

Penerapan Algoritma Clustering K-Means Data Mining dalam Pengelompokan Mahasiswa Penerima Beasiswa

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—Beasiswa merupakan sebuah program yang diperuntukan untuk membantu bagi para pelajar dalam permasalahan ekonomi. Bagi perguruan tinggi khususnya perguruan tinggi swasta beasiswa merupakan sebuah daya tarik ataupun ajang promosi kampus untuk menarik minat calon mahasiswa untuk mendaftar pada kampus tersebut. Beasiswa yang diberikan oleh pihak kampus merupakan beasiswa mandiri yang dimana berdasarkan dengan pendanaan oleh yayasan dari perguruan tinggi tersebut. Hal tersebut menjadi sangat penting untuk diperhatikan, dimana selain memiliki prestasi calon mahasiswa nantinya juga harus dipertimbangkan terhadap kesiapan ataupun kemampuan dalam mengikuti proses pembelajaran yang berlangsung di perguruan tinggi tersebut. Maka dari itu, memperhatikan nilai yang didapatkan dari calon mahasiswa sangatlah penting untuk diperhatikan. Permasalahan lainnya adalah kuota yang diberikan dari yayasan untuk beasiswa jumlah terbatas, dimana tidak tercover untuk semua calon mahasiswa yang mendaftar ataupun mengajukan dalam permohonan beasiswa. Dalam hal penentuan ataupun pemberian beasiswa ini belum terdapat sebuah standar acuan yang dipergunakan untuk penentuan dalam proses pengambilan keputusan, seringkali pemberian beasiswa menjadi salah sasaran. Kesalahan dalam pemberian beasiswa tersebut tentu saja sangat merugikan bagi pihak kampus. Oleh sebab itu, permasalahan ini sudah seharusnya membutuhkan perhatian serta penanganan khusus. Permasalahan tersebut dapat mudah diselesaikan dengan meneumkan sebuah pola aturan dalam penerimaan beasiswa. Data mining merupakan sebuah cara proses yang banyak dipergunakan saat ini, hal tersebut dikarenakan data mining sangat membantu dalam proses pengambilan keputusan. Pada proses yang dilakukan data mining, terbagi atas beberapa teknik didalamnya seperti Clustering. Clustering merupakan sebuah cara untuk melakukan pengelompokan data baru. Algoritma K-Means melakukan proses penyelesaian berdasarkan dengan pengelompokan, maka dari itu algoritma K-Means ini digolongkan menjadi bagian clustering pada data mining. Tujuan penelitian yang akan dilakukan adalah untuk membantu dalam proses pengelompokan terhadap calon mahasiswa yang akan diprioritaskan dalam penerimaan beasiswa. Berdasarkan dengan hasil penelitian tersebut nantinya dapat membantu untuk mendapatkan mahasiswa yang benar – benar layak untuk mendapatkan beasiswa tersebut. Hasil yang didapatkan dari penelitian bahwasannya terdapat 2 (cluster) yang didapatkan dari proses algoritma K-Means. Dimana pada cluster 1 terdapat 10 data pengelompokan dan pada cluster 2 terdapat 5 data pengelompokan.

Kata Kunci: Mahasiswa; Beasiswa; Algoritma K-Means; Data Mining

Abstract—Scholarships are a program intended to help students with economic problems. For universities, especially private universities, scholarships are an attraction or a campus promotional event to attract prospective students to register at the campus. The scholarships provided by the campus are independent scholarships which are based on funding from the university's foundation. This is very important to pay attention to, where apart from the achievements of prospective students, they must also consider their readiness or ability to participate in the learning process that takes place at the university. Therefore, paying attention to the grades obtained from prospective students is very important to pay attention to. Another problem is that the quota given by the foundation for scholarships is also limited, which is not covered by all prospective students who register or submit scholarship applications. In terms of determining or awarding scholarships, there is not yet a reference standard that is used for determination in the decision-making process, so scholarship awards are often misdirected. Mistakes in awarding scholarships are of course very detrimental to the campus. Therefore, this problem should require special attention and treatment. This problem can be easily resolved by finding a pattern of rules for accepting scholarships. Data mining is a process method that is widely used today, this is because data mining is very helpful in the decision making process. The process carried out by data mining is divided into several techniques such as Clustering. Clustering is a way to group new data. The K-Means algorithm carries out a solution process based on grouping, therefore the K-Means algorithm is classified as a clustering part of data mining. The aim of the research to be carried out is to assist in the process of grouping prospective students who will be prioritized in receiving scholarships. Based on the results of this research, it can later help to find students who are truly worthy of receiving the scholarship. The results obtained from the research are that there are 2 (clusters) obtained from the K-Means algorithm process. Where in cluster 1 there are 10 grouping data and in cluster 2 there are 5 grouping data.

