

Penerapan Data Mining Dengan Menggunakan Algoritma Clustering K-Means Untuk Pembagian Jurusan Pada Sekolah Menengah Atas

Ikbal Danu Setiawan, Agung Triayudi*

Prodi Magister Teknologi Informasi, Fakultas Teknologi Komunikasi dan Informatika, Universitas Nasional, Jakarta, Indonesia

Email: ¹ikbaldanusetiawan2017@student.unas.ac.id, ^{2,*}agungtriayudi@civitas.unas.ac.id

Email Penulis Korespondensi: agungtriayudi@civitas.unas.ac.id

Submitted: 23/02/2024; Accepted: 27/02/2024; Published: 28/02/2024

Abstrak—Sekolah Menengah Atas merupakan jenjang terakhir yang wajib ditempuh sebelum melanjutkan pendidikan pada tingkat lebih tinggi seperti Diploma ataupun Sarjana. Dimana pada umumnya Sekolah Menengah Atas terdapat penjurusan kelas untuk siswa/I yang akan naik ke kelas XI dari kelas X. Pembagian jurusan terhadap siswa/I bertujuan untuk meningkatkan mutu pendidikan dari para siswa/I keberlangsungan proses belajar mengajar yang terjadi didalam kelas. Peningkatan kualitas pendidikan yang dilakukan pada proses penjurusan kelas terdapat dimana para siswa nantinya akan lebih fokus sesuai dengan bidang minat jurusan yang seharusnya diambil oleh para siswa/I. Proses yang terjadi dalam penentuan jurusan hanya berdasarkan dengan keinginan yang disampaikan oleh para siswa/I tanpa memperhitungkan nilai – nilai akademik dari mata pelajaran yang telah dilalui ataupun diselesaikan para siswa/I di kelas X. Permasalahan tersebut bukanlah sebuah masalah kecil yang harus diabaikan, hal tersebut merupakan sebuah permasalahan penting yang harus segera diselesaikan dikarenakan jika permasalahan tersebut tidak diselesaikan dengan segera maka akan menimbulkan dampak yang berkelanjutan nantinya. Proses penentuan dari pembagian jurusan para siswa/I dapat dilihat berdasarkan dengan pola atau nilai – nilai dari para siswa/I sebelumnya. Data mining merupakan sebuah proses yang dipergunakan untuk menyelesaikan proses terhadap data yang besar. Data yang diproses merupakan kumpulan – kumpulan data yang menjadi sebuah Big Data dari data dimasa lampau yang tersimpan pada sebuah wadah penyimpanan dan kemudian dapat dipergunakan kembali dengan melakukan proses pengolahan. Clustering merupakan cara yang sesuai untuk menyelesaikan permasalahan. Dimana pada clustering dilakukan pengelompokan berdasarkan dengan jarak terhadap setiap objek data. Algoritma K-Means merupakan bagian dari Clustering Data Mining, dimana pada algoritma tersebut dapat digunakan untuk melakukan pengelompokan baru berdasarkan dengan cara pembentukan cluster. Dari hasil yang didapatkan bahwasannya terdapat 2 (dua) cluster pembentukan baru. Pada cluster 1 terdapat 9 (sembilan) siswa dan pada cluster 2 terdapat 6 (enam) siswa.

Kata Kunci: Sekolah Menengah Atas; Penjurusan; Kelas; Data Mining; Algoritma K-Means

Abstract—Senior High School is the last level that must be taken before continuing education at a higher level such as a Diploma or Bachelor's degree. Where in general high schools have class majors for students who will move up to class XI from class Improving the quality of education carried out in the class majoring process means that students will be more focused in accordance with the field of interest of the major that the student/I should take. The process that occurs in determining majors is only based on the wishes expressed by the students without taking into account the academic grades of the subjects that the students have passed or completed in class X. This problem is not a small problem that should be ignored, it This is an important problem that must be resolved immediately because if the problem is not resolved immediately it will have lasting impacts later. The process of determining the division of majors for students can be seen based on the patterns or values of previous students. Data mining is a process used to complete processing of large data. The data that is processed is a collection of data that becomes Big Data from past data that is stored in a storage container and can then be reused by processing it. Clustering is an appropriate way to solve problems. Where in clustering grouping is carried out based on the distance to each data object. The K-Means algorithm is part of Clustering Data Mining, where this algorithm can be used to carry out new groupings based on how clusters are formed. From the results obtained, there are 2 (two) new formation clusters. In cluster 1 there are 9 (nine) students and in cluster 2 there are 6 (six) students.

Keywords: Senior High School; Majoring; Class; Data Mining; K-Means Algorithm

1. PENDAHULUAN

Sekolah menengah atas merupakan pendidikan tingkat atas yang wajib ditempuh oleh anak – anak di Indonesia sesuai dengan program kerja yang di instruksikan dari pemerintah yaitu wajib belajar 12 (dua belas) tahun. Sekolah Menengah Atas merupakan jenjang terakhir yang wajib ditempuh sebelum melanjutkan pendidikan pada tingkat lebih tinggi seperti Diploma ataupun Sarjana[1], [2].

Pada Sekolah Menengah Atas ditempuh pada kelas X, XI dan XII. Dimana pada umumnya Sekolah Menengah Atas terdapat penjurusan kelas untuk siswa/I yang akan naik ke kelas XI dari kelas X. Penjurusan kelas bertujuan untuk mengelompokan para siswa/I untuk masuk dalam kelas Ilmu Pengetahuan Alam (IPA) ataupun Ilmu Pengetahuan Sosial (IPS)[3], [4].

Pembagian jurusan terhadap siswa/I bertujuan untuk meningkatkan mutu pendidikan dari para siswa/I keberlangsungan proses belajar mengajar yang terjadi didalam kelas. Dimana proses pembagian jurusan kelas terhadap siswa/I berkenaan langsung ataupun berkaitan langsung dengan para siswa/I yang akan naik tingkat menjadi kelas XI. Penjurusan kelas pada Sekolah Menengah Atas biasanya dilakukan oleh bagian kesiswaan berkolaborasi dengan para wali kelas[5], [6].

Peningkatan kualitas pendidikan yang dilakukan pada proses penjurusan kelas terdapat dimana para siswa nantinya akan lebih fokus sesuai dengan bidang minat jurusan yang seharusnya diambil oleh para siswa/I. Pembagian jurusan yang tidak sesuai dengan kemampuan dari para siswa/I tentu akan sangat merugikan bagi para siswa/I dimana hasil nilai akademik yang didapatkan pada keberlanjutan kelas akan berantakan dan juga tidak sesuai dengan studi lanjut pada perguruan tinggi nantinya[7], [8].

Proses yang terjadi dalam penentuan jurusan hanya berdasarkan dengan keinginan yang disampaikan oleh para siswa/I tanpa memperhitungkan nilai – nilai akademik dari mata pelajaran yang telah dilalui ataupun diselesaikan para siswa/I di kelas X. Penentuan pembagian jurusan tersebut yang memiliki banyak kelemahan dan juga kerugian bukan saja bagi siswa/I tetapi juga bagi para guru yang mengajar nantinya. Dimana seharusnya proses penjurusan kelas berdasarkan dengan nilai akademik yang didapatkan dari para siswa/I.

