

DATA MINING USING K-MEANS CLUSTERING ALGORITHM TO KNOW STUDENTS' INTERESTS AND TALENTS IN EXTRACURRICULAR ACTIVITIES (Case Study: SMK Setia Budi Binjai)

¹Dita Sahputri, ²Yani Maulita, ³Novriyenni

¹Program Studi Sistem Informasi, STMIK KAPUTAMA Binjai

²Program Studi Sistem Informasi, STMIK KAPUTAMA Binjai

³Program Studi Manajemen Informatika, STMIK KAPUTAMA Binjai

Jl. Veteran No. 4A-9A, Tangsi, Binjai Kota Binjai, Sumatera Utara 20714

e-mail: ¹ditasahputri07@gmail.com, ²yani.maulita@gmail.com ,

³novriyenni.sikumbang@gmail.com

*e-mail: ²yani.maulita@gmail.com

Received: July 17, 2023
Revised: August 2, 2023
Accepted: August 9, 2023

Page : 107-115

Abstrak : Pemahaman tentang minat dan bakat siswa/siswi pada kegiatan ekstrakurikuler memiliki peran penting dalam pengembangan pendidikan dan pembelajaran yang holistik. Dalam penelitian ini, kami mengusulkan penggunaan metode Data Mining dengan menggunakan algoritma K-means Clustering untuk menganalisis minat dan bakat siswa/siswi dalam konteks kegiatan ekstrakurikuler.

Penelitian ini dilakukan dengan menggunakan data survei yang dikumpulkan dari siswa/siswi sekolah di sebuah daerah. Data yang dikumpulkan meliputi informasi tentang minat dan bakat siswa/siswi terkait berbagai kegiatan ekstrakurikuler yang tersedia di sekolah. Tujuan utama penelitian ini adalah untuk mengklasifikasikan siswa/siswi berdasarkan pola minat dan bakat mereka menggunakan algoritma K-means Clustering. Diharapkan bahwa penelitian ini akan memberikan wawasan yang berharga bagi pihak sekolah dalam memahami minat dan bakat siswa/siswi, sehingga dapat mengoptimalkan penawaran kegiatan ekstrakurikuler yang sesuai dengan kebutuhan dan preferensi siswa/siswi. Hasil penelitian ini juga dapat membantu siswa/siswi dalam memilih kegiatan ekstrakurikuler yang sesuai dengan minat dan bakat mereka, sehingga dapat meningkatkan partisipasi mereka dalam kegiatan tersebut.

Kata kunci : Data Mining, K-Means ,Clustering, Minat Dan Bakat Pada Kegiatan Estrakulikuler

Abstract : *An understanding of students' interests and talents in extracurricular activities has an important role in the development of holistic education and learning. In this study, we propose the use of the Data Mining method using the K-means Clustering algorithm to analyze students' interests and talents in*

the context of extracurricular activities.

This research was conducted using survey data collected from school students in an area. The data collected includes information about the interests and talents of students related to various extracurricular activities available at school. The main objective of this research is to classify students based on their interest patterns and talents using the K-means clustering algorithm. It is hoped that this research will provide valuable insights for the school in understanding the interests and talents of students, so as to optimize the offer of extracurricular activities that suit the needs and preferences of students. The results of this study can also assist students in choosing extracurricular activities that suit their interests and talents, so as to increase their participation in these activities.

Keywords : *Keywords: Data Mining, K-Means, Clustering, Interests and Talents in Extracurricular Activities*



Journal of Engineering, Technology and Computing (JETCom) This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-nc-sa/4.0/).

1 Pendahuluan (or Introduction)

Ekstrakurikuler adalah kegiatan pendidikan yang dilakukan oleh peserta didik diluar jam belajar kurikulum standar sebagai perluasan dari kegiatan kurikulum dan dilakukan di bawah bimbingan sekolah dengan tujuan untuk mengembangkan kepribadian bakat, minat, dan kemampuan peserta didik yang lebih luas atau di luar minat yang dikembangkan oleh kurikulum. Dengan adanya kegiatan ekstrakurikuler diharapkan peserta didik dapat melakukannya di luar jam sekolah tanpa mengganggu jam pelajaran lainnya. Namun sayangnya masih ada peserta didik yang tidak memanfaatkan waktu luang contohnya seperti halnya dengan kegiatan ekstrakurikuler olahraga. Hal ini bisa disebabkan salah satunya karena minat peserta didik terhadap kegiatan ekstrakurikuler yang ada.