Keywords: Students; Scholarship; K-Means Algorithm; Data Mining

1. PENDAHULUAN

Beasiswa merupakan sebuah program yang diperuntukan untuk membantu bagi para pelajar dalam permasalahan ekonomi. Dengan beasiswa nantinya para pelajar tidak perlu takut lagi untuk memikirkan terhadap biaya pendidikan yang perlu dikeluarkan, sehingga para pelajar lebih terfokus dalam melaksanakan pembelajaran yang dilakukannya[1], [2].

Program beasiswa sendiri bukan saja terdapat pada jenjang sekolah baik Sekolah Dasar, Sekolah Menengah Pertama, Sekolah Menengah Atas ataupun Sekolah Menengah Kejuruan, namun beasiswa juga

terdapat pada jenjang perguruan tinggi. Bagi perguruan tinggi khususnya perguruan tinggi swasta beasiswa merupakan sebuah daya tarik ataupun ajang promosi kampus untuk menarik minat calon mahasiswa untuk mendaftar pada kampus tersebut[3], [4].

Beasiswa yang diberikan oleh pihak kampus merupakan beasiswa mandiri yang dimana berdasarkan dengan pendanaan oleh yayasan dari perguruan tinggi tersebut. Pada umumnya beasiswa tersebut biasanya diberikan kepada calon mahasiswa yang memiliki prestasi baik dibidang akademik ataupun non akademik. Seharusnya beasiswa juga harus mempertimbangkan terhadap faktor lainnya seperti nilai yang didapatkan oleh calon mahasiswa tersebut[5], [6].

Hal tersebut menjadi sangat penting untuk diperhatikan, dimana selain memiliki prestasi calon mahasiswa nantinya juga harus dipertimbangkan terhadap kesiapan ataupun kemampuan dalam mengikuti proses pembelajaran yang berlangsung di perguruan tinggi tersebut. Dimana calon mahasiswa diharuskan untuk menyelesaikan studi tepat waktu dan juga tidak terdapat nilai yang baik[7], [8].

Maka dari itu, memperhatikan nilai yang didapatkan dari calon mahasiswa di masa Sekolah Menengah Atas ataupun Sekolah Menengah Kejuruan sangatlah penting untuk diperhatikan. Dimana dengan mempertimbangkan nilai yang didapatkan dimasa Sekolah Menengah Atas ataupun Sekolah Menengah Kejuruan nantinya sebagai tolak ukur yang digunakan untuk memperkirakan calon mahasiswa tersebut mampu atau tidak dalam mengikuti pelajaran nantinya[9], [10].

Permasalahan lainnya adalah kuota yang diberikan dari yayasan untuk beasiswa jugalah terbatas, dimana tidak tercover untuk semua calon mahasiswa yang mendaftar ataupun mengajukan dalam permohonan beasiswa. Maka dari itu pihak penyelenggaraan beasiswa ataupun tim penerima mahasiswa baru di perguruan tinggi haruslah bijak dalam pengambilan keputusan terhadap calon mahasiswa yang mendapatkan beasiswa.

Dalam hal penentuan ataupun pemberian beasiswa ini belum terdapat sebuah standar acuan yang dipergunakan untuk penentuan dalam proses pengambilan keputusan, seringkali pemberian beasiswa menjadi salah sasaran. Dimana pemberian beasiswa yang tidak tepat tentu akan merugikan bagi pihak kampus yang mengharapakan mahasiswa nantinya dapat mengikuti perkuliahan dengan baik dan juga tetap berprestasi demi menunjang kualitas dari kampus tersebut.