Permasalahan tersebut bukanlah sebuah masalah kecil yang harus diabaikan, hal tersebut merupakan sebuah permasalahan penting yang harus segera diselesaikan dikarenakan jika permasalahan tersebut tidak diselesaikan dengan segera maka akan menimbulkan dampak yang berkelanjutan nantinya. Dampak yang dihadapi bagi sekolah adalah tidak dapat mengeluarkan potensi akademik secara keseluruhan yang dimiliki oleh para siswa/I.

Proses penentuan dari pembagian jurusan para siswa/I dapat dilihat berdasarkan dengan pola atau nilai – nilai dari para siswa/I sebelumnya. Dari pola tersebut nantinya akan membantu dan memudahkan bagi pihak sekolah untuk melakukan proses penentuan jurusan berdasarkan dengan nilai yang didapatkan dari para siswa/I sebelumnya. Maka dari itu penyelesaian terhadap permasalahan sudah seharusnya mempertimbangkan terhadap sebuah cara. Salah satu cara yang mungkin dapat dipergunakan untuk menyelesaikan permasalahan tersebut dengan menggunakan data mining.

Data mining merupakan sebuah proses yang dipergunakan untuk menyelesaikan proses terhadap data yang besar. Data yang diproses merupakan kumpulan – kumpulan data yang menjadi sebuah Big Data dari data dimasa lampau yang tersimpan pada sebuah wadah penyimpanan dan kemudian dapat dipergunakan kembali dengan melakukan proses pengolahan. Proses pengolahan terhadap data pada data mining bertujuan untuk menemukan sebuah informasi kembali ataupun informasi yang penting tersimpan dari data dan nantinya dapat dipergunakan dalam proses pengambilan keputusan. Pada data mining sendiri hasil informasi yang didapatkan dapat divisualisasi dalam berbagai macam bentuk sesuai dengan informasi seperti apa yang akan diinginkan. Dalam data mining sendiri, proses penyelesaian terbagi atas beberapa kelompok cara dan pada penelitian ini proses penyelesaian berdasarkan dengan clustering[9]–[11].

Pada data mining clustering merupakan sebuah cara yang merupakan bagian dari pada unsupervised learning dimana data yang dilakukan pengolahan merupakan data yang masih belum memiliki tujuan kelasnya. Clustering merupakan cara yang sesuai untuk menyelesaikan permasalahan. Dimana pada clustering dilakukan pengelompokan berdasarkan dengan jarak terhadap setiap objek data[12], [13]. Pengelompokan yang dilakukan pada clustering juga berdasarkan dengan kesamaan karakteristik yang dimiliki oleh setiap data. Maka dari itu, pada penelitian ini clustering lebih tepat untuk menyelesaikan permasalahan tersebut. Clustering pada data mining sendiri terdapat beberapa algoritma yang dapat dipergunakan untuk menyelesaikan permasalahan, salah satunya adalah algoritma K-Means[14], [15].

Algoritma K-Means merupakan bagian dari Clustering Data Mining, dimana pada algoritma tersebut dapat digunakan untuk melakukan pengelompokan baru berdasarkan dengan cara pembentukan cluster. Pada algoritma K-Means proses pembentukan cluster berdasarkan dengan karakteristik dari setiap objek yang terdapat pada data. Karakteristik tersebut dapat dilihat berdasarkan dengan nilai jarak terdekat yang dimiliki pada data. Proses yang dilakukan pada algoritma K-Means berdasarkan dengan proses iterasi yang dilakukan, dimana proses iterasi dilakukan untuk menghitung kedekatan dari setiap jarak antar objek data[16], [17]. Pada proses iterasi yang dilakukan terlebih dahulu ditentukan nilai titik pusat (centeroid) sebagai dasar penentuan kelompok. Dari nilai tersebut nantinya akan dilakukan proses pengolahan data secara keseluruhan. Proses iterasi yang dilakukan pada algoritma K-Means dilakukan berulang dimana akan berhenti jika cluster yang terbentuk sudah tidak terdapat perubahan lagi[18], [19].

Untuk membantu dalam pelaksanaan penelitian, diperlukan beberapa penguatan berdasarkan dengan penelitian terdahulu yang telah dilakukan seperti penelitian dari Preddy Marpaung, dkk pada tahun 2023 dengan judul penelitian yang dilakukan Penerapan Data Mining Untuk Pengelompokan Kepadatan Penduduk Kabupaten Deli Serdang Menggunakan Algoritma K-Means dan hasil yang diapatkan dari penelitian bahwa Dengan hasil penggunaan algoritma K-Means dapat menghasilkan, bahwa dari 22 kecamatan yang ada, terdapat 3 kecamatan Sangat Padat (cluster1), 4 kecamatan padat (cluster2), dan 15 kecamatan sedang (cluster3)[20].

Penelitian lainnya yang juga telah dilakukan oleh Suharmanto, dkk pada tahun 2023 dengan judul penelitian Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Clustering Perokok Usia Lebih dari 15 Tahun hasil yang didapatkan dengan menggunakan Clustering K-Means untuk mengetahui pengelompokan perokok dengan usia lebih dari 15 tahun di setiap daerah pengelompokan perokok dengan rentan usia lebih dari 15 tahun dapat dilihat melalui 3 cluster[21].

Selain itu, juga terdapat penelitian yang dilakukan oleh Davit Ariyanto pada tahun 2022 dimana judul penelitian yaitu Data Mining Menggunakan Algoritma K-Means untuk Klasifikasi Penyakit Infeksi Saluran

Pernafasan Akut serta hasil dari proses penelitian berupa hasil terdapat proses pembentukan cluster yaitu terdapat 3 (tiga) cluster terhadap penyakit ISPA[22].

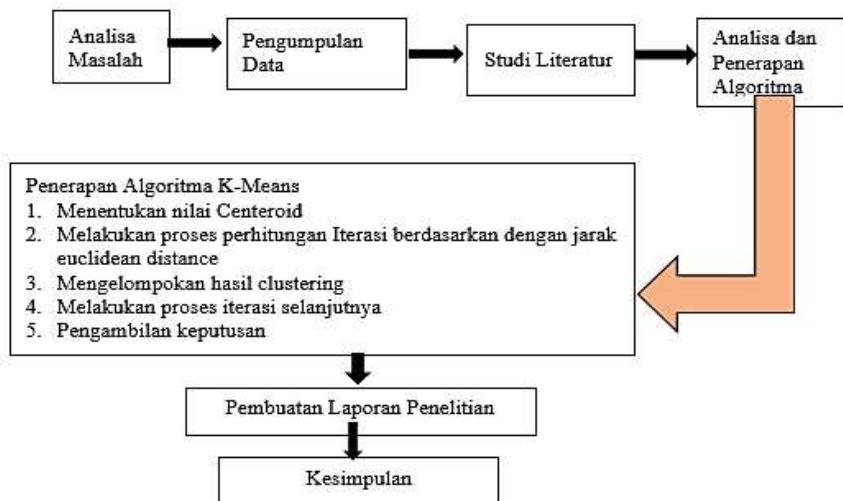
Penelitian terakhir yang digunakan sebagai acuan penelitian yang dilakukan oleh Rafki Maulidi pada tahun 2022 dengan judul penelitian Data Mining Menggunakan Algoritma K-Means Clustering dalam Analisis Tingkat Potongan Harga Terhadap Harga Jual Sepeda Motor Honda dan hasil dari penelitian hasil dari pengujian terhadap algoritma ini adalah 3 cluster dengan 42 data uji, pada cluster 1 terdapat 34 tipe kendaraan mendapat potongan harga, kemudian cluster 2 sebanyak 7 tipe kendaraan boleh dapat potongan harga dan cluster 3 sebanyak 1 tipe kendaraan tidak dapat potongan harga[23].

