil menampilkan identitas dari pembina ekstrakurikuler sehingga dapat diketahui biodata pada pembina ekstrakurikuler tersebut sebagai admin pengelola aplikasi.

Hasil pengujian program menggunakan pengujian *white box* dengan teknik basic path testing yang bertujuan untuk memastikan semua jalur independen dalam sebuah program telah diuji setidaknya sekali untuk membantu menguji semua jalur logis yang dilalui program selama eksekusi. Langkah pertama dalam pengujian *white box* yaitu dengan *flow graph* dan perhitungan *cyclomatic complexity* (CC) (Praniffa et al., 2023). Berikut gambar 12 code fitur k-means dan gambar 13 flow graph untuk perhitungan pada fitur k-means.

```
X = ['L/P', 'Minat', 'Bakat', 'Hobi']
initial_centroids = np.array([
    [2, 2, 2, 2],
    [2, 3, 3, 3],
    [1, 1, 3, 3],
    [2, 1, 1, 3],
    [2, 2, 2, 3],
    [2, 3, 3, 5]
])
for tahun in df['Tahun'].unique():
    data_tahun = df[df['Tahun'] == tahun][X]
    # Melakukan klasterisasi dengan K-Means
    n_clusters = 6
    if len(data_tahun) >= n_clusters:
        # Melakukan klasterisasi dengan K-Means
        kmeans = KMeans(n_clusters=n_clusters, init=initial_centroids, n_init=1, random_state=42).fit(data_tahun)
        df.loc[df['Tahun'] == tahun, 'Cluster'] = kmeans.labels_ + 1
    else:
        print(f"Skip tahun {tahun} karena kurangnya sampel ({len(data_tahun)}) dari cluster ({n_clusters})")
```

Gambar 12. Code untuk Klasterisasi Data Menggunakan K-means

Pada gambar 12. merupakan code untuk melakukan klasterisasi data dengan menggunakan algoritma k-means.



Gambar 13. Flow Graph Menghitung K-means

Pada gambar 13. merupakan representasi flowchart dan flow graph pada fitur k-means proses iterasi untuk mengklusterisasi data tahunan. Setiap tahun diperiksa apakah masih ada data yang harus diproses, kemudian data diambil dan jumlah sampel diperiksa. Jika jumlah sampel mencukupi, klusterisasi dilakukan dan hasilnya disimpan. Jika tidak mencukupi data tahun tersebut mencetak pesan dan dilewati. Proses berulang sampai semua data tahunan diproses. Setelah itu melanjutkan proses selanjutnya untuk menghitung cyclomatic complexity (CC). Berikut adalah perhitungan cyclomatic complexity serta penjelasannya :

Flow graph : 1 → 2 → 3 → 4 → 5 → 6 → 7 → 8 → 9 → 10

1. E (jumlah edge) = 11 (1→2, 2→3, 3→4, 3→5, 5→6, 6→7, 7→8, 7→9, 9→10, 10→3, 8→3)
2. N (jumlah node) = 10 (node 1,2,3,4,5,6,7,8,9,10)
3. P (jumlah komponen terhubung) = 1

$$V(G) = E - N + 2P = 11 - 10 + 2(1) = 3$$

Dari perhitungan tersebut didapatkan hasil cyclomatic complexity pada code ini adalah 3. Hasil pengujian pada fitur k-means menunjukkan bahwa code tersebut memiliki CC yang rendah sehingga aplikasi yang dibuat mengindikasikan risiko rendah terhadap cacat kode dan mudah dipelihara.

V. KESIMPULAN

Berdasarkan aplikasi yang telah dibuat dengan implementasi algoritma k-means diperoleh dengan mengklusterisasi minat dan bakat siswa pada SMPN 2 Sedong, ditemukan bahwa siswa yang mengalami keraguan dalam memilih ekstrakurikuler dapat lebih diarahkan, sehingga potensi siswa dalam mengikuti kegiatan ekstrakurikuler dapat ditingkatkan.

Pembina ekstrakurikuler dapat menghemat waktu dalam menyeleksi siswa dengan adanya sistem ini, selain itu pembina dapat memastikan bahwa siswa ditempatkan pada ekstrakurikuler yang sesuai dengan minat dan bakat mereka sehingga pelaksanaan kegiatan ekstrakurikuler menjadi lebih banyak diminati.