Kesalahan dalam pemberian beasiswa tersebut tentu saja sangat merugikan bagi pihak kampus, dimana dalam pembatalan pemberian beasiswa juga diharuskan membutuhkan dasar atau landasan pendukung dalam pengambilan keputusan. Selain itu juga, kesalahan dalam pemberian beasiswa yang ditakutkan adalah mahasiswa tidak dapat mengikuti sebagaimana mestinya yang diharapkan dan juga nilai dari setiap matakuliah mendapatkan hasil yang buruk.

Oleh sebab itu, permasalahan ini sudah seharusnya membutuhkan perhatian serta penanganan khusus. Bagi tim penyelenggara beasiswa ataupun penerimaan mahasiswa baru sudah seharusnya memiliki sebuah dasar acuan dalam proses pengambilan keputusan untuk menentukan siapa calon mahasiswa yang layak untuk menerima beasiswa dengan mempertimbangkan terhadap nilai yang didapatkan ketika masa Sekolah Menengah Atas ataupun Sekolah Menengah Kejuruan.

Permasalahan tersebut dapat mudah diselesaikan dengan menemukan sebuah pola aturan dalam penerimaan beasiswa, dimana pola tersebut dapat terlihat berdasarkan dengan data penerimaan sebelumnya. Dimana proses penerimaan beasiswa sebelumnya dapat menjadi sebagai pertimbangan dalam pengambilan keputusan yang berlaku. Data – data sebelumnya dapat dipergunakan kembali dengan melakukan sebuah pemrosesan sehingga didapatkan hasil yang dapat digunakan sebagai pengambilan keputusan. Penyelesaian terhadap permasalahan tersebut juga membutuhkan sebuah cara atau biasa disebut dengan data mining.

Data mining merupakan sebuah cara proses yang banyak dipergunakan saat ini, hal tersebut dikarenakan data mining sangat membantu dalam proses pengambilan keputusan. Pada data mining nantinya akan dilakukan proses penggalian informasi berdasarkan dengan data – data yang telah tersimpan sebelumnya, kemudian data tersebut dipergunakan kembali sehingga didapatkan hasil dalam pengambilan keputusan. Sama halnya dengan proses penyelesaian yang diharapkan, oleh sebab itu data mining merupakan sebuah cara yang tepat digunakan untuk menyelesaikan permasalahan berkaitan dengan pemrosesan data – data dimasa lampau[11]–[13].

Pada proses yang dilakukan data mining, terbagi atas beberapa teknik didalamnya seperti Clustering, Prediksi, Estimasi, Asosiasi dan Klasifikasi. Pada penyelesaian ini terfokus pada clustering. Clustering merupakan sebuah cara untuk melakukan pengelompokan data baru. Dimana pengelompokan yang dilakukan pada clustering berdasarkan dengan karakteristik yang dimiliki oleh data. Karakteristik tersebut nantinya dikelompokkan menjadi sebuah kelompok – kelompok baru. Proses tersebut juga sesuai dipergunakan untuk proses penyelesaian yang dilakukan dalam penelitian, dimana nantinya dapat mengelompokkan calon mahasiswa yang diprioritaskan untuk menerima beasiswa[14], [15].

Dalam clustering data mining sendiri, terdapat algoritma penyelesaian yang digunakan salah satunya adalah algoritma K-Means. Algoritma K-Means melakukan proses penyelesaian berdasarkan dengan pengelompokan, maka dari itu algoritma K-Means ini digolongkan menjadi bagian clustering pada data mining. Algoritma K-Means data mining dapat dipergunakan untuk mengelompokkan data yang tidak memiliki target kelasnya. Dimana data yang dipergunakan bukan merupakan data supervised learning[16], [17].