Berdasarkan dengan penjabaran diatas, maka dari itu pada penelitian ini akan dilakukan sebuah proses penelitian untuk melakukan proses penjurusan terhadap para siswa/I pada Sekolah Menengah Atas. Penjurusan kelas dilakukan dengan mempertimbangkan nilai akademik yang didapatkan dari para siswa/I di kelas X. Tujuan dari penelitian agar kiranya tidak lagi terdapat kekeliruan terhadap proses penjurusan para siswa/I yang memiliki dampak cukup panjang bagi keberlangsungan masa depan dari para siswa/I

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian merupakan suatu kegiatan ilmiah yang bertujuan untuk memperoleh data yang kemudian digunakan untuk kepentingan tertentu. Seseorang yang melakukan penelitian disebut juga dengan istilah peneliti. Penelitian yang termasuk kegiatan ilmiah, maka dalam melakukan kegiatan ini, harus dilakukan dengan cara-cara yang sistematis, rasional atau masuk akal, dan data-data yang valid atau sesuai dengan fakta. Supaya penelitian tidak dilakukan secara asal-asalan, maka bagi peneliti harus menggunakan yang namanya metodologi penelitian. Metodologi penelitian adalah suatu cara atau teknik untuk mendapatkan informasi dan sumber data yang akan digunakan dalam penelitian. Dengan adanya tahapan penelitian ini penelitian yang dilakukan dapat terstruktur dengan baik dan tidak ada proses yang dilakukan berulang. Informasi atau data ini bisa dalam bentuk apa saja, literatur, seperti jurnal, artikel, tesis, buku, koran, dan sebagainya. Bahkan sumber data bisa juga diperoleh dari survei atau wawancara. Dalam melakukan penelitian ini, ada beberapa tahap yang dipakai agar mendapatkan data yang sangat diperlukan sebagai berikut:



Gambar 1. Tahapan Penelitian

Berdasarkan dengan gambar 1, maka dapat diketahui terhadap penjelasan secara ringkas dari setiap tahapan penelitian yang dilakukan:

1. Analisa Masalah

Analisa masalah digunakan untuk memecahkan suatu masalah penelitian serta menganalisa data dalam melakukan suatu kejadian sebelum melakukan perancangan ataupun perhitungan.

2. Pengumpulan data

Pada tahap ini dilakukan pengumpulan data-data yang diperlukan sebagai bahan untuk memecahkan masalah yang telah dirumuskan. Pengumpulan data dilakukan dengan observasi secara langsung dan wawancara

3. Studi Literatur Digunakan untuk mempelajari serta menambah pemahaman penelitian tentang dalam menganalisa secara khusus serta membaca jurnal-jurnal ataupun referensi lainnya yang terkait dengan penelitian .

4. Analisa dan Penerapan Metode dalam penelitian ini yaitu dimulai dengan menganalisa permasalahan yang terjadi dalam proses pemilihan jurusan siswa/I pada Sekolah Menengah Atas. Dimulai dari menganalisa

sistem lama, setelah itu dilanjutkan dengan menganalisa dengan menggunakan cara baru dan menerapkan algoritma K-Means.

5. Laporan Penelitian Pada tahap ini membuat laporan dari keseluruhan penelitian ini, untuk melihat apakah hasil dari penelitian ini sudah sesuai dengan yang diharapkan, dan dilanjut dengan membuat suatu kesimpulan dari penelitian tersebut.

2.2 Data Mining

Data mining merupakan proses pengolahan data yang banyak digunakan pada saat ini. Dimasa sekarang ini peran data mining sangatlah krusial, hal tersebut dikarenakan perannya yang cukup penting dalam proses pengambilan keputusan. Peran data mining dalam pengambilan keputusan berdasarkan dengan pengolahan data yang dilakukan[24], [25]. Pengolahan data yang dilakukan pada data mining merupakan data yang besar dari kumpulan – kumpulan data sebelumnya. Kumpulan data tersebut nantinya dilakukan penggalian kembali untuk mendapatkan informasi yang berharga dari kumpulan data tersebut. Informasi tersebut nantinya dapat membantu dalam proses pengambilan keputusan yang dilakukan. Pada data mining sendiri tidak berjalan dengan sendirian, dimana pada data mining terdapat berbagai macam cara penyelesaian seperti Asosiasi, Estimasi, Prediksi, Klasifikasi dan Klasterisasi[8]–[10].

2.3 Algoritma K-Means

K-Means adalah salah satu algoritma clustering yang sangat popular karenakesederhanaan dan kemampuannya dalam menangani data dengan skala besar. Namun demikian algoritma ini sangat sensitif terhadap centroid awal. Perbedaan centroid awal akan memberikan perbedaan hasil clustering dan apabila centroid awal yang diberikan adalah centroid yang tidak baik maka dapat dipastikan hasil clusteringnya juga tidak baik [26][27][28].

Dengan metode pengelompokan berdasarkan jarak, teknik K-means yang dianggap sebagai metode pengelompokan klasik. Salah satu fitur utama yang membuat K-means berguna dalam penelitian ini adalah kemungkinan menentukan terlebih dahulu berapa banyak cluster yang sedang dicari. Jumlah ini mengacu pada parameter K. Dari jumlah ini, poin K dipilih dalam cluster random sebagai pusat cluster. Semua variabel ditugaskan ke pusat cluster terdekat mereka sesuai dengan metrik jarak Euclidean. Dengan rumus jarak Euclidean Distance sebagai berikut[29]–[31]:

$$d(x, y) = |x - y| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

3. HASIL DAN PEMBAHASAN

Penjuruan kelas bagi siswa kelas XI Sekolah Menengah Atas sangatlah penting, hal tersebut dikarenakan proses penjurusan kelas merupakan salah satu bagian cara untuk meningkatkan kualitas pendidikan. Dimana dengan penjurusan kelas bagi para siswa/I akan sangat membantu untuk meningkatkan minat belajar yang dimiliki oleh siswa hingga dapat memaksimalkan proses pembelajaran yang dilakukan. Selama ini proses penjurusan kelas berdasarkan dengan keinginan dari para siswa/I tanpa mempertimbangkan kemampuan ataupun nilai akademik mata pelajaran yang didapat oleh para siswa/I. Kesalahan dalam penjurusan kelas tentu akan memiliki dampak yang cukup besar bukan saja bagi para siswa/I tetapi juga bagi para guru. Dampak yang paling nyata dihadapi adalah ketidaksesuaian jurusan dengan kemampuan siswa/I maka tidak dapat memaksimalkan kemampuan yang dimiliki ataupun juga proses kelanjutan studi pada tingkat perguruan tinggi. Permasalahan tersebut dapat diselesaikan dengan menemukan pola berdasarkan dengan data – data sebelumnya. Maka dari itu, data mining merupakan sebuah cara yang tepat dipergunakan untuk menyelesaikan permasalahan tersebut. Dimana data mining merupakan sebuah cara yang dilakukan untuk melakukan penggalian data hingga didapatkan informasi – informasi baru. Salah satu cara pada data mining yang dapat dipergunakan yaitu clustering. Clustering pada data mining berutujuan untuk membentuk kelompok baru berdasarkan dengan karakteristik data ataupun kedekatan dari setiap objek data. Salah satu algoritma yang dapat dipergunakan untuk menyelesaikan permasalahan adalah algoritma K-Means. Algoritma K-Means bagian clustering data mining yang bertujuan untuk membentuk data baru berdasarkan dengan euclidean distance, pada algoritma K-Means dilakukan perhitungan jarak terdekat berdasarkan dengan nilai centeroid awal dan juga data yang tersedia sebagai dataset.

