

Penerapan Data Mining Untuk Penjurusan Kelas dengan Menggunakan Algoritma K-Medoids

Jhiro Faran¹, Rima Tamara Aldisa^{2*}

¹Fakultas Teknologi Komunikasi Dan Informatika, Teknologi Informasi, Universitas Nasional, Jakarta, Indonesia

²Fakultas Teknologi Komunikasi Dan Informatika, Informatika, Universitas Nasional, Jakarta, Indonesia

Email: ¹jhirofaran2022@student.unas.ac.id, ^{2*}rimatamaraa@gmail.com

Email Penulis Korespondensi: rimatamaraa@gmail.com

Submitted: 21/09/2023; Accepted: 30/09/2023; Published: 30/09/2023

Abstrak—Penjurusan kelas dilakukan untuk memfokuskan bagi siswa/I terhadap mata pelajaran yang akan dipelajari pada masa Sekolah Menengah Atas (SMA). Penjurusan kelas pada umumnya dilakukan pada kelas XI, dimana proses penjurusan dilakukan berdasarkan dengan nilai mata pelajaran Bahasa Indonesia, Matematika, Bahasa Inggris, IPA dan IPS sewaktu dikelas X. Proses penjurusan kelas bukan hanya semata dilihat dari nilai IPA dan IPS saja tetapi harus juga diperhatikan terhadap seluruh nilai utama yang digunakan pada proses penjurusan kelas. Hal ini menjadi sebuah masalah dari proses penjurusan kelas, dimana seringkali terjadi kekeliruan terhadap proses penjurusan kelas. Kesalahan terhadap penjurusan kelas yang dilakukan terhadap siswa/I akan berdampak cukup fatal terhadap siswa/I tersebut, selain tidak dapat merubah kelas akan juga menimbulkan efek kemalasan bagi siswa/I karena tidak sesuai dengan kemampuan yang dimiliki oleh siswa/I tersebut. Penyelesaian permasalahan tersebut dapat dilakukan dengan menggunakan sebuah teknik yang bernama data mining. Pada penyelesaian permasalahan ini dilakukan dengan menggunakan cara klusterisasi. Algoritma K-Medoids merupakan algoritma yang digunakan untuk menyelesaikan permasalahan pada penelitian ini. Proses pengelompokan ataupun pembentukan klaster pada algoritma K-Medoids berdasarkan dengan perhitungan jarak terdekat dari setiap objek, perhitungan jarak terdekat berdasarkan dengan penentuan nilai centeroid terlebih dahulu. Algoritma K-Medoids dapat membentuk 2 (dua) klaster sesuai dengan penjurusan kelas yang ada. Hasil yang didapatkan bahawasanya terdapat 3 (tiga) alternatif yang termasuk dalam klaster 1 dan juga 12 (dua belas) alternatif termasuk dalam klaster 2.

Kata Kunci: Penjurusan Kelas; Data Mining; Klaster; Algoritma K-Medoids

Abstract—Class assignments are carried out to focus students on the subjects that will be studied during Senior High School (SMA). Class majors are generally carried out in class of all the main values used in the class majoring process. This is a problem with the class majoring process, where mistakes often occur in the class majoring process. Mistakes regarding class majors made by students will have quite a fatal impact on the student, apart from not being able to change classes, it will also have a laziness effect on the student because it does not match the student's abilities. Solving this problem can be done using a technique called data mining. The solution to this problem is done using clustering. The K-Medoids algorithm is the algorithm used to solve the problems in this research. The process of grouping or forming clusters in the K-Medoids algorithm is based on calculating the closest distance to each object, calculating the closest distance is based on determining the centeroid value first. The K-Medoids algorithm can form 2 (two) clusters according to existing class majors. The results obtained show that there are 3 (three) alternatives included in cluster 1 and also 12 (twelve) alternatives included in cluster 2.

Keywords: Class Major; Data Mining; Cluster; K-Medoids algorithm

1. PENDAHULUAN

Sekolah Menengah Atas (SMA) merupakan tingkatan pendidikan paling tinggi pada program wajib belajar 12 tahun. Pada Sekolah Menengah Atas (SMA) dimulai dari kelas X sampai dengan kelas XII. Pada umumnya Sekolah Menengah Atas (SMA) terdapat penjurusan kelas bagi seluruh siswa/I baik untuk kelas Ilmu Pengetahuan Alam (IPA) ataupun Ilmu Pengetahuan Sosial (IPS)[1]. Penjurusan kelas dilakukan untuk memfokuskan bagi siswa/I terhadap mata pelajaran yang akan dipelajari pada masa Sekolah Menengah Atas (SMA). Tujuan dari pada penjurusan kelas yaitu untuk lebih memudahkan bagi siswa/I untuk memilih terhadap bidang pendidikan pada tingkat perguruan tinggi nantinya. Dimana dengan dilakukan penjurusan kelas pada Sekolah Menengah Atas (SMA) akan memudahkan bagi siswa/I untuk menentukan arah bidang apa yang akan diambil setelah menyelesaikan pendidikan pada Sekolah Menengah Atas (SMA)[2][3].

Penjurusan kelas pada umumnya dilakukan pada kelas XI, dimana proses penjurusan dilakukan berdasarkan dengan nilai mata pelajaran Bahasa Indonesia, Matematika, Bahasa Inggris, IPA dan IPS sewaktu dikelas X. Penjurusan dilakukan dengan melihat dan mempertimbangkan dari nilai – nilai yang didapatkan oleh siswa tersebut. Proses penjurusan kelas bukan hanya semata dilihat dari nilai IPA dan IPS saja tetapi harus juga diperhatikan terhadap seluruh nilai utama yang digunakan pada proses penjurusan kelas[4]. Hal ini menjadi sebuah masalah dari proses penjurusan kelas, dimana seringkali terjadi kekeliruan terhadap proses penjurusan kelas. Dimana proses penjurusan kelas yang dilakukan hanya berdasarkan dengan nilai dari mata pelajaran IPA dan IPS saja tanpa mempertimbangkan nilai dari mata pelajaran lainnya. Hal itu tentu saja merugikan bagi siswa/I yang tidak sesuai dengan penjurusan yang sebenarnya.