Dengan memanfaatkan algoritma K-Means yang diterapkan dalam sistem ini berhasil mengklusterisasikan minat dan bakat siswa dengan akurat dan hasil klusterisasi yang ditampilkan sesuai dengan data yang diinput sehingga menunjukkan keandalan sistem dalam memberikan rekomendasi yang tepat berdasarkan analisis data.

VI. REFERENSI

- Alfarizi, M. Z., T., M. R. S., A., M., A. G., & Elgar, M. (2023). Penggunaan Python Sebagai Bahasa Pemrograman untuk Machine Learning dan Deep Learning . *Karimah Tauhid*, 2(1), 1–6.
- Ariawan, P. A. (2019). Optimasi pengelompokan data pada metode K-means dengan analisis outlier. *J. Nas. Teknol. Dan Sist. Inf.*, 5(2), 88–95.
- Arifudin, O. (2022). Optimalisasi Kegiatan Ekstrakurikuler dalam Membina Karakter Peserta Didik. In *JIP-Jurnal Ilmiah Ilmu Pendidikan* (Vol. 5, Issue 3). <http://Jiip.stkipyapisdmpu.ac.id>
- Fahmi, M. N. (2023). Implementasi Mechine Learning menggunakan Python Library : Scikit-Learn (Supervised dan Unsupervised Learning). *Sains Data Jurnal Studi Matematika Dan Teknologi*, 1(2), 87–96. <https://doi.org/10.52620/sainsdata.v1i2.31>
- Hidayat, I., Tolago, A. I., Dako, R. D. R., & Ilham, J. (2023). Analisis Data Eksploratif Capaian Indikator Kinerja Utama 3 Fakultas Teknik. *Jambura Journal of Electrical and Electronics Engineering*, 5(2), 185–191.
- Iqbal, M. (2022). Metode MFEP Dalam Meningkatkan Kualitas Penentuan Rekomendasi Ekstrakurikuler. *Jurnal Teknis*, 2(1), 5–10.
- Lubis, S. S., & Hendrik, B. (2023). Implementasi Data Mining Pengelompokan Data Penjualan Berdasarkan Pembelian dengan Menggunakan Algoritma K-Means pada UD. *Martua. Journal of Information System and Education Development*, 1(3), 36–41.
- Nahjan, M. R., H. N., & Voutama, A. (2023). Implementasi Rapidminer Dengan Metode Clustering K-Means Untuk Analisa Penjualan Pada Toko Oj Cell. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 101–104.
- Nalendra, A. K., Mujiono, M., A. R., & Utama, A. S. W. (2020). Implementasi Algoritma K-Mean dalam Pengelompokan Data Kecelakaan di Kabupaten Kediri. *VOCATECH: Vocational Education and Technology Journal*, 1(2), 118–125.
- Praniffa, A. C., Syahri, A., S., F., F. U., Giansyah, Q. A., & Hamzah, M. (2023). Pengujian Sistem Informasi Parkir Berbasis Web Pada UIN SUSKA RIAU Menggunakan White Box dan Black Box Testing. *Jurnal Testing Dan Implementasi Sistem Informasi*, 1(1), 1–16.
- Rivaldi, M. A., Nurudin, A. A., & Nugraheni, W. (2022). Survei Minat Siswa Dalam Mengikuti Kegiatan Ekstrakurikuler Futsal SMA Negeri Se-Kota Sukabumi. *Jurnal Educatio*, 8(2), 498–504.
- T. Magrisa, K. D. K. Wardhani, & M. R. A. Saf. (2018). “Implementasi Metode SMART pada Sistem Pendukung Keputusan Pemilihan Kegiatan Ekstrakurikuler untuk Siswa SMA,.” *Inform. Mulawarman J. Ilm. Ilmu Komput*, 13(1), 49–55.
- Voutama, A., & Novalia, E. (2022). Perancangan Sistem Informasi Plakat Wisuda Berbasis Web Menggunakan UML dan Model Waterfall. In *Syntax: Jurnal Informatika* (Vol. 11, Issue 01).