3.1 Penerapan Algoritma K-Means

Setelah dilakukan proses pendahuluan secara matang, maka selanjutnya dilakukan proses penyelesaian permasalahan dengan menggunakan algoritma K-Means. Sebelum dilakukan proses penyelesaian tersebut terlebih dahulu diketahui terhadap data yang akan digunakan dalam proses penyelesaian. Data yang digunakan dapat dilihat pada tabel 1 berikut.

Tabel 1. Data Nilai Siswa/I

Siswa	NMP1	NMP2	NMP3	NMP4	NMP5
Siswa 1	98	78	92	70	84

Siswa	NMP1	NMP2	NMP3	NMP4	NMP5
Siswa 2	98	90	74	88	71
Siswa 3	80	99	97	77	89
Siswa 4	94	94	87	70	78
Siswa 5	88	88	70	78	75
Siswa 6	82	91	81	71	70
Siswa 7	94	96	84	71	71
Siswa 8	78	78	97	94	92
Siswa 9	89	83	91	86	90
Siswa 10	78	99	89	72	82
Siswa 11	98	99	84	93	78
Siswa 12	85	90	72	83	72
Siswa 13	93	94	97	99	92
Siswa 14	74	84	98	79	91
Siswa 15	98	77	81	95	79

Pada tabel 1 tersebut dapat dilihat bahwasannya terdapat 15 (lima belas) nilai data siswa/I selain itu juga terdapat masing – masing nilai mata pelajaran dari siswa, dimana pada data tersebut terdapat 5 (lima) nilai mata pelajaran siswa. Selanjutnya setelah diketahui terhadap data yang akan digunakan pada penelitian maka selanjutnya sudah dapat dilakukan proses penyelesaian permasalahan dengan menggunakan algoritma K-Means. Adapun proses penyelesaian algoritma K-Means dapat dilihat berikut:

Iterasi – 1

1. Menentukan jumlah clusters; Dalam hal ini jumlah cluster yang ditentukan dalam penelitian yaitu 2 (dua) dimana cluster tersebut merupakan pembagian terhadap kelas Ilmu Pengetahuan Alam (IPA) dan Ilmu Pengetahuan Sosial (IPS).
2. Langkah selanjutnya yaitu menentukan nilai centroid awal; Nilai centroid awal berdasarkan dengan data set awal nilai siswa. Dimana pada hal ini ditentukan 2 (dua) nilai centroid awal untuk mewakili Cluster 1 dan Cluster 2. Disini untuk nilai centroid awal diambil data Siswa 7 sebagai Cluster 1 dan Siswa 11 sebagai Cluster 11. Adapun untuk nilai centroid awal dapat dilihat pada tabel 2 berikut:

Tabel 2. Nilai Centeroid Awal

Siswa	NMP1	NMP2	NMP3	NMP4	NMP5
Siswa 7	94	96	84	71	71
Siswa 11	98	99	84	93	78

3. Langkah ketiga setelah diketahui nilai centroid awal yaitu menghitung nilai kedekatan jarak dari centroid awal terhadap jarak yang terdapat pada data. Dimana perhitungan tersebut dilakukan dengan menggunakan rumus euclidean distance:

$$d(x, y) = |x - y| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Cluster – 1

$$\begin{aligned} d(x_1, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(98 - 94)^2 + (78 - 96)^2 + (92 - 84)^2 + (70 - 71)^2 + (84 - 71)^2} = 23,96 \end{aligned}$$

$$\begin{aligned} d(x_2, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(98 - 94)^2 + (90 - 96)^2 + (74 - 84)^2 + (88 - 71)^2 + (71 - 71)^2} = 21 \end{aligned}$$

$$\begin{aligned} d(x_3, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(80 - 94)^2 + (99 - 96)^2 + (97 - 84)^2 + (77 - 71)^2 + (89 - 71)^2} = 27,09 \end{aligned}$$

$$\begin{aligned} d(x_4, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(94 - 94)^2 + (94 - 96)^2 + (87 - 84)^2 + (70 - 71)^2 + (78 - 71)^2} = 7,94 \end{aligned}$$

$$\begin{aligned} d(x_5, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(88 - 94)^2 + (88 - 96)^2 + (70 - 84)^2 + (78 - 71)^2 + (75 - 71)^2} = 19 \end{aligned}$$

$$\begin{aligned} d(x_6, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(81 - 94)^2 + (91 - 96)^2 + (81 - 84)^2 + (71 - 71)^2 + (70 - 71)^2} = 13,38 \end{aligned}$$

$$\begin{aligned}
 d(x7, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(94 - 94)^2 + (96 - 96)^2 + (84 - 84)^2 + (71 - 71)^2 + (71 - 71)^2} = 0 \\
 d(x8, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(78 - 94)^2 + (78 - 96)^2 + (97 - 84)^2 + (94 - 71)^2 + (92 - 71)^2} = 41,46 \\
 d(x9, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(89 - 94)^2 + (83 - 96)^2 + (91 - 84)^2 + (86 - 71)^2 + (90 - 71)^2} = 28,79 \\
 d(x10, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(78 - 94)^2 + (99 - 96)^2 + (89 - 84)^2 + (72 - 71)^2 + (82 - 71)^2} = 20,30 \\
 d(x11, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(98 - 94)^2 + (99 - 96)^2 + (84 - 84)^2 + (93 - 71)^2 + (78 - 71)^2} = 23,62 \\
 d(x12, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(85 - 94)^2 + (90 - 96)^2 + (72 - 84)^2 + (83 - 71)^2 + (72 - 71)^2} = 20,15 \\
 d(x13, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(93 - 94)^2 + (94 - 96)^2 + (97 - 84)^2 + (99 - 71)^2 + (92 - 71)^2} = 37,40 \\
 d(x14, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(74 - 94)^2 + (84 - 96)^2 + (98 - 84)^2 + (79 - 71)^2 + (91 - 71)^2} = 34,70 \\
 d(x15, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(98 - 94)^2 + (77 - 96)^2 + (81 - 84)^2 + (95 - 71)^2 + (79 - 71)^2} = 32,03
 \end{aligned}$$