Kesalahan terhadap penjurusan kelas yang dilakukan terhadap siswa/I akan berdampak cukup fatal terhadap siswa/I tersebut, selain tidak dapat merubah kelas akan juga menimbulkan efek kemalasan bagi siswa/I karena tidak sesuai dengan kemampuan yang dimiliki oleh siswa/I tersebut. Maka dari itu permasalahan tersebut harus diselesaikan

secara baik agar kiranya tidak menimbulkan dampak yang lebih buruk bagi siswa/I atau penjurusan kelas dikemudian harinya. Penyelesaian permasalahan tersebut dapat dilakukan dengan menggunakan sebuah teknik yang bernama data mining. Data mining merupakan sebuah teknik yang dipergunakan untuk menyelesaikan permasalahan terutama terhadap pengolahan data. Pada data mining pengolahan data dilakukan berdasarkan dengan kumpulan – kumpulan data yang telah tersimpan sebelumnya. Pada kumpulan data tersebut tersimpan sebuah informasi yang berharga bagi pemilik datanya, maka dari itu peran dari data mining adalah untuk mendapatkan kembali informasi yang tersimpan tersebut[5][6].

Pada data mining sendiri terdapat berbagai macam cara untuk proses penyelesaian masalah baik itu estimasi, asosiasi, prediksi, klasifikasi dan klasterisasi. Pada penyelesaian permasalahan ini dilakukan dengan menggunakan cara klasterisasi. Klasterisasi merupakan proses yang dilakukan untuk pengelompokan data menjadi beberapa klaster, dimana pada setiap klaster terdapat kemiripan dari setiap data. Proses pengelompokan berdasarkan kemiripan tersebut dapat dilakukan dengan menggunakan jarak terdekat dari setiap data. Berbagai macam algoritma yang dapat dipergunakan untuk menyelesaikan permasalahan klaster tersebut seperti K-Means, Support Vector Machine dan juga beberapa algoritma lainnya[7], [8].

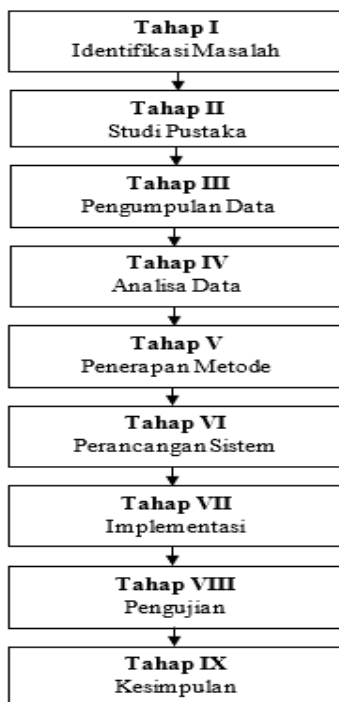
Algoritma K-Medoids merupakan algoritma yang digunakan untuk menyelesaikan permasalahan pada penelitian ini. Proses pengelompokan ataupun pembentukan klaster pada algoritma K-Medoids berdasarkan dengan perhitungan jarak terdekat dari setiap objek, perhitungan jarak terdekat berdasarkan dengan penentuan nilai centroid terlebih dahulu. Pada setiap iterasi yang dilakukan dan juga pada setiap klaster nilai centroid yang digunakan berbeda – beda. Proses pembentukan iterasi yang dilakukan pada algoritma K-Medoids berhenti ketika nilai cost pada iterasi selanjutnya lebih besar dibandingkan dengan nilai cost pada iterasi sebelumnya. Berdasarkan dengan hal tersebut maka dari itu algoritma K-Medoids lebih layak untuk digunakan menyelesaikan permasalahan pada penjurusan kelas tersebut[9], [10]. Beberapa penelitian terdahulu yang mendukung terhadap penyelesaian dengan algoritma K-Medoids seperti yang dilakukan oleh Sukma Sindi, dkk pada tahun 2020 dengan judul penelitian “Analisis Algoritma K-Medoids Clustering Dalam Pengelompokan Penyebaran Covid-19 Di Indonesia” didapatkan hasil bahwasannya Algoritma K-Medoids dapat dipergunakan untuk membantu pembentukan ataupun pengelompokan penyebaran covid di Indonesia[11]. Penelitian lainnya yang dilakukan oleh Ivana Indriani Putri Damanik, dkk pada tahun 2019 dengan judul penelitian yang dilakukan “Algoritma K-Medoids untuk Mengelompokkan Desa yang Memiliki Fasilitas Sekolah di Indonesia” dimana hasil penelitian yang didapatkan bahwasannya dari proses yang dilakukan terdapat 3 klaster yang terbentuk, dimana setiap klaster terdapat jumlah provinsi yang berbeda – beda[12]. Selain itu, penelitian lainnya yang dilakukan oleh Anggi Ayu Dwi Sulistyawati dan Mujiono Sadikin pada tahun 2021 dengan judul penelitian “Penerapan Algoritma K-Medoids Untuk Menentukan Segmentasi Pelanggan” dengan hasil penelitian yang didapatkan bahwasannya setelah dilakukan pengujian berdasarkan dengan nilai Silhouette Index dan Davies Boulbin Index bahwasannya terdapat 3 klaster yang ideal dapat dipergunakan untuk pengelompokan[13]. Terdapat penelitian lainnya yang dilakukan oleh Novita Lestari Anggreini pada tahun 2019 dengan judul penelitian “Teknik Clustering Dengan Algoritma K-Medoids Untuk Menangani Strategi Promosi Di Politeknik TEDC Bandung” didapatkan hasil penelitian bahwasannya proses strategi promosi dilakukan berdasarkan dengan asal daerah terbanyak dari para mahasiswa/I[14]. Penelitian terakhir yang digunakan sebagai referensi yang dilakukan oleh Insanul Kamila, dkk pada tahun 2019 dengan judul penelitian “Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau” dengan hasil penelitian bahwasannya dari pengujian yang dilakukan nilai Davies Boulbin Index lebih besar dibandingkan dengan nilai Davies Boulbin Index pada algoritma K-Means. Hal tersebut menandakan bahwasannya algoritma K-Medoids lebih baik dibandingkan dengan algoritma K-Means[15].