Untuk mendukung terhadap pelaksanaan penelitian, maka dari itu diperlukan sebuah referensi dari penelitian terdahulu terkait dengan penelitian yang dilakukan. Dimana hal tersebut dapat dilihat dari penelitian

yang dilakukan oleh Ronal Watrianthos, dkk pada tahun 2022 dengan judul penelitian Penerapan Algoritma K-Means Pada Pemetaan Kemampuan Penggunaan Teknologi Informasi Remaja dan Dewasa di Indonesia dari penelitian yang dilakukan dimana hasil yang didapatkan Sebagai salah satu metode pada unsupervised learning, algoritma K-Means digunakan agar data set dapat diorganisasikan ke dalam kelompok dengan item lain yang sebanding dengannya. Hasil penelitian menunjukkan bahwasannya terdapat 7 (tujuh) cluster yang didapatkan[18].

Selain penelitian tersebut, juga dilakukan sebuah penelitian oleh Tri Bayu Pamungkas, dkk pada tahun 2023 dengan judul penelitian yang dilakukan Implementasi Data Mining Pada Stok Penggunaan Barang Di Gmf Aeroasia Menggunakan Algoritma K-Means Clustering serta hasil yang didapatkan adalah Hasil analisis dan penelitian terhadap data penggunaan barang menggunakan algoritma k-means menunjukkan bahwa data berhasil dikelompokkan menjadi tiga klaster[19].

Penelitian selanjutnya yang digunakan sebagai dasar penelitian yang dilakukan oleh Yosi Mayona, dkk pada tahun 2022 dengan judul penelitian Data Mining Clustering Tingkat Kejahatan Dengan Metode Algoritma K-Means (Studi Kasus: Kejaksaan Negeri Binjai) dimana hasil yang didapatkan bahwasannya Algoritma K-Means merupakan algoritma non hirarki dari metode clustering yang mempunyai kemampuan mengelompokkan data dalam jumlah yang cukup besar, cepat dan efisien. Berdasarkan perhitungan dapat diketahui bahwa pada cluster 3 Pengelompokan tindak kejahatan[20].

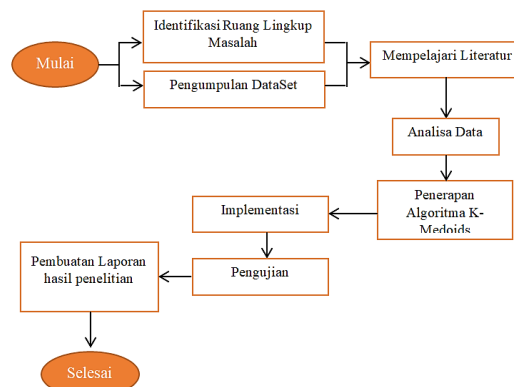
Dan penelitian terakhir yang digunakan sebagai pendukung penelitian dilakukan oleh Riski Askia Kurniawan, dkk pada tahun 2022 dengan judul penelitian Penerapan Algoritma K-Means Untuk Clustering Tempat Makan Di Batubara serta hasil yang didapatkan dari penelitian bahwasannya terdapat 2 (dua) cluster yang terbentuk dari hasil proses penelitian[21].

Berdasarkan dengan penjabaran permasalahan yang terjadi maka dari itu tujuan penelitian yang akan dilakukan adalah untuk membantu dalam proses pengelompokan terhadap calon mahasiswa yang akan diprioritaskan dalam penerimaan beasiswa. Berdasarkan dengan hasil penelitian tersebut nantinya dapat membantu untuk mendapatkan mahasiswa yang benar – benar layak untuk mendapatkan beasiswa tersebut.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Tahapan penelitian merupakan alur proses yang dilaksanakan pada penelitian. Dengan tahapan penelitian maka alur proses akan tersusun secara baik. Alur proses pada tahapan penelitian dimulai dari tahapan identifikasi masalah sampai dengan tahapan laporan. Adapun tahapan penelitian dapat dilihat pada gambar 1 berikut.



Gambar 1. Tahapan Penelitian

2.2 Data Mining

Pada masa sekarang ini data berkembang dengan sangat pesat, dimana data sangat berperan penting bagi berbagai macam kalangan. Berdasarkan dengan data maka akan memudahkan dalam proses pengambilan keputusan yang konkret dan jelas. Namun nyatanya proses dalam pengambilan keputusan terhadap data bukan sebuah hal yang mudah, hal tersebut dikarenakan saling berkaitan antara atribut yang dimiliki dari data[2]–[7].