Cluster – 2

$$\begin{aligned}
 d(x1, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(98 - 98)^2 + (78 - 99)^2 + (92 - 84)^2 + (70 - 93)^2 + (84 - 78)^2} = 31,71 \\
 d(x2, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(98 - 98)^2 + (90 - 99)^2 + (74 - 84)^2 + (88 - 93)^2 + (71 - 78)^2} = 15,97 \\
 d(x3, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(80 - 98)^2 + (99 - 99)^2 + (97 - 84)^2 + (77 - 93)^2 + (89 - 78)^2} = 29,50 \\
 d(x4, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(94 - 98)^2 + (94 - 99)^2 + (87 - 84)^2 + (70 - 93)^2 + (78 - 78)^2} = 24,06 \\
 d(x5, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(88 - 98)^2 + (88 - 99)^2 + (70 - 84)^2 + (78 - 93)^2 + (75 - 78)^2} = 25,51 \\
 d(x6, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(82 - 98)^2 + (91 - 99)^2 + (81 - 84)^2 + (71 - 93)^2 + (70 - 78)^2} = 29,61 \\
 d(x7, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(94 - 98)^2 + (96 - 99)^2 + (84 - 84)^2 + (71 - 93)^2 + (71 - 78)^2} = 23,62 \\
 d(x8, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(78 - 98)^2 + (78 - 99)^2 + (97 - 84)^2 + (94 - 93)^2 + (92 - 78)^2} = 34,74
 \end{aligned}$$

Journal of Computer System and Informatics (JoSYC)

ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)

Volume 5, No. 2, February 2024, Page 380-392

<https://ejurnal.seminar-id.com/index.php/josyc>

DOI 10.47065/josyc.v5i2.4970

$$\begin{aligned}
 d(x9, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(89 - 98)^2 + (83 - 99)^2 + (91 - 84)^2 + (86 - 93)^2 + (90 - 78)^2} = 24,06 \\
 d(x10, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(78 - 98)^2 + (99 - 99)^2 + (89 - 84)^2 + (72 - 93)^2 + (82 - 78)^2} = 29,70 \\
 d(x11, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(98 - 98)^2 + (99 - 99)^2 + (84 - 84)^2 + (93 - 93)^2 + (78 - 78)^2} = 0 \\
 d(x12, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(85 - 98)^2 + (90 - 99)^2 + (72 - 84)^2 + (83 - 93)^2 + (72 - 78)^2} = 23,02 \\
 d(x13, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(93 - 98)^2 + (94 - 99)^2 + (97 - 84)^2 + (99 - 93)^2 + (92 - 78)^2} = 21,24 \\
 d(x14, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(74 - 98)^2 + (84 - 99)^2 + (98 - 84)^2 + (79 - 93)^2 + (91 - 78)^2} = 36,91 \\
 d(x15, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(98 - 98)^2 + (77 - 99)^2 + (81 - 84)^2 + (95 - 93)^2 + (79 - 78)^2} = 22,32
 \end{aligned}$$

Setelah proses perhitungan jarak terdekat dengan menggunakan ecludian distance, selanjutnya maka hasil perhitungan disusun kembali dan kemudian memilih untuk jarak terdekat terdapat pada cluster mana. Adapun hasil tersebut dapat dilihat pada tabel 3 berikut:

Tabel 3. Hasil Perhitungan Jarak Iterasi 1

Siswa	NMP1	NMP2	NMP3	NMP4	NMP5	DC1	DC2	Hasil Cluster
Siswa 1	98	78	92	70	84	23,96	32,71	C1
Siswa 2	98	90	74	88	71	21,00	15,97	C2
Siswa 3	80	99	97	77	89	27,09	29,50	C1
Siswa 4	94	94	87	70	78	7,94	24,06	C1
Siswa 5	88	88	70	78	75	19,00	25,51	C1
Siswa 6	82	91	81	71	70	13,38	29,61	C1
Siswa 7	94	96	84	71	71	0,00	23,62	C1
Siswa 8	78	78	97	94	92	41,46	34,74	C2
Siswa 9	89	83	91	86	90	28,79	24,06	C2
Siswa 10	78	99	89	72	82	20,30	29,70	C1
Siswa 11	98	99	84	93	78	23,62	0,00	C2
Siswa 12	85	90	72	83	72	20,15	23,02	C1
Siswa 13	93	94	97	99	92	37,40	21,24	C2
Siswa 14	74	84	98	79	91	34,70	36,91	C1
Siswa 15	98	77	81	95	79	32,03	22,32	C2

Iterasi – 2

Proses selanjutnya yang dilakukan yaitu melakukan proses perhitungan pada iterasi – 2, samal halnya dengan iterasi – 1 tetapi pada iterasi – 2 nilai centeroid didapatkan berdasarkan dengan nilai rata – rata yang didapatkan dari hasil pembentukan cluster sebelumnya. Nilai centeroid cluster 1 terlihat pada tabel 4 berikut ini.

Tabel 4. Nilai Rata – Rata Cluster 1

Siswa	NMP1	NMP2	NMP3	NMP4	NMP5
Siswa 1	98	78	92	70	84
Siswa 3	80	99	97	77	89
Siswa 4	94	94	87	70	78
Siswa 5	88	88	70	78	75
Siswa 6	82	91	81	71	70
Siswa 7	94	96	84	71	71
Siswa 10	78	99	89	72	82

Siswa	NMP1	NMP2	NMP3	NMP4	NMP5
Siswa 12	85	90	72	83	72
Siswa 14	74	84	98	79	91
Nilai Rata - Rata	85,89	91,00	85,56	74,56	79,11

Selanjutnya, setelah didapatkan nilai rata – rata cluster 1 maka juga dilakukan perhitungan nilai rata – rata cluster 2. Adapun nilai rata – rata dapat dilihat pada tabel 5 berikut.

Tabel 5. Nilai Rata – Rata Cluster 2

Siswa	NMP1	NMP2	NMP3	NMP4	NMP5
2	98	90	74	88	71
8	78	78	97	94	92
9	89	83	91	86	90
11	98	99	84	93	78
13	93	94	97	99	92
15	98	77	81	95	79
Nilai Rata - Rata	92,33	86,83	87,33	92,50	83,67

Pada tabel 5 dapat dilihat nilai rata – rata untuk cluster 2, setelah didapatkan nilai rata untuk cluster 1 dan cluster 2 maka nilai rata – rata tersebut dipergunakan sebagai nilai centroid baru pada iterasi 2. Adapun nilai centroid baru dapat dilihat pada tabel 6 berikut.

Tabel 6. Nilai Centeroid Baru

Cluster	NMP1	NMP2	NMP3	NMP4	NMP5
C1	85,89	91,00	85,56	74,56	79,11
C2	92,33	86,83	87,33	92,50	83,67

Setelah didapatkan nilai centroid baru seperti pada tabel 6 diatas, maka selanjut dilakukan proses yang sama dengan menghitung kembali jarak antara objek dengan nilai centroid. Adapun proses tersebut dapat dilihat berikut ini:

- Menghitung jarak objek berdasarkan dengan rumus euclidean distance

Cluster – 1

$$\begin{aligned} d(x_1, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(98 - 85,89)^2 + (78 - 91)^2 + (92 - 85,56)^2 + (70 - 74,56)^2 + (84 - 79,11)^2} = 20,05 \end{aligned}$$

$$\begin{aligned} d(x_2, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(98 - 85,89)^2 + (90 - 91)^2 + (74 - 85,56)^2 + (88 - 74,56)^2 + (71 - 79,11)^2} = 22,97 \end{aligned}$$

$$\begin{aligned} d(x_3, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(80 - 85,89)^2 + (99 - 91)^2 + (97 - 85,56)^2 + (77 - 74,56)^2 + (89 - 79,11)^2} = 18,26 \end{aligned}$$

$$\begin{aligned} d(x_4, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(94 - 85,89)^2 + (94 - 91)^2 + (87 - 85,56)^2 + (70 - 74,56)^2 + (78 - 79,11)^2} = 9,94 \end{aligned}$$

$$\begin{aligned} d(x_5, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(88 - 85,89)^2 + (88 - 91)^2 + (70 - 85,56)^2 + (78 - 74,56)^2 + (75 - 79,11)^2} = 16,86 \end{aligned}$$

$$\begin{aligned} d(x_6, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(82 - 85,89)^2 + (91 - 91)^2 + (81 - 85,56)^2 + (71 - 74,56)^2 + (70 - 79,11)^2} = 11,47 \end{aligned}$$

$$\begin{aligned} d(x_7, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(94 - 85,89)^2 + (96 - 91)^2 + (84 - 85,56)^2 + (71 - 74,56)^2 + (71 - 79,11)^2} = 13,10 \end{aligned}$$

$$\begin{aligned} d(x_8, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(78 - 85,89)^2 + (78 - 91)^2 + (97 - 85,56)^2 + (94 - 74,56)^2 + (92 - 79,11)^2} = 30,11 \end{aligned}$$

$$d(x_9, c1) = \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2}$$

$$= \sqrt{(89 - 85,89)^2 + (83 - 91)^2 + (91 - 85,56)^2 + (86 - 74,56)^2 + (90 - 79,11)^2} = 18,78$$

$$\begin{aligned} d(x10, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(78 - 85,89)^2 + (99 - 91)^2 + (89 - 85,56)^2 + (72 - 74,56)^2 + (82 - 79,11)^2} = 12,37 \end{aligned}$$

$$\begin{aligned} d(x11, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(98 - 85,89)^2 + (99 - 91)^2 + (84 - 85,56)^2 + (93 - 74,56)^2 + (78 - 79,11)^2} = 23,55 \end{aligned}$$

$$\begin{aligned} d(x12, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(85 - 85,89)^2 + (90 - 91)^2 + (72 - 85,56)^2 + (83 - 74,56)^2 + (72 - 79,11)^2} = 17,53 \end{aligned}$$

$$\begin{aligned} d(x13, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(93 - 85,89)^2 + (94 - 91)^2 + (97 - 85,56)^2 + (99 - 74,56)^2 + (92 - 79,11)^2} = 30,89 \end{aligned}$$

$$\begin{aligned} d(x14, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(74 - 85,89)^2 + (84 - 91)^2 + (98 - 85,56)^2 + (79 - 74,56)^2 + (91 - 79,11)^2} = 22,50 \end{aligned}$$

$$\begin{aligned} d(x15, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(98 - 85,89)^2 + (77 - 91)^2 + (81 - 85,56)^2 + (95 - 74,56)^2 + (79 - 79,11)^2} = 27,95 \end{aligned}$$

Cluster - 2

$$\begin{aligned} d(x1, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(98 - 92,33)^2 + (78 - 86,83)^2 + (92 - 87,33)^2 + (72 - 92,50)^2 + (84 - 83,67)^2} = 25,26 \end{aligned}$$

$$\begin{aligned} d(x2, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(98 - 92,33)^2 + (90 - 86,83)^2 + (74 - 87,33)^2 + (88 - 92,50)^2 + (71 - 83,67)^2} = 20,02 \end{aligned}$$

$$\begin{aligned} d(x3, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(80 - 92,33)^2 + (99 - 86,83)^2 + (97 - 87,33)^2 + (77 - 92,50)^2 + (89 - 83,67)^2} = 25,73 \end{aligned}$$

$$\begin{aligned} d(x4, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(94 - 92,33)^2 + (94 - 86,83)^2 + (87 - 87,33)^2 + (70 - 92,50)^2 + (78 - 83,67)^2} = 24,34 \end{aligned}$$

$$\begin{aligned} d(x5, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(88 - 92,33)^2 + (88 - 86,83)^2 + (70 - 87,33)^2 + (78 - 92,50)^2 + (75 - 83,67)^2} = 24,62 \end{aligned}$$

$$\begin{aligned} d(x6, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(82 - 92,33)^2 + (91 - 86,83)^2 + (81 - 87,33)^2 + (71 - 92,50)^2 + (70 - 83,67)^2} = 28,52 \end{aligned}$$

$$\begin{aligned} d(x7, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(94 - 92,33)^2 + (96 - 86,83)^2 + (84 - 87,33)^2 + (71 - 92,50)^2 + (71 - 83,67)^2} = 26,84 \end{aligned}$$

$$\begin{aligned} d(x8, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(78 - 92,33)^2 + (78 - 86,83)^2 + (97 - 87,33)^2 + (94 - 92,50)^2 + (92 - 83,67)^2} = 21,18 \end{aligned}$$

$$\begin{aligned} d(x9, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(89 - 92,33)^2 + (83 - 86,83)^2 + (91 - 87,33)^2 + (86 - 92,50)^2 + (90 - 83,67)^2} = 11,03 \end{aligned}$$

$$\begin{aligned} d(x10, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(78 - 92,33)^2 + (99 - 86,83)^2 + (89 - 87,33)^2 + (72 - 92,50)^2 + (82 - 83,67)^2} = 27,92 \end{aligned}$$

$$d(x11, c2) = \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2}$$

Journal of Computer System and Informatics (JoSYC)

ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)

Volume 5, No. 2, February 2024, Page 380-392

<https://ejurnal.seminar-id.com/index.php/josyc>

DOI 10.47065/josyc.v5i2.4970

$$= \sqrt{(98 - 92,33)^2 + (99 - 86,83)^2 + (84 - 87,33)^2 + (93 - 92,50)^2 + (78 - 83,67)^2} = 14,95$$

$$\begin{aligned} d(x12, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(85 - 92,33)^2 + (90 - 86,83)^2 + (72 - 87,33)^2 + (83 - 92,50)^2 + (72 - 83,67)^2} = 22,92 \end{aligned}$$

$$\begin{aligned} d(x13, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(93 - 92,33)^2 + (94 - 86,83)^2 + (97 - 87,33)^2 + (99 - 92,50)^2 + (92 - 83,67)^2} = 16,03 \end{aligned}$$

$$\begin{aligned} d(x14, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(74 - 92,33)^2 + (84 - 86,83)^2 + (98 - 87,33)^2 + (79 - 92,50)^2 + (91 - 83,67)^2} = 26,34 \end{aligned}$$

$$\begin{aligned} d(x15, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(98 - 92,33)^2 + (77 - 86,83)^2 + (81 - 87,33)^2 + (95 - 92,50)^2 + (79 - 83,67)^2} = 14,03 \end{aligned}$$

Setelah proses perhitungan jarak terdekat dengan menggunakan eculidean distance, selanjutnya maka hasil perhitungan disusun kembali dan kemudian memilih untuk jarak terdekat terdapat pada cluster mana. Adapun hasil tersebut dapat dilihat pada tabel 7 berikut.