Dari penjelasan serta penjabaran yang telah dilakukan, maka dari itu tujuan yang akan dilakukan pada penelitian ini untuk menyelesaikan permasalahan terhadap penjurusan kelas. Proses yang dilakukan untuk membantu menyelesaikan masalah tersebut yaitu dengan menggunakan algoritma K-Medoids. Hasil akhir yang ingin dicapai nantinya yaitu proses penjurusan kelas bagi siswa/I tidak lagi terjadi kekeliruan yang merugikan bagi siswa/I tersebut.

2. METODOLOGI PENELITIAN

2.1 Kerangka Kerja Penelitian

Kerangka kerja penelitian merupakan sebuah alur proses tahapan yang dilakukan pada penelitian. Pada kerangka kerja penelitian dapat terlihat jelas setiap tahapan yang dilalui pada penelitian dimulai dari tahapan awal hingga tahapan akhir kesimpulan pada penelitian. Dengan adanya kerangka kerja penelitian ini, penelitian yang dilakukan nantinya dapat tersusun serta terjadwal dengan baik. Adapun gambaran kerangka kerja penelitian dapat dilihat berikut:



Gambar 1. Kerangka Kerja Penelitian

2.2 Data Mining

Data mining merupakan bahasa lain dari Knowledge Discovery in Database (KDD). Pada data mining dilakukan sebuah proses penggalian data dari kumpulan – kumpulan data dalam jangka waktu yang lama dan juga dalam kapasitas yang besar. Pada data mining dilakukan proses pengolahan terhadap keseluruhan data yang tersimpan. Pengolahan data yang dilakukan pada data mining untuk menemukan kembali sebuah informasi – informasi penting yang tersimpan pada data hingga dapat dipergunakan kembali bagi pemilik data [16], [17].

2.3 Algoritma K-Medoids

Algoritma K-Medoids salah satu algoritma yang menjadi bagian klusterisasi pada data mining. Algoritma K-Medoids dipergunakan untuk melakukan proses pengelompokan atau kluster dari data, dengan menggunakan algoritma K-Medoids nantinya data dapat disusun ataupun dibagi berdasarkan dengan berbagai kelompok kluster tertentu. Proses pengelompokan pada algoritma K-Medoids berdasarkan dengan perhitungan nilai jarak terdekat dari setiap data, dengan terlebih dahulu diketahui nilai centeroid dari setiap kluster. Proses dari pada algoritma K-Medoids nantinya akan membentuk Iterasi – Iterasi, iterasi yang terbentuk nantinya akan berhenti jika nilai cost yang didapatkan pada iterasi selanjutnya lebih besar dari pada nilai cost dari iterasi sebelumnya. Adapun rumus untuk mencari nilai jarak dapat dilihat pada berikut [18]–[20]:

$$d_{ij} = \sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2 + \dots + (x_{ki} - x_{kj})^2} \quad (1)$$

Adapun tahapan penyelesaian yang dilakukan oleh algoritma K-Medoids dapat dilihat pada berikut ini :

1. Menentukan banyaknya cluster (k) yang akan dibentuk
2. Menghitung jarak antar objek dengan menggunakan persamaan (1)
3. Menentukan nilai centroid baru untuk proses perhitungan jarak dan pembentukan cluster
4. Hitung nilai rata – rata ataupun cost dari proses cluster
5. Jika nilai cost baru lebih besar dari cost lama, maka proses telah berakhir.

3. HASIL DAN PEMBAHASAN

Penjurusan kelas merupakan sebuah proses yang sangat penting dilakukan pada tingkat Sekolah Menengah Atas (SMA). Proses penjurusan kelas dilakukan untuk mengelompokkan para siswa/I apakah masih kedalam kelas Ilmu Pengetahuan Alam (IPA) ataupun Ilmu Pengetahuan Sosial (IPS). Penjurusan kelas dilakukan untuk lebih memfokuskan mata pelajaran kepada siswa/I terhadap calon bidang pendidikan yang akan diambil pada tingkat perguruan tinggi nantinya. Maka dari itu proses penjurusan kelas harus dilakukan dengan sebaik mungkin tanpa terdapat kesalahan. Kesalahan dalam proses penentuan penjurusan kelas tentunya akan berdampak cukup fatal bagi siswa, selain membuat siswa/I menjadi tidak semangat tetapi juga tidak dapat memaksimalkan kemampuan yang dimiliki oleh siswa/I tersebut. Penjurusan kelas yang dilakukan tidak semata berdasarkan dengan nilai pelajaran IPA

dan IPS tetapi juga mempertimbangkan nilai mata pelajaran inti lainnya seperti Matematika, Bahasa Indonesia dan juga Bahasa Inggris. Nilai pelajaran IPA merupakan rata – rata nilai yang didapatkan dari mata pelajaran Kimia, Fisika dan Biologi dan juga nilai pelajaran IPS merupakan rata – rata nilai yang didapatkan dari mata pelajaran Sosiologi, Sejarah dan Geografi. Nilai – nilai mata pelajaran tersebut merupakan nilai mata pelajaran yang didapatkan oleh para siswa/I ketika kelas X. Maka dari itu perlu dilakukan penyelesaian permasalahan dengan menggunakan sebuah cara khusus, salah satu cara yang dapat digunakan yaitu dengan memanfaatkan teknik data mining. Data mining sendiri merupakan sebuah cara yang dilakukan pada kumpulan data untuk mendapatkan kembali informasi dan juga pola yang tersimpan pada kumpulan data tersebut. Proses pengolahan data yang dilakukan pada data mining tidak lepas dari penggunaan algoritma didalamnya. Algoritma K-Medoids merupakan salah satu algoritma yang dapat digunakan pada penyelesaian data mining. Dimana pada algoritma K-Medoids dapat menyelesaikan permasalahan pengelompokan, pengelompokan yang dilakukan pada algoritma K-Medoids berdasarkan dengan perhitungan jarak terdekat dari setiap objek pada data.