2 Tinjauan Literatur (or Literature Review)

2.1 Data Mining

Menurut Turban et al dalam jurnal Kennedy Tampubolon, dkk (2013 : 96), menjelaskan bahwa: “Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar/Data Warehouse”.

2.2 Clustering

Clustering merupakan metode penganalisaan data, yang sering dimasukan sebagai salah satu metode data mining, yang tujuannya adalah untuk mengelompokan data dengan karakteristik yang sama. Tujuan utama dari metode *clustering* adalah pengelompokan sejumlah data /objek ke dalam *cluster* (kelompok) sehingga dalam setiap *cluster* akan berisi data yang semirip mungkin. Metode *clustering* berusaha untuk menempatkan obyek yang mirip (jaraknya dekat) dalam datu kelompok dan membuat jarak antar kelompok sejauh mungkin.

2.3 Algoritma K-means

(Eko Psetyo, 2012) menyatakan bahwa “*K-Means* adalah salah satu metode pengelompokan non hirarki (sekatan) yang berusaha mempartisi data ke dalam *cluster*/kelompok sehingga data yang memiliki karakteristik yang sama akan dimasukkan ke dalam satu *cluster* yang sama dan data yang memiliki karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain”.

2.4 Matlab

Menurut Santosa (2007, h. 1) menyatakan bahwa “*Matlab* adalah suatu bahasa pemrograman tingkat tinggi yang diperuntukkan untuk komputasi teknis. *Matlab* mengintegrasikan aspek komputasi, visualisasi dan dipergunakan untuk aplikasi”.

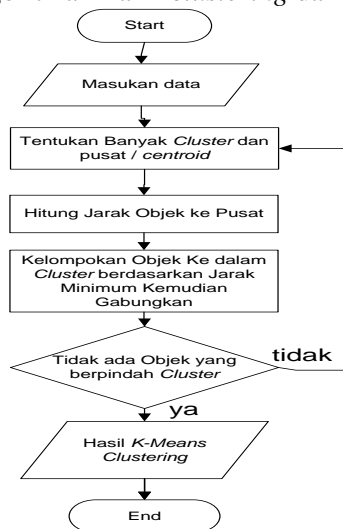
2.5 Pengertian GUI

Menurut Dr. Eng. Ags Naba (2009, h:40), menyatakan ”GUIDE atau GUI builder merupakan sebuah *graphical user interface*(GUI) yang dibangun dengan obyek grafik seperti tombol (*button*), kotak teks, slider, menu dan lain-lain”.

3 Metode Penelitian (or Research Method)

Metodeologi penelitian merujuk pada kidentifikasi variabel yang relevan untuk mengetahui minat dan bakat siswa siswi, seperti preferensi kegiatan ekstrakurikuler, partisipasi sebelumnya, nilai akademik, usia, jenis kelamin, dll. Kumpulkan data tersebut melalui survei, wawancara, atau sumber data yang tersedia. Evaluasi dan periksa kualitas data, identifikasi dan perbaiki nilai yang hilang atau tidak valid Lakukan transformasi data jika diperlukan, misalnya, normalisasi data numerik. Seleksi fitur: Jika ada fitur yang tidak relevan atau redundan, hapus atau kurangi dimensi data.

Adapun *flowcart* algoritma hirarki *clustering* dan *K-means* adalah :



Gamnbar 3.1 Flowchart Algoritma Hirarki *Clustering* Dan *K-means*

Pengelompokan data dengan metode *K-Means* ini secara umum dilakukan dengan cara sebagai berikut :

1. Masukan data
2. Tentukan jumlah kelompok
3. Alokasi data ke dalam kelompok secara acak.
4. Hitung pusat kelompok (centroid/rata-rata) dari data yang ada di masing- masing kelompok.
5. Alokasikan masing-masing data ke centroid/rata-rata terdekat.
6. Kembali ke langkah 3 apabila masih ada data yang berpindah kelompok, atau apabila ada peubahan nilai centroid di atas nilai ambang yang ditentukan, atau apabila perubahan nilai fungsi objektif yang digunakan masih di atas nilai ambang yang ditentukan.