Tabel 7. Hasil Perhitungan Jarak Iterasi 2

Siswa	NMP1	NMP2	NMP3	NMP4	NMP5	DC1	DC2	Hasil Cluster
Siswa 1	98	78	92	70	84	20,05	25,26	C1
Siswa 2	98	90	74	88	71	22,97	20,02	C2
Siswa 3	80	99	97	77	89	18,26	25,73	C1
Siswa 4	94	94	87	70	78	9,94	24,34	C1
Siswa 5	88	88	70	78	75	16,86	24,62	C1
Siswa 6	82	91	81	71	70	11,47	28,52	C1
Siswa 7	94	96	84	71	71	13,10	26,84	C1
Siswa 8	78	78	97	94	92	30,11	21,18	C2
Siswa 9	89	83	91	86	90	18,78	11,03	C2
Siswa 10	78	99	89	72	82	12,37	27,92	C1
Siswa 11	98	99	84	93	78	23,55	14,95	C2
Siswa 12	85	90	72	83	72	17,53	22,92	C1
Siswa 13	93	94	97	99	92	30,89	16,03	C2
Siswa 14	74	84	98	79	91	22,50	26,34	C1
Siswa 15	98	77	81	95	79	27,95	14,03	C2

Setelah dilakukan proses hasil perhitungan jarak pada iterasi 2, selanjutnya melakukan perbandingan hasil antara iterasi 1 dan hasil pada iterasi 2. Apakah terdapat perubahan hasil cluster atau tidak. Adapun perbandingan hasil dapat dilihat berikut pada tabel 8.

Tabel 8. Hasil Perbandingan Cluster Iterasi 1 dan Iterasi 2

Siswa	NMP1	NMP2	NMP3	NMP4	NMP5	Hasil Iterasi 1	Hasil Iterasi 2
Siswa 1	98	78	92	70	84	C1	C1
Siswa 2	98	90	74	88	71	C2	C2
Siswa 3	80	99	97	77	89	C1	C1
Siswa 4	94	94	87	70	78	C1	C1
Siswa 5	88	88	70	78	75	C1	C1
Siswa 6	82	91	81	71	70	C1	C1
Siswa 7	94	96	84	71	71	C1	C1
Siswa 8	78	78	97	94	92	C2	C2
Siswa 9	89	83	91	86	90	C2	C2
Siswa 10	78	99	89	72	82	C1	C1
Siswa 11	98	99	84	93	78	C2	C2
Siswa 12	85	90	72	83	72	C1	C1
Siswa 13	93	94	97	99	92	C2	C2
Siswa 14	74	84	98	79	91	C1	C1
Siswa 15	98	77	81	95	79	C2	C2

Pada tabel 8 diatas dapat dilihat hasil perbandingan iterasi, dapat dilihat dari hasil tersebut bahwasannya tidak terdapat perubahan hasil cluster. Maka dari itu proses iterasi dihentikan dan hasil cluster dapat dibagi menjadi kelompok berikut:

Tabel 9. Hasil Pengelompokan Cluster Siswa

Hasil Cluster 1	Hasil Cluster 2
Siswa 1	Siswa 2
Siswa 3	Siswa 8
Siswa 4	Siswa 9
Siswa 5	Siswa 11
Siswa 6	Siswa 13
Siswa 7	Siswa 15
Siswa 10	
Siswa 12	
Siswa 14	

Pada tabel 9 dapat dilihat hasil terhadap pembentukan cluster dimana pada cluster 1 terdapat 9 (sembilan) siswa yaitu Siswa 1, Siswa 3, Siswa 4, Siswa 5, Siswa 6, Siswa 7, Siswa 10, Siswa 12 dan Siswa 14. Pada cluster 2 terdapat 6 (enam) siswa yaitu Siswa 2, Siswa 8, Siswa 9, Siswa 11, Siswa 13 dan Siswa 15.

4. KESIMPULAN

Tahapan akhir dari sebuah penelitian adalah sebuah penarikan sebuah kesimpulan. Adapun yang menjadi kesimpulan dari penelitian bahwa data mining dapat dipergunakan untuk membantu penyelesaian terhadap pengolahan data dimasa lampau hingga didapatkan informasi baru. Dalam proses pembagian penjurusan siswa/I di Sekolah Menengah Atas dilakukan proses pengelompokan baru berdasarkan dengan proses clustering. Hasil yang didapatkan dari pengelompokan baru yaitu algoritma K-Means berhasil melakukan pembentukan cluster dimana pada cluster 1 terdapat 9 (sembilan) siswa yaitu Siswa 1, Siswa 3, Siswa 4, Siswa 5, Siswa 6, Siswa 7, Siswa 10, Siswa 12 dan Siswa 14. Pada cluster 2 terdapat 6 (enam) siswa yaitu Siswa 2, Siswa 8, Siswa 9, Siswa 11, Siswa 13 dan Siswa 15.