3.1 Penerapan Algoritma K-Medoids

Langkah awal dalam proses penerapan algoritma K-Medoids yaitu kumpulan data yang akan digunakan. Berdasarkan hal tersebut maka terlebih dahulu untuk mengetahui data yang akan digunakan dalam proses penyelesaian masalah. Maka dari itu, data yang tersedia merupakan data nilai mata pelajaran siswa/I. Adapun data tersebut dapat dilihat pada tabel 1 berikut:

Tabel 1. Data Nilai Siswa/I

No	Alternatif	Bahasa Indonesia	Bahasa Inggris	Matematika	IPA	IPS
1	A1	80	90	80	80	85
2	A2	85	90	85	85	90
3	A3	90	90	85	85	80
4	A4	90	85	80	85	85
5	A5	95	80	90	80	90
6	A6	75	90	90	85	85
7	A7	80	95	75	90	80
8	A8	85	90	80	85	80
9	A9	85	90	85	80	85
10	A10	90	95	80	80	90
11	A11	85	90	80	90	85
12	A12	80	95	90	85	80
13	A13	80	85	85	80	80
14	A14	90	80	95	85	85
15	A15	95	80	90	80	80

Dari tabel 1 dapat dilihat nilai – nilai yang didapatkan dari setiap siswa/I, nilai – nilai tersebut nantinya yang akan diproses untuk proses pembentuka klaster pada algoritma K-Medoids. Proses yang dilakukan pada algoritma K-Medoids berdasakan dengan beberapa iterasi, adapun proses tersebut dapat dilihat berikut:

a. Iterasi I

Iterasi I merupakan proses iterasi awal yang dilakukan pada algoritma K-Medoids, sebelum dilakukan proses pada iterasi ini terlebih dahulu diketahui nilai centroid awal, adapun nilai centroid awal dapat dilihat pada tabel 2 berikut ini:

Tabel 2. Nilai Centroid Awal Iterasi I

No	Alternatif	Bahasa Indonesia	Bahasa Inggris	Matematika	IPA	IPS	Cluster
1	A8	85	90	80	85	80	C1
2	A12	80	95	90	85	80	C2

Setelah diketahui nilai centroid yang akan digunakan pada proses iterasi, maka tahapan selanjutnya melakukan proses perhitungan jarak. Adapun tahapan perhitungan tersebut dapat dilihat berikut:

Perhitungan Jarak Cluster 1

$$\begin{aligned}
 & d(a1,c1) \\
 &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(80 - 85)^2 + (90 - 90)^2 + (80 - 80)^2 + (80 - 85)^2 + (85 - 80)^2} \\
 &= 12,25 \\
 & d(a2,c1) \\
 &= \sqrt{(85 - 85)^2 + (90 - 90)^2 + (85 - 80)^2 + (85 - 85)^2 + (90 - 80)^2} \\
 &= 12,25 \\
 & d(a3,c1)
 \end{aligned}$$



$$= \sqrt{(90 - 85)^2 + (90 - 90)^2 + (85 - 80)^2 + (85 - 85)^2 + (80 - 80)^2}$$

$$= 16,58$$

d(a4,c1)

$$= \sqrt{(90 - 85)^2 + (85 - 90)^2 + (80 - 80)^2 + (85 - 85)^2 + (85 - 80)^2}$$

$$= 18,71$$

d(a5,c1)

$$= \sqrt{(95 - 85)^2 + (80 - 90)^2 + (90 - 80)^2 + (80 - 85)^2 + (90 - 80)^2}$$

$$= 23,45$$

d(a6,c1)

$$= \sqrt{(75 - 85)^2 + (90 - 90)^2 + (90 - 80)^2 + (85 - 85)^2 + (85 - 80)^2}$$

$$= 0$$

d(a7,c1)

$$= \sqrt{(90 - 85)^2 + (95 - 90)^2 + (75 - 80)^2 + (90 - 85)^2 + (80 - 80)^2}$$

$$= 18,03$$

d(a8,c1)

$$= \sqrt{(85 - 85)^2 + (90 - 90)^2 + (80 - 80)^2 + (85 - 85)^2 + (80 - 80)^2}$$

$$= 15,00$$

d(a9,c1)

$$= \sqrt{(85 - 85)^2 + (90 - 90)^2 + (85 - 80)^2 + (80 - 85)^2 + (85 - 80)^2}$$

$$= 12,25$$

d(a10,c1)

$$= \sqrt{(90 - 85)^2 + (95 - 90)^2 + (80 - 80)^2 + (80 - 85)^2 + (90 - 80)^2}$$

$$= 20,00$$

d(a11,c1)

$$= \sqrt{(85 - 85)^2 + (90 - 90)^2 + (80 - 80)^2 + (90 - 85)^2 + (85 - 80)^2}$$

$$= 15,00$$

d(a12,c1)

$$= \sqrt{(80 - 85)^2 + (95 - 90)^2 + (90 - 80)^2 + (85 - 85)^2 + (80 - 80)^2}$$

$$= 8,66$$

d(a13,c1)

$$= \sqrt{(80 - 85)^2 + (85 - 90)^2 + (85 - 80)^2 + (80 - 85)^2 + (80 - 80)^2}$$

$$= 11,18$$

d(a14,c1)

$$= \sqrt{(90 - 85)^2 + (80 - 90)^2 + (95 - 80)^2 + (85 - 85)^2 + (85 - 80)^2}$$

$$= 18,71$$

d(a15,c1)