Peran dari data mining sangatlah penting terhadap hal tersebut, hal tersebut dikarenakan data mining dipergunakan untuk membantu mempermudah dalam penyelesaian ataupun pengolahan data yang cukup besar dan memiliki keterkaitan atribut yang cukup kompleks. Hal tersebut didasari dengan proses yang dilakukan dari data mining yaitu dengan melakukan pengolahan kembali terhadap data yang tersimpan serta mendapatkan sebuah informasi baru yang dapat dipergunakan untuk proses pengambilan keputusan[18], [20], [22].

Oleh sebab itu, data mining dimasa sekarang ini merupakan proses yang penting dan sudah hampir dipergunakan oleh berbagai macam kalangan untuk membantu dalam proses pengambilan keputusan yang dilakukan. Hasil yang didapatkan pada data mining juga cukup beragam, dimana hasil yang didapatkan

bergantung dengan bentuk seperti apa yang diinginkan, sehingga menyebabkan data mining lebih fleksibel untuk dipergunakan.

2.3 Algoritma K-Means

Algoritma K-Means merupakan algoritma clustering pada data mining, dimana pada algoritma tersebut membagikan data dalam kelompok – kelompok tertentu ataupun disebut dengan cluster berdasarkan dengan karakteristik ataupun kemiripan dari data. Pada clustering, algoritma K-Means saat ini banyak dipergunakan. Hal tersebut dikarenakan kemudahan dalam penggunaannya dan kesederhanaan dalam proses yang dilakukannya[23], [24].

Algoritma K-Means dapat dipergunakan untuk menyelesaikan data yang cukup besar dengan mudahnya, dimana proses pada algoritma K-Means didasarkan dengan euclidean distance yang artinya berdasarkan dengan jarak terdekat yang dimiliki antara objek yang satu dengan objek yang lainnya. Adapun tahapan proses yang dilakukan pada algoritma K-Means dapat dilihat seperti berikut[22], [25]–[27]:

1. Menentukan nilai K atau banyaknya cluster yang diinginkan dari hasil
2. Menentukan nilai centeroid awal dari setiap cluster. Dimana nilai centeroid awal ditentukan secara acak (random) dari data yang tersedia
3. Menghitung jarak terdekat dari data yang ada terhadap data pada nilai centeroid cluster dengan menggunakan euclidean distance hingga didapatkan jarak paling terdekat. Adapun rumus yang digunakan untuk melakukan perhitungan jarak menggunakan rumus berikut:

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

4. Mengelompokkan data berdasarkan dengan nilai jarak terdekat yang dimiliki antara data centeroid dengan data yang terdapat pada dataset.
5. Selanjutnya, melanjutkan proses pada iterasi selanjutnya. Untuk melakukan iterasi selanjutnya terlebih dahulu mencari nilai centeroid baru berdasarkan hasil pembentukan cluster. Adapun rumus yang digunakan untuk mencari nilai centeroid baru dari setiap cluster yaitu:

$$y_i(t + 1) = \frac{1}{N_{sj}} \sum_{j \in s_j} x_j \quad (2)$$

6. Mengulang kembali proses langkah 2 sampai dengan langkah 5. Dimana proses akan berhenti ketika sudah tidak terdapat perubahan terhadap hasil cluster.