REFERENCES

- [1] T. Hidayat, "Klasifikasi Data Jamaah Umroh Menggunakan Metode K-Means Clustering," *J. Sistim Inf. dan Teknol.*, vol. 4, pp. 19–24, 2022, doi: 10.37034/jsisfotek.v4i1.115.
- [2] P. S. Hasugian and J. Raphita Sagala, "Penerapan Data Mining Untuk Pengelompokan Siswa Berdasarkan Nilai Akademik dengan Algoritma K-Means," *KLICK Kaji. Ilm. Inform. dan Komput.*, vol. 3, no. 3, pp. 262–268, 2022, [Online]. Available: <https://djournals.com/klik>.
- [3] D. Zakiyah, N. Merlina, and N. A. Mayangky, "Penerapan Algoritma K-Means Clustering Untuk Mengetahui Kemampuan Karyawan IT," *Comput. Sci.*, vol. 2, no. 1, pp. 59–67, 2022, doi: 10.31294/coscience.v2i1.623.
- [4] M. Y. Matdoan, U. A. Matdoan, and M. Saleh Far-Far, "Algoritma K-Means Untuk Klasifikasi Provinsi di Indonesia Berdasarkan Paket Pelayanan Stunting," *PANRITA J. Sci. Technol. Arts*, vol. 1, no. 2, pp. 41–46, 2022, [Online]. Available: <https://journal.dedikasi.org/pjsta>.
- [5] E. Widodo and W. Hadikristanto, "Pengelompokan Untuk Penjualan Obat Dengan Menggunakan Algoritma K-Means," *Bull. Inf. Technol.* ..., vol. 4, no. 2, pp. 408–413, 2023, [Online]. Available: <https://journal.fkpt.org/index.php/BIT/article/view/848%0Ahttps://journal.fkpt.org/index.php/BIT/article/download/848/468>.
- [6] H. L. Sari and I. Y. Beti, "Penerapan Data Mining Dalam Pengelompokan Buku Yang Dipinjam Menggunakan Algoritma K-Means," *KLICK Kaji. Ilm. Inform. dan Komput.*, vol. 3, no. 6, pp. 925–933, 2023, doi: 10.30865/klik.v3i6.826.
- [7] 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.
- [8] J. Faran and R. T. Aldisa, "Penerapan Data Mining Untuk Penjurusan Kelas dengan Menggunakan Algoritma K-Medoids," *Build. Informatics, Technol. Sci.*, vol. 5, no. 2, pp. 543–552, 2023, doi: 10.47065/bits.v5i2.4313.
- [9] N. M. Purnamasari, A. Syauqi, 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.
- [10] 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.
- [11] 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.
- [12] V. Darsono, Amroni, and A. Andrianti, “Penerapan Data Mining Algoritma K-Means Untuk Rekomendasi Pemilihan Bidang Studi Perguruan Tinggi Pada Siswa SMKN 1 Kota Jambi,” *J. Inform. Dan Rekayasa Komputer(JAKAKOM)*, vol. 2, no. 2, pp. 161–171, 2022, doi: 10.33998/jakakom.2022.2.2.80.
- [13] A. Supriatna, W. Dharmawan, C. Juliane, S. Tinggi, M. Informatika, and & Komputer -Likmi, “Algoritma K-Means Clustering Pada Pengelompokan Minat Bakat Siswa SMK PGRI 2 Karawang,” *J. Tek. Inform. dan Sist. Inf.*, vol. 10, No. 1, no. 1, pp. 38–50, 2023, [Online]. Available: <http://jurnal.mdp.ac.id>.
- [14] Q. A. A'yuniyah and M. Reza, “Penerapan Algoritma K-Nearest Neighbor Untuk Klasifikasi Jurusan Siswa Di Sma Negeri 15 Pekanbaru,” *Indones. J. Inform. Res. Softw. Eng.*, vol. 3, no. 1, pp. 39–45, 2023, doi: 10.57152/ijirse.v3i1.484.
- [15] J. Hutagalung, “Pemetaan Siswa Kelas Unggulan Menggunakan Algoritma K-Means Clustering,” *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 1, pp. 606–620, 2022, doi: 10.35957/jatisi.v9i1.1516.
- [16] S. Ningsih, “Penerapan Algoritma K-Means Clustering Untuk Menentukan Kelas Kelompok Bimbingan Belajar Tambahan,” *J. Sist. Inf.*, vol. 3, no. 2, pp. 73–82, 2022, [Online]. Available: <http://ojs.itb-ad.ac.id/index.php/JUSIN/article/view/1961%0Ahttps://ojs.itb-ad.ac.id/index.php/JUSIN/article/download/1961/435>.
- [17] M. Azzam Al Fauzie and J. Akhir Putra, “Clustering Data Menggunakan Metode K-Means untuk Rekomendasikan Pembelajaran Akademik bagi Siswa Aktif dalam Ekstrakurikuler,” *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 1, pp. 642–648, 2023, doi: 10.30865/klik.v4i1.1116.
- [18] P. Subekti, T. D. Andini, and M. Islamiyah, “Sistem Penentuan Konsentrasi Jurusan Bagi Mahasiswa Informatika Menggunakan Metode K-Means Di Institut Asia Malang,” *J. Manaj. Inform.*, vol. 12, no. 1, pp. 25–39, 2022, doi: 10.34010/jamika.v12i1.6452.
- [19] M. R. Palevi and Z. Indra, “Implementasi Algoritma K-Means Clustering Dengan Pendekatan Active Learning Pada Siswa SMA Untuk Menentukan Jurusan Ke Perguruan Tinggi,” *J. SAINTIKOM (Jurnal Sains Manaj. Inform. dan Komputer)*, vol. 23, no. 1, pp. 26–36, 2024.
- [20] P. Marpaung, I. Pebrian, and W. Putri, “Penerapan Data Mining Untuk Pengelompokan Kepadatan Penduduk Kabupaten Deli Serdang Menggunakan Algoritma K-Means,” *J. Ilmu Komput. dan Sist. Inf.*, vol. 6, no. 2, pp. 64–70, 2023.
- [21] Suaharmanto, W. S. Utami, N. Pratiwi, and M. Faisal, “Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Clustering Perokok Usia Lebih dari 15 Tahun,” *Bull. Inf. Technol.*, vol. 4, no. 4, pp. 501–507, 2023.
- [22] D. Ariyanto, “Data Mining Menggunakan Algoritma K-Means untuk Klasifikasi Penyakit Infeksi Saluran Pernafasan Akut,” *J. Sistim Inf. dan Teknol.*, vol. 4, no. 1, pp. 13–18, 2022, doi: 10.37034/jsisfotek.v4i1.117.
- [23] R. Mauliadi, “Data Mining Menggunakan Algoritma K-Means Clustering dalam Analisis Tingkat Potongan Harga Terhadap Harga Jual Sepeda Motor Honda,” *J. Inform. Ekon. Bisnis*, vol. 4, no. 4, pp. 7–9, 2022, doi: 10.37034/infeb.v4i4.156.
- [24] A. Pangestu and T. Ridwan, “PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA K-MEANS PENGELOMPOKAN PELANGGAN BERDASARKAN KUBIKASI AIR TERJUAL MENGGUNAKAN WEKA,” vol. 11, no. 3, pp. 67–71, 2021.
- [25] N. W. Utami and I. G. J. E. Putra, “TEXT MINIG CLUSTERING UNTUK PENGELOMPOKAN TOPIK DOKUMEN PENELITIAN MENGGUNAKAN ALGORITMA K-MEANS DENGAN COSINE SIMILARITY,” *JINTEKS (Jurnal Inform. Teknol. dan Sains)*, vol. 4, no. 3, pp. 255–259, 2022.
- [26] R. Prasojo, Y. R. W. Utami, and R. T. Vulandari, “Implementasi K-Means Clustering Pada Pengelompokan Potensi Kerjasama Pelanggan,” *J. Teknol. Inf. dan Komun.*, vol. 7, no. 2, 2019, doi: 10.30646/tikomsin.v7i2.435.
- [27] E. Ramadanti and M. Muslih, “Penerapan Data Mining Algoritma K-Means Clustering Pada Populasi Ayam Petelur Di Indonesia,” *Rabit J. Teknol. dan Sist. Inf. Univrab*, vol. 7, no. 1, pp. 1–7, 2022, doi: 10.36341/rabit.v7i1.2155.
- [28] D. Anggarwati, O. Nurdianwan, I. Ali, and D. A. Kurnia, “Penerapan Algoritma K-Means Dalam Prediksi Penjualan Karoseri,” *J. Data Sci. Inform.*, vol. 1, no. 2, pp. 58–62, 2021.
- [29] S. M. Hutabarat and A. Sindar, “Data Mining Penjualan Suku Cadang Sepeda Motor Menggunakan Algoritma K-Means,” *J. Nas. Komputasi dan Teknol. Inf.*, vol. 2, no. 2, p. 126, 2019, doi: 10.32672/jnkti.v2i2.1555.
- [30] C. S. D. B. Sembiring, L. Hanum, and S. P. Tamba, “Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Menentukan Judul Skripsi Dan Jurnal Penelitian (Studi Kasus Ftik Unpri),” *J. Sist. Inf. dan Ilmu Komput. Prima(JUSIKOM PRIMA)*, vol. 5, no. 2, pp. 80–85, 2022, doi: 10.34012/journalsisteminformasidilmukomputer.v5i2.2393.

Journal of Computer System and Informatics (JoSYC)

ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)

Volume 5, No. 2, February 2024, Page 380-392

<https://ejurnal.seminar-id.com/index.php/josyc>

DOI 10.47065/josyc.v5i2.4970

- [31] R. NOVIANTO, “Penerapan Data Mining menggunakan Algoritma K-Means Clustering untuk Menganalisa Bisnis Perusahaan Asuransi,” *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 6, no. 1, pp. 85–95, 2019, doi: 10.35957/jatisi.v6i1.150.