$$= \sqrt{(95 - 85)^2 + (80 - 90)^2 + (90 - 80)^2 + (80 - 85)^2 + (80 - 80)^2}$$

$$= 23,45$$

Perhitungan Jarak Cluster 2

d(a1,c2)

$$= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2}$$

$$= \sqrt{(80 - 80)^2 + (90 - 95)^2 + (80 - 90)^2 + (80 - 85)^2 + (85 - 80)^2}$$

$$= 10,00$$

d(a2,c2)

$$= \sqrt{(85 - 80)^2 + (90 - 95)^2 + (85 - 90)^2 + (85 - 85)^2 + (90 - 80)^2}$$

$$= 0$$

d(a3,c2)

$$= \sqrt{(90 - 80)^2 + (90 - 95)^2 + (85 - 90)^2 + (85 - 85)^2 + (80 - 80)^2}$$

$$= 11,18$$

d(a4,c2)

$$= \sqrt{(90 - 80)^2 + (85 - 95)^2 + (80 - 90)^2 + (85 - 85)^2 + (85 - 80)^2}$$

$$= 10,00$$

d(a5,c2)

$$= \sqrt{(95 - 80)^2 + (80 - 95)^2 + (90 - 90)^2 + (80 - 85)^2 + (90 - 80)^2}$$

$$= 15,81$$

d(a6,c2)



$$= \sqrt{(75 - 80)^2 + (90 - 95)^2 + (90 - 90)^2 + (85 - 85)^2 + (85 - 80)^2}$$

$$= 12,25$$

d(a7,c2)

$$= \sqrt{(80 - 80)^2 + (95 - 95)^2 + (75 - 90)^2 + (90 - 85)^2 + (80 - 80)^2}$$

$$= 16,58$$

d(a8,c2)

$$= \sqrt{(85 - 80)^2 + (90 - 95)^2 + (80 - 90)^2 + (85 - 85)^2 + (80 - 80)^2}$$

$$= 11,18$$

d(a9,c2)

$$= \sqrt{(85 - 80)^2 + (90 - 95)^2 + (85 - 90)^2 + (80 - 85)^2 + (85 - 80)^2}$$

$$= 7,07$$

d(a10,c2)

$$= \sqrt{(90 - 80)^2 + (95 - 95)^2 + (80 - 90)^2 + (80 - 85)^2 + (90 - 80)^2}$$

$$= 10,00$$

d(a11,c2)

$$= \sqrt{(85 - 80)^2 + (90 - 95)^2 + (80 - 90)^2 + (90 - 85)^2 + (85 - 80)^2}$$

$$= 8,66$$

d(a12,c2)

$$= \sqrt{(80 - 80)^2 + (95 - 95)^2 + (90 - 90)^2 + (85 - 85)^2 + (80 - 80)^2}$$

$$= 13,23$$

d(a13,c2)

$$= \sqrt{(80 - 80)^2 + (85 - 95)^2 + (85 - 90)^2 + (80 - 85)^2 + (80 - 80)^2}$$

$$= 13,23$$

d(a14,c2)

$$= \sqrt{(90 - 80)^2 + (80 - 95)^2 + (95 - 90)^2 + (85 - 85)^2 + (85 - 80)^2}$$

$$= 15,81$$

d(a15,c2)

$$= \sqrt{(95 - 80)^2 + (80 - 95)^2 + (90 - 90)^2 + (80 - 85)^2 + (80 - 80)^2}$$

$$= 18,71$$

Adapun hasil perhitungan jarak dan penentuan cluster dapat dilihat berikut

Tabel 3. Hasil Perhitungan Jarak Iterasi I

Alternatif	dc1	dc2	Cluster
A1	8,66	13,23	C1
A2	11,18	13,23	C1
A3	7,07	12,25	C1
A4	8,66	18,03	C1
A5	20,62	23,98	C1
A6	15,00	8,66	C2
A7	10,00	15,81	C1
A8	0,00	12,25	C1
A9	8,66	11,18	C1
A10	13,23	18,03	C1
A11	7,07	14,14	C1
A12	12,25	0,00	C2
A13	10,00	12,25	C1
A14	19,36	19,36	C1
A15	18,03	21,79	C1

Proses penentuan cluster berdasarkan nilai jarak terdekat jika $dc1 < dc2$ maka hasil pembentukan cluster yaitu C1. Sebaliknya jika $dc2 < dc1$ hasil pembentukan cluster yaitu C2. Setelah didapatkan proses pembentukan cluster berikutnya yaitu proses perhitungan nilai cost berdasarkan dengan jarak cluster yang terpilih. Adapun perhitungan nilai cost dapat dilihat berikut:

$$\text{Cost Iterasi I} = 8,66 + 11,18 + 7,07 + 8,66 + 20,62 + 8,66 + 10,00 + 0,00 + 8,66 + 13,23 + 7,07 + 0,00 + 10,00 + 19,36 + 18,03 = \mathbf{151,20}$$

Setelah dilakukan proses perhitungan nilai cost, maka dapat dilakukan proses untuk iterasi selanjutnya

b. Iterasi II

Proses yang dilakukan pada iterasi II sama dengan yang dilakukan pada proses iterasi I. Dimana langkah awal yaitu menentukan nilai centroid awal. Adapun nilai centroid awal pada iterasi II dapat dilihat pada tabel 4 berikut:

Tabel 4. Nilai Centroid Awal Iterasi II

No	Alternatif	Bahasa Indonesia	Bahasa Inggris	Matematika	IPA	IPS	Cluster
1	A6	75	90	90	85	85	C1
2	A2	85	90	85	85	90	C2

Setelah didapatkan nilai centroid awal maka dapat dihitung nilai jarak antar tiap data berdasarkan dengan perhitungan rumus jarak terdekat. Adapun proses perhitungan jarak dapat dilihat berikut:

Perhitungan Jarak Cluster 1

$d(a1,c1)$

$$= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2}$$

$$= \sqrt{(80 - 75)^2 + (90 - 90)^2 + (80 - 90)^2 + (80 - 85)^2 + (85 - 85)^2}$$

$$= 12,25$$

$d(a2,c1)$

$$= \sqrt{(85 - 75)^2 + (90 - 90)^2 + (85 - 90)^2 + (85 - 85)^2 + (90 - 85)^2}$$

$$= 12,25$$

$d(a3,c1)$

$$= \sqrt{(90 - 75)^2 + (90 - 90)^2 + (85 - 90)^2 + (85 - 85)^2 + (80 - 85)^2}$$

$$= 16,58$$

$d(a4,c1)$

$$= \sqrt{(90 - 75)^2 + (85 - 90)^2 + (80 - 90)^2 + (85 - 85)^2 + (85 - 85)^2}$$

$$= 18,71$$

$d(a5,c1)$

$$= \sqrt{(95 - 75)^2 + (80 - 90)^2 + (90 - 90)^2 + (80 - 85)^2 + (90 - 85)^2}$$

$$= 23,45$$

$d(a6,c1)$

$$= \sqrt{(75 - 75)^2 + (90 - 90)^2 + (90 - 90)^2 + (85 - 85)^2 + (85 - 85)^2}$$

$$= 0$$

$d(a7,c1)$

$$= \sqrt{(90 - 75)^2 + (95 - 90)^2 + (75 - 90)^2 + (90 - 85)^2 + (80 - 85)^2}$$

$$= 18,03$$

$d(a8,c1)$

$$= \sqrt{(85 - 75)^2 + (90 - 90)^2 + (80 - 90)^2 + (85 - 85)^2 + (80 - 85)^2}$$

$$= 15,00$$

$d(a9,c1)$

$$= \sqrt{(85 - 75)^2 + (90 - 90)^2 + (85 - 90)^2 + (80 - 85)^2 + (85 - 85)^2}$$

$$= 12,25$$

$d(a10,c1)$

$$= \sqrt{(90 - 75)^2 + (95 - 90)^2 + (80 - 90)^2 + (80 - 85)^2 + (90 - 85)^2}$$

$$= 20,00$$

$d(a11,c1)$

$$= \sqrt{(85 - 75)^2 + (90 - 90)^2 + (80 - 90)^2 + (90 - 85)^2 + (85 - 85)^2}$$

$$= 15,00$$

$d(a12,c1)$

$$= \sqrt{(80 - 75)^2 + (95 - 90)^2 + (90 - 90)^2 + (85 - 85)^2 + (80 - 85)^2}$$

$$= 8,66$$

$d(a13,c1)$

$$= \sqrt{(80 - 75)^2 + (85 - 90)^2 + (85 - 90)^2 + (80 - 85)^2 + (80 - 85)^2}$$

$$= 11,18$$

$d(a14,c1)$

$$= \sqrt{(90 - 75)^2 + (80 - 90)^2 + (95 - 90)^2 + (85 - 85)^2 + (85 - 85)^2}$$

$$= 18,71$$

$d(a15,c1)$

$$= \sqrt{(95 - 75)^2 + (80 - 90)^2 + (90 - 90)^2 + (80 - 85)^2 + (80 - 85)^2}$$

$$= 23,45$$

Perhitungan Jarak Cluster 2

$$\begin{aligned}
 & d(a1,c2) \\
 &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(80 - 85)^2 + (90 - 90)^2 + (80 - 85)^2 + (80 - 85)^2 + (85 - 90)^2} \\
 &= 10,00 \\
 & d(a2,c2) \\
 &= \sqrt{(85 - 85)^2 + (90 - 90)^2 + (85 - 85)^2 + (85 - 85)^2 + (90 - 90)^2} \\
 &= 0 \\
 & d(a3,c2) \\
 &= \sqrt{(90 - 85)^2 + (90 - 90)^2 + (85 - 85)^2 + (85 - 85)^2 + (80 - 90)^2} \\
 &= 11,18 \\
 & d(a4,c2) \\
 &= \sqrt{(90 - 85)^2 + (85 - 90)^2 + (80 - 85)^2 + (85 - 85)^2 + (85 - 90)^2} \\
 &= 10,00 \\
 & d(a5,c2) \\
 &= \sqrt{(95 - 85)^2 + (80 - 90)^2 + (90 - 85)^2 + (80 - 85)^2 + (90 - 90)^2} \\
 &= 15,81 \\
 & d(a6,c2) \\
 &= \sqrt{(75 - 85)^2 + (90 - 90)^2 + (90 - 85)^2 + (85 - 85)^2 + (85 - 90)^2} \\
 &= 12,25 \\
 & d(a7,c2) \\
 &= \sqrt{(80 - 85)^2 + (95 - 90)^2 + (75 - 85)^2 + (90 - 85)^2 + (80 - 90)^2} \\
 &= 16,58 \\
 & d(a8,c2) \\
 &= \sqrt{(85 - 85)^2 + (90 - 90)^2 + (80 - 85)^2 + (85 - 85)^2 + (80 - 90)^2} \\
 &= 11,18 \\
 & d(a9,c2) \\
 &= \sqrt{(85 - 85)^2 + (90 - 90)^2 + (85 - 85)^2 + (80 - 85)^2 + (85 - 90)^2} \\
 &= 7,07 \\
 & d(a10,c2) \\
 &= \sqrt{(90 - 85)^2 + (95 - 90)^2 + (80 - 85)^2 + (80 - 85)^2 + (90 - 90)^2} \\
 &= 10,00 \\
 & d(a11,c2) \\
 &= \sqrt{(85 - 85)^2 + (90 - 90)^2 + (80 - 85)^2 + (90 - 85)^2 + (85 - 90)^2} \\
 &= 8,66 \\
 & d(a12,c2) \\
 &= \sqrt{(80 - 85)^2 + (95 - 90)^2 + (90 - 85)^2 + (85 - 85)^2 + (80 - 90)^2} \\
 &= 13,23 \\
 & d(a13,c2) \\
 &= \sqrt{(80 - 85)^2 + (85 - 90)^2 + (85 - 85)^2 + (80 - 85)^2 + (80 - 90)^2} \\
 &= 13,23 \\
 & d(a14,c2) \\
 &= \sqrt{(90 - 85)^2 + (80 - 90)^2 + (95 - 85)^2 + (85 - 85)^2 + (85 - 90)^2} \\
 &= 15,81 \\
 & d(a15,c2) \\
 &= \sqrt{(95 - 85)^2 + (80 - 90)^2 + (90 - 85)^2 + (80 - 85)^2 + (80 - 90)^2} \\
 &= 18,71
 \end{aligned}$$