4. Hasil dan Pembahasan (or Results and Analysis)

Tabel III.1 Ekstrakurikuler

Kode	Ekstrakurikuler
1	Atletik
2	Karate
3	Rohis Putra Putri
4	Futsal
5	Seni Tari
6	Bola Volly
7	Pramuka
8	Paskibra
9	Paduan Suara

Tabel III.2 Minat

Kode	Minat
1	Pendidikan Jasmani
2	PAI (Pendidikan Agama Islam)
3	Seni Budaya
4	PKN (Pendidikan Kewarganegaraan)

Tabel III.3 Bakat

Kode	Bakat
1	Olah Raga
2	Belajar Keagamaan
3	Menari/Menyanyi
4	PBB (Peraturan Baris Berbaris)

Tabel III.4 Data yang akan di-Clustering

No	Nama	Ekstrakurikuler	Minat	Bakat
1	Bela Anisa	Atletik	Pendidikan Jasmani	Olah Raga
2	Citra Ramadhani	Karate	Pendidikan Jasmani	Olah Raga
3	Dea Amelia	Futsal	Pendidikan Jasmani	Olah Raga
4	Dilla Yolanda	Seni Tari	Seni Budaya	Menari/ Menyanyi
5	Dwi Salwa Zannah	Rohis Putra Dan Putri	PAI	Belajar Keagamaan
6	Jelita Ariska	Pramuka	PKN	Peraturan Baris-berbaris

7	Laila Asy Syifa	Paskibra	PKN	Peraturan Baris-berbaris
8	Melani Br. Sinulingga	Paduan Suara	Seni Budaya	Menari/ Menyanyi
9	Nadia Meilani	Rohis Putra Dan Putri	PAI	Belajar Keagamaan
10	Nurmah Sitah Amanda Br Harahap	Karate	Pendidikan Jasmani	Olah Raga
11	Saqira Putri Nabila	Futsal	Pendidikan Jasmani	Olah Raga
12	Alvida Novrianti	Seni Tari	Seni Budaya	Menari/ Menyanyi
13	Ankesa Srinaomi	Bola Voly	Pendidikan Jasmani	Olah Raga
14	Bella Aurellia Amanda	Rohis Putra Dan Putri	PAI	Belajar Keagamaan
15	Cantika Ardila	Paskibra	PKN	Peraturan Baris-berbaris
16	Davina Malinda	Paduan Suara	Seni Budaya	Menari/ Menyanyi
17	Dedek Wulandari	Atletik	Pendidikan Jasmani	Olah Raga
18	Devika Artanti	Paskibra	PKN	Peraturan Baris-berbaris
19	Dinda Khairi Lubis	Paduan Suara	Seni Budaya	Menari/ Menyanyi
20	Dwi Andini	Rohis Putra Dan Putri	PAI	Belajar Keagamaan

Tabel III.5 Data yang di Trasformasi

No	Nama	Ekstrakulikuler	Minat	Bakat
1	A	1	1	1
2	B	2	1	1
3	C	4	1	1
4	D	5	3	3
5	E	3	2	2
6	F	7	4	4
7	G	8	4	4
8	H	9	3	3
9	I	3	2	2
10	J	2	1	1

11	K	4	1	1
12	L	5	3	3
13	M	6	1	1
14	N	3	2	2
15	O	8	4	4
16	P	9	3	3
17	Q	1	1	1
18	R	8	4	4
19	S	9	3	3
20	T	3	2	2

Hasil perhitungan iterasi 1 di dapat dari tabel yang sudah di transformasikan kemudian dilakukan perhitungan untuk mendapatkan hasil iterasi 1. Setelah didapat hasil iterasi 1 lalu dimasukan kedalam tabel iterasi 1. Berikut ini adalah tabel hasil iterasi 1 :

Tabel III. 6 Hasil Perhitungan Iterasi 1

No	X	Y	Z	C ₁	C ₂	C ₃	Group
1	1	1	1	4.90	7.35	5.00	1
2	2	1	1	4.12	6.56	4.00	3
3	4	1	1	3.00	5.20	2.00	3
4	5	3	3	0.00	2.45	3.00	1
5	3	2	2	2.45	4.90	3.32	1
6	7	4	4	2.45	0.00	4.36	2
7	8	4	4	3.32	1.00	4.69	2
8	9	3	3	4.00	2.45	4.12	2
9	3	2	2	2.45	4.90	3.32	1
10	2	1	1	4.12	6.56	4.00	3
11	4	1	1	3.00	5.20	2.00	3
12	5	3	3	0.00	2.45	3.00	1
13	6	1	1	3.00	4.36	0.00	3
14	3	2	2	2.45	4.90	3.32	1
15	8	4	4	3.32	1.00	4.69	2