Setelah dilakukan proses perhitungan jarak, maka selanjutnya yaitu penentuan cluster. Adapun hasil perhitungan jarak dan penentuan cluster dapat dilihat pada tabel 5 berikut:

Tabel 5. Hasil Perhitungan Jarak Iterasi II

Alternatif	dc1	dc2	Cluster
A1	12,25	10,00	C2
A2	12,25	0,00	C2
A3	16,58	11,18	C2



Alternatif	dc1	dc2	Cluster
A4	18,71	10,00	C2
A5	23,45	15,81	C2
A6	0,00	12,25	C1
A7	18,03	16,58	C2
A8	15,00	11,18	C2
A9	12,25	7,07	C2
A10	20,00	10,00	C2
A11	15,00	8,66	C2
A12	8,66	13,23	C1
A13	11,18	13,23	C1
A14	18,71	15,81	C2
A15	23,45	18,71	C2

Proses penentuan cluster berdasarkan nilai jarak terdekat jika $dc1 < dc2$ maka hasil pembentukan cluster yaitu C1. Sebaliknya jika $dc2 < dc1$ hasil pembentukan cluster yaitu C2. Setelah didapatkan proses pembentukan cluster berikutnya yaitu proses perhitungan nilai cost berdasarkan dengan jarak cluster yang terpilih. Adapun perhitungan nilai cost dapat dilihat berikut:

$$\text{Cost Iterasi II} = 10,00 + 0,00 + 11,18 + 10,00 + 15,81 + 0,00 + 16,58 + 11,18 + 7,07 + 10,00 + 8,66 + 8,66 + 11,18 + 15,81 + 18,71 = \mathbf{154,84}$$

Setelah didapatkan nilai cost pada Iterasi II, maka selanjutnya dapat dilakukan perbandingan dari nilai cost lama dengan nilai cost baru sebagai penentuan berakhir atau tidaknya proses iterasi. Berdasarkan proses perhitungan dapat dilihat bahwasannya **nilai cost baru (Iterasi II) 154,84** > dari **nilai cost lama (Iterasi I) 151,20**. Maka dari itu proses telah dihentikan dan selanjutnya dapat diambil keputusan dalam proses pembentukan cluster seperti terlihat pada tabel 6 berikut

Tabel 6. Pembentukan Cluster Prioritas

Cluster 1	Cluster 2
A6	A1
A12	A2
A13	A3
	A4
	A5
	A7
	A8
	A9
	A10
	A11
	A14
	A15

Dari tabel 6 dapat dilihat bahwasannya A6, A12 dan A13 termasuk dalam cluster 1 yang menandakan bahwasannya alternatif tersebut menjadi penjurusan kelas IPA. Sedangkan alternatif A1, A2, A3, A4, A5, A7, A8, A9, A10, A11, A14 dan A5 termasuk dalam cluster 2 dimana alternatif tersebut menjadi penjurusan kelas IPS

4. KESIMPULAN

Hasil akhir dalam sebuah penelitian merupakan kesimpulan, adapun yang menjadi kesimpulan dari penelitian ini yaitu Data mining dapat digunakan untuk menyelesaikan permasalahan terhadap proses penjurusan kelas. Dimana proses yang dilakukan pada data mining dengan melakukan pengolahan terhadap seluruh data. Proses pembentukan klaster dilakukan dengan menggunakan algoritma K-Medoids. Algoritma K-Medoids dapat membentuk 2 (dua) klaster sesuai dengan penjurusan kelas yang ada. Hasil yang didapatkan bahwasannya terdapat 3 (tiga) alternatif yang termasuk dalam klaster 1 dan juga 12 (dua belas) alternatif termasuk dalam klaster 2.

REFERENCES

- [1] V. Anestiviya and A. F. O. Pasaribu, "Analisis Pola Menggunakan Metode C4.5 Untuk Peminatan Jurusan Siswa Berdasarkan Kurikulum (Studi Kasus : Sman 1 Natar)," *J. Teknol. dan Sist. Inf.*, vol. 2, no. 1, pp. 80-85, 2021, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTSI>.
- [2] D. N. Hidayat and A. B. Setiawan, "Sistem Pendukung Keputusan Penjurusan Siswa Berbasis Web Di Smk Intensif Baitussalam Menggunakan K-Means Clustering," in *Seminar Nasional Inovasi Teknologi*, 2022, pp. 221-225, [Online]. Available:



<https://proceeding.unpkediri.ac.id/index.php/inotek/article/view/2488%0Ahttps://proceeding.unpkediri.ac.id/index.php/inotek/article/download/2488/1551>.