16	9	3	3	4.00	2.45	4.12	2
17	1	1	1	4.90	7.35	5.00	1
18	8	4	4	3.32	1.00	4.69	2
19	9	3	3	4.00	2.45	4.12	2
20	3	2	2	2.45	4.90	3.32	1

Keterangan :

Setelah dilakukan perhitungan menggunakan rumus *cluster* yang ada, maka *group* berdasarkan jarak minimal *Centroid* terdekat adalah:

Group Lama : {0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0}

Group Baru : {1,3,3,1,1,2,2,2,1,3,3,1,3,1,2,2,1}

Terjadi perubahan group dilanjutkan ke iterasi berikut :

Tabel III.7 Hasil Perhitungan Iterasi 2

No	X	Y	Z	C ₁	C ₂	C ₃	Group
1	1	1	1	2.45	8.14	2.60	1
2	2	1	1	1.73	7.26	1.60	3
3	4	1	1	1.73	5.62	0.40	3
4	5	3	3	2.45	3.38	3.16	1
5	3	2	2	0.00	5.73	1.54	1
6	7	4	4	4.90	1.42	5.44	2
7	8	4	4	5.74	0.67	6.11	2
8	9	3	3	6.16	1.08	6.10	2
9	3	2	2	0.00	5.73	1.54	1
10	2	1	1	1.73	7.26	1.60	3
11	4	1	1	1.73	5.62	0.40	3
12	5	3	3	2.45	3.38	3.16	1
13	6	1	1	3.32	4.30	2.40	3
14	3	2	2	0.00	5.73	1.54	1
15	8	4	4	5.74	0.67	6.11	2
16	9	3	3	6.16	1.08	6.10	2
17	1	1	1	2.45	8.14	2.60	1
18	8	4	4	5.74	0.67	6.11	2

19	9	3	3	6.16	1.08	6.10	2
20	3	2	2	0.00	5.73	1.54	1

Keterangan :

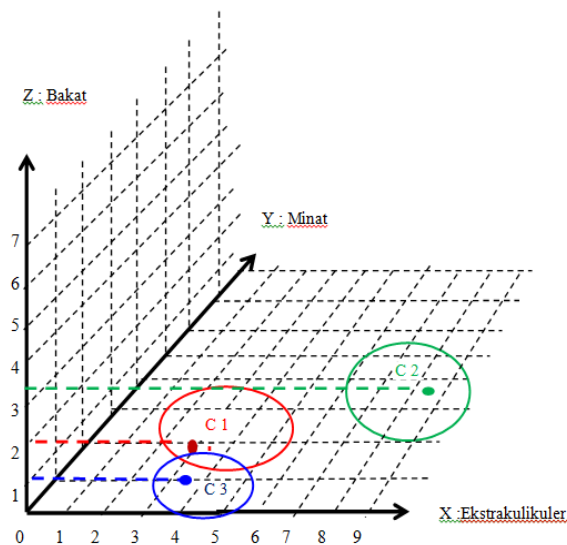
Setelah dilakukan perhitungan menggunakan rumus *cluster* yang ada, pada iterasi 1 terjadi sama seperti iterasi 2 dan Tidak ada data yang berpindah group lagi sehingga perhitungan dapat dihentikan. Sehingga dapat dibuat grafik cluster-nya. Adapun hasil group yang diperoleh dari perhitungan Iterasi 1 dan 2 adalah sebagai berikut :

Group Lama : {1,3,3,1,1,2,2,2,1,3,3,1,3,1,2,2,1,2,2,1}

Group Baru : {1,3,3,1,1,2,2,2,1,3,3,1,3,1,2,2,1,2,2,1}

4.1 Grafik Clustering

Membuat grafik *cluster* berdasarkan perhitungan yang telah dilakukan. Adapun grafik yang diperoleh adalah sebagai berikut :



Gambar III.3 Grafik Cluster berdasarkan perhitungan yang telah dilakukan

- Cluster 1 : 3,00; 2,00; 2,00
- Cluster 2 : 8,29; 3,57; 3,57
- Cluster 3 : 3,60; 1,00; 1,00

Penjelasan Grafik :

Dari 20 data diperoleh 3 *group*, Cluster 1 terdapat 8 data, Cluster 2 terdapat 7 data, dan Cluster 3 terdapat 5 data. Dan diperoleh Group terbanyak adalah *cluster* 1.