- [3] M. Syahril, S. Kusnasari, A. Muhazir, and A. Syahputri, “Implementasi Data Mining Untuk Rekomendasi Jurusan Menggunakan Algoritma K-Means Clustering Muhammad,” *J. Teknol. Sist. Inf. dan Sist. Komput. TGD*, vol. 6, no. 1, pp. 235–245, 2023.
- [4] N. M. Purnamasari, A. Syaqui, and D. A. Pramana, “Pengelompokan Data Calon Siswa Baru Di Sekolah Menengah Kejuruan menggunakan Algoritma K-Means,” *J. Sist. Inf. dan Teknol. Perad.*, vol. 4, no. 1, pp. 24–30, 2023.
- [5] R. Sidik, N. Suarna, and A. R. Dikananda, “ANALISA DATA SET PEMINATAN SISWA MENGGUNAKAN ALGORITMA K- MEANS DENGAN OPTIMIZE PARAMETER DI SEKOLAH MENENGAH KEJURUAN (STUDI KASUS: SMK PUI GEGESIK),” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 2, pp. 1197–1203, 2023.
- [6] W. Kokoh Andriyan, “Penerapan Data Mining Dengan Menggunakan Metode K-Means Clustering Dalam Pengelompokan Data Nilai Pada SMA YKPP PENDOPO Untuk Menentukan Jurusan Ipa Dan Ips,” *J. Jupiter*, vol. 15, no. 1, pp. 452–461, 2023.
- [7] R. T. Aldisa, “Data Mining Penentuan Jurusan Siswa Menggunakan Metode Agglomerative Hierarchical Clustering (AHC),” *J. Media Inform. Budidarma*, vol. 7, no. 2, pp. 873–880, 2023, doi: 10.30865/mib.v7i2.6092.
- [8] T. Syahputra, “Penerapan Data Mining Dalam Pengelompokan Kelas Mahasiswa Berdasarkan Hasil Ujian Saringan Masuk Dengan Algoritma K-Means,” *JURTEKSI (Jurnal Teknol. dan Sist. Informasi)*, vol. 5, no. 2, pp. 161–166, 2019, doi: 10.33330/jurteksi.v5i2.350.
- [9] A. Hasnan and M. Arif, “Penerapan Data Mining Untuk Pengelompokan Kelas Siswa Unggulan Berdasarkan Nilai Raport Menggunakan Metode K- Means Clustering (Studi Kasus : SMK Muhammadiyah 3 Weleri),” *J. Tek. Inform. dan Desain Komun. Vis.*, vol. 1, no. 2, pp. 82–93, 2022.
- [10] Firza and Sarjono, “Penerapan Algoritma K-Means Dalam Metode Clustering Untuk Peminatan Jurusan Bagi Siswa Swasta Pelita Raya Kota Jambi,” *J. Manaj. Sist. Inf.*, vol. 5, no. 3, pp. 371–382, 2020, [Online]. Available: <http://ejournal.stikom-db.ac.id/index.php/manajemensisteminformasi/article/view/907>.
- [11] S. Sindi, W. R. O. Ningse, I. A. Sihombing, F. I. R.H.Zer, and D. Hartama, “Analisis Algoritma K-Medoids Clustering Dalam Pengelompokan Penyebaran Covid-19 Di Indonesia,” *J. Teknol. Inf.*, vol. 4, no. 1, pp. 166–173, 2020, doi: 10.36294/jurti.v4i1.1296.
- [12] I. I. P. Damanik, S. Solikhun, I. S. Saragih, I. Parlina, D. Suhendro, and A. Wanto, “Algoritma K-Medoids untuk Mengelompokkan Desa yang Memiliki Fasilitas Sekolah di Indonesia,” *Pros. Semin. Nas. Ris. Inf. Sci.*, vol. 1, no. September, p. 520, 2019, doi: 10.30645/senaris.v1i0.58.
- [13] A. A. D. Sulistyawati and M. Sadikin, “Penerapan Algoritma K-Medoids Untuk Menentukan Segmentasi Pelanggan,” *Sistemasi*, vol. 10, no. 3, p. 516, 2021, doi: 10.32520/stmsi.v10i3.1332.
- [14] N. L. Anggreini, “Teknik Clustering Dengan Algoritma K-Medoids Untuk Menangani Strategi Promosi Di Politeknik Tedc Bandung,” *J. Teknol. Inf. dan Pendidik.*, vol. 12, no. 2, pp. 1–7, 2019, doi: 10.24036/tip.v12i2.215.
- [15] I. Kamila, U. Khairunnisa, and M. Mustakim, “Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau,” *J. Ilm. Rekayasa dan Manaj. Sist. Inf.*, vol. 5, no. 1, p. 119, 2019, doi: 10.24014/rmsi.v5i1.7381.
- [16] F. Y. Rahman, I. I. Purnomo, and N. Hijriana, “PENERAPAN ALGORITMA DATA MINING UNTUK KLASIFIKASI KUALITAS AIR,” *Technologia*, vol. 13, no. 3, pp. 228–232, 2022.
- [17] S. Ucha Putri, E. Irawan, and F. Rizky, “Implementasi Data Mining Untuk Prediksi Penyakit Diabetes Dengan Algoritma C4.5,” *Januari*, vol. 2, no. 1, pp. 39–46, 2021.
- [18] F. Faisal, L. A. Giopani, M. Fitriah, Z. C. Dwyne, and S. Syahidatul, “Comparison of K-Means and K-Medoids Algorithms for Temperature Grouping in Riau Province Perbandingan Algoritma K-Means dan K-Medoids Untuk Pengelompokan Suhu di Provinsi Riau,” *IJRSE Indones. J. Inform. Res. Softw. Eng.*, vol. 2, no. 2, pp. 128–134, 2022.
- [19] S. D. Nirwana, M. I. Jambak, and A. Bardadi, “Perbandingan Algoritma K-Means Dan K-Medoids Dalam Clustering Rata-Rata Penambahan Kasus Covid-19 Berdasarkan Kota/Kabupaten Di Provinsi Sumatera Selatan,” *JSII (Jurnal Sist. Informasi)*, vol. 9, no. 2, pp. 126–131, 2022, doi: 10.30656/jsii.v9i2.5127.
- [20] R. K. Purba and E. Bu’ulolo, “Implementasi Algoritma K-Medoids dalam Pengelompokan Mahasiswa yang Layak Mendapat Bantuan Uang Kuliah Tunggal,” *INSOLOGI J. Sains dan Teknol.*, vol. 1, no. 2, pp. 79–86, 2022, doi: 10.55123/insologi.v1i2.195.