1. Cluster 1 Terdapat 8 Data

3,00; 2,00; 2,00

Dapat diketahui pada *cluster* 1 terdapat 8 data yaitu jenis ekstrakurikuler Rohis Putra Putri minatnya terarah pada PAI (Pendidikan Agama Islam) bakat yang terlihat yaitu Belajar Keagamaan.

2. Cluster 2 Terdapat 7 Data

8,29; 3,57; 3,57

Dapat diketahui pada *cluster 2* terdapat 7 data yaitu jenis ekstrakurikuler Paskibra minatnya terarah pada PKN (Pendidikan Kewarganegaraan) bakat yang terlihat yaitu PBB (Peraturan Baris Berbaris).

3. Cluster 3 Terdapat 5 Data

3,60; 1,00; 1,00

Dapat diketahui pada *cluster 3* terdapat 5 data yaitu jenis ekstrakurikuler Futsal minatnya terarah pada Pendidikan Jasmani bakat yang terlihat yaitu Olah Raga.

4 Kesimpulan (or Conclusion)

Penelitian ini bertujuan untuk menerapkan algoritma k-means clustering dalam menganalisis minat dan bakat siswa-siswi terhadap kegiatan ekstrakurikuler di sekolah. Metode k-means clustering digunakan untuk mengelompokkan siswa-siswi berdasarkan pola minat dan bakat yang mereka tunjukkan dalam berbagai kegiatan ekstrakurikuler.

Hasil penelitian menunjukkan bahwa algoritma k-means clustering berhasil mengidentifikasi beberapa kelompok siswa-siswi berdasarkan minat dan bakat mereka. Kelompok-kelompok ini mencerminkan pola minat dan bakat yang berbeda-beda terhadap kegiatan ekstrakurikuler, seperti paskibra, pramuka, seni tari dan lain-lain. Hal ini memberikan wawasan yang berharga bagi sekolah dalam menyusun program ekstrakurikuler yang sesuai dengan minat dan bakat siswa-siswi.

Dalam kesimpulannya, penelitian ini menyoroti manfaat penggunaan algoritma k-means clustering dalam menganalisis minat dan bakat siswa-siswi pada kegiatan ekstrakurikuler. Hasil penelitian ini dapat menjadi dasar bagi sekolah untuk mengembangkan program ekstrakurikuler yang lebih efektif dan sesuai dengan minat serta bakat siswa-siswi, sehingga meningkatkan partisipasi dan kepuasan siswa-siswi dalam kegiatan ekstrakurikuler di sekolah. Namun, perlu dilakukan lebih banyak penelitian untuk memperkuat temuan ini dan mengatasi keterbatasan yang ada.

DAFTAR PUSTAKA

- [1] Ai Ilah Warnilah. 2016. "Analisis Algoritma K-Means Clustering Untuk Pemetaan Prestasi Siswa Studi Kasus SMP N 1 Sukahening." *Jurnal Computer and Information Technology* Vol. 1 No.1 Mei 2016.
- [2] Fani Oktavianti, A.Busyairi. 2019. "Manajemen Peserta Didik Dalam Pengembangan Minat dan Bakat Melalui Kegiatan Ekstrakurikuler" *Joyful Learning Jurnal* Vol. 8 No.4 2019.
- [2] Nurdiana Saputri, Nurrus Sa'adah. 2021. "Pengembangan Minat dan Bakat Peserta Didik Melalui Kegiatan Ekstrakurikuler." *Jurnal Bimbingan Konseling Islam* Vol. 2 No.2 Desember 2021.
- [3] Dr. Eng. Ags Naba, *Pengertian GUIDE atau GUI* Yogyakarta, 2008.
- [4]Eko Prasetyo, *Data Mining Konsep Dan Aplikasi Menggunakan Matlab*, PT.Andi, Yogyakarta, 2012.
- [5] Kusri, Emha Taufiq Luthfi, *Algoritma Data Mining*, Penerbit C.V Andi, Yogyakarta 2009.
- [6] Yatini B, *Definisi Flowchart*, Jakarta PT. Gramedia Widia Sarana Indonesia, 2006
([www.http://modul-guideuploader-by-teuinsuska2009-wordpress-com](http://modul-guideuploader-by-teuinsuska2009-wordpress-com)) diakses pada tanggal : 15 Mei 2023