3. HASIL DAN PEMBAHASAN

Beasiswa merupakan program yang sangat membantu dalam melanjutkan studi. Beasiswa bukan saja diberikan oleh pemerintah, tetapi juga dapat diberikan secara mandiri. Bagi perguruan tinggi swasta, program beasiswa menjadi sebuah media promosi atau ajang untuk menarik calon mahasiswa untuk mendaftar pada perguruan tinggi tersebut. Pada umumnya pemberian beasiswa dilakukan berdasarkan dengan prestasi akademik dan nonakademik, seharusnya pemberian beasiswa juga harus mempertimbangkan terhadap hasil nilai yang didapatkan oleh calon mahasiswa di masa Sekolah Menengah Atas ataupun Sekolah Menengah Kejuruan. Pertimbangan terhadap aspek nilai dikarenakan ketika menjadi mahasiswa nantinya diharuskan dan dituntut untuk dapat menyelesaikan studi secara tepat waktu serta nilai – nilai yang didapatkan dari hasil pembelajaran matakuliah juga harus memuaskan. Sering kali pemberian beasiswa masih dilakukan tidak sesuai dengan sasarannya, hal tersebut mengakibatkan mahasiswa tidak dapat mengikuti pembelajaran dengan baik yang menyebabkan masa studi menjadi terhambat serta nilai perkuliahan yang berantakan. Maka dari itu perlu sebuah penanganan khusus terhadap permasalahan yang dihadapi. Penyelesaian dapat dilakukan dengan melihat terhadap data – data mahasiswa sebelumnya yang mampu untuk mengikuti perkuliahan secara keseluruhan, dari data tersebut nantinya akan didapatkan sebuah pola ataupun informasi yang dapat dipergunakan dan dimanfaatkan dalam proses pengambilan keputusan. Berdasarkan dengan hal tersebut maka data mining sangat pantas dipergunakan untuk menyelesaikan permasalahan. Data mining merupakan sebuah proses penggalian data hingga didapatkan sebuah informasi yang dipergunakan untuk proses pengambilan keputusan. Dalam data mining sendiri, terdapat beberapa macam teknik yang dipergunakan untuk membantu dalam penyelesaian permasalahan, salah satunya adalah clustering. Clustering merupakan sebuah teknik yang dipergunakan untuk melakukan pengelompokan berdasarkan dengan pembentukan cluster baru. Pembentukan cluster tersebut berdasarkan dengan karakteristik yang dimiliki dari data tersebut. Maka dari itu, penelitian ini bertujuan untuk membantu penyelesaian permasalahan dalam pengelompokan calon mahasiswa yang akan menerima beasiswa.

3.1 Penerapan Algoritma K-Means

Analisa terhadap masalah merupakan sebuah hal penting yang harus dilakukan dalam penelitian, tanpa adanya analisa masalah maka penelitian tidak mungkin dapat terselesaikan. Setelah dilakukannya analisa masalah seperti tampak pada pembahasan sebelumnya, maka dapat dilakukan proses penyelesaian terhadap penggunaan

algoritma K-Means. Namun sebelum dilakukan penyelesaian terlebih dahulu harus diketahui terhadap dataset yang akan digunakan dari penelitian. Adapun dataset yang akan digunakan pada penelitian dapat dilihat pada tabel 1.

Tabel 1. Dataset Penelitian

Nama Calon Mahasiswa	Nilai 1	Nilai 2	Nilai 3	Nilai 4	Nilai 5
Calon Mahasiswa 1	84	96	78	79	75
Calon Mahasiswa 2	88	96	95	78	82
Calon Mahasiswa 3	97	75	77	95	86
Calon Mahasiswa 4	76	93	98	77	93
Calon Mahasiswa 5	95	94	77	93	92
Calon Mahasiswa 6	87	92	85	78	85
Calon Mahasiswa 7	80	81	88	75	92
Calon Mahasiswa 8	86	79	93	80	86
Calon Mahasiswa 9	95	91	85	89	85
Calon Mahasiswa 10	98	94	88	88	87
Calon Mahasiswa 11	78	97	90	90	95
Calon Mahasiswa 12	96	80	82	78	75
Calon Mahasiswa 13	83	92	95	84	94
Calon Mahasiswa 14	82	90	84	94	89
Calon Mahasiswa 15	86	81	96	81	89

Pada tabel 1 merupakan dataset yang akan digunakan pada penelitian, dari tabel tersebut dapat terlihat bahwasannya terdapat 15 (lima belas) data calon mahasiswa yang akan menerima beasiswa. Kemudian setelah didapatkan nilai pada dataset, maka selanjutnya dapat diselesaikan dengan menggunakan algoritma K-Means tersebut. Adapun penyelesaian dapat dilakukan seperti berikut:

1. Menentukan nilai K atau banyaknya cluster yang diinginkan dari hasil; Pada penelitian banyaknya cluster yang akan dibentuk sebanyak 2 (dua) cluster.
2. Menentukan nilai centeroid awal dari setiap cluster. Dimana nilai centeroid awal ditentukan secara acak (random) dari data yang tersedia; Pada tahapan ini nilai centeroid awal untuk setiap cluster diambil berdasarkan dengan data pada dataset. Maka dari itu untuk nilai centeroid awal yang digunakan pada Cluster 1 merupakan nilai calon mahasiswa 10. Sedangkan nilai centeroid awal yang digunakan pada Cluster 2 merupakan nilai calon mahasiswa 12. Untuk lebih memudahkan dalam melihat data nilai centeroid awal, maka dapat dilihat pada tabel 2 berikut.

Tabel 2. Nilai Centeroid Awal

Nama Calon Mahasiswa	Nilai 1	Nilai 2	Nilai 3	Nilai 4	Nilai 5
Calon Mahasiswa 10	98	94	88	88	87
Calon Mahasiswa 12	96	80	82	78	75

3. Setelah diketahui nilai centeroid awal untuk setiap cluster maka proses selanjutnya menghitung jarak terdekat dari data yang ada terhadap data pada nilai centeroid cluster dengan menggunakan euclidean distance hingga didapatkan jarak paling terdekat. Adapun rumus yang digunakan untuk melakukan perhitungan jarak menggunakan rumus berikut:

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

Cluser – 1

$$d(x1, c1)$$

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

$$= \sqrt{(84 - 98)^2 + (96 - 94)^2 + (78 - 88)^2 + (79 - 88)^2 + (75 - 87)^2}$$

$$= 22,91$$

$$d(x2, c1)$$

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

$$= \sqrt{(88 - 98)^2 + (96 - 94)^2 + (95 - 88)^2 + (78 - 88)^2 + (82 - 87)^2}$$

$$= 16,67$$

$$d(x3, c1)$$

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

$$= \sqrt{(97 - 98)^2 + (75 - 94)^2 + (77 - 88)^2 + (95 - 88)^2 + (86 - 87)^2}$$

$$= 23,09$$

$$\begin{aligned} d(x4, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(76 - 98)^2 + (93 - 94)^2 + (98 - 88)^2 + (77 - 88)^2 + (93 - 87)^2} \\ &= 27,24 \end{aligned}$$

$$\begin{aligned} d(x5, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(95 - 98)^2 + (94 - 94)^2 + (77 - 88)^2 + (93 - 88)^2 + (92 - 87)^2} \\ &= 13,42 \end{aligned}$$

$$\begin{aligned} d(x6, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(87 - 98)^2 + (92 - 94)^2 + (85 - 88)^2 + (78 - 88)^2 + (85 - 87)^2} \\ &= 15,43 \end{aligned}$$

$$\begin{aligned} d(x7, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(80 - 98)^2 + (81 - 94)^2 + (88 - 88)^2 + (75 - 88)^2 + (92 - 87)^2} \\ &= 26,21 \end{aligned}$$

$$\begin{aligned} d(x8, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(86 - 98)^2 + (79 - 94)^2 + (93 - 88)^2 + (80 - 88)^2 + (86 - 87)^2} \\ &= 21,42 \end{aligned}$$

$$\begin{aligned} d(x9, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(95 - 98)^2 + (91 - 94)^2 + (85 - 88)^2 + (89 - 88)^2 + (85 - 87)^2} \\ &= 5,66 \end{aligned}$$

$$\begin{aligned} d(x10, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(98 - 98)^2 + (94 - 94)^2 + (88 - 88)^2 + (88 - 88)^2 + (87 - 87)^2} \\ &= 0 \end{aligned}$$

$$\begin{aligned} d(x11, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(78 - 98)^2 + (97 - 94)^2 + (90 - 88)^2 + (90 - 88)^2 + (95 - 87)^2} \\ &= 21,93 \end{aligned}$$

$$\begin{aligned} d(x12, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(96 - 98)^2 + (80 - 94)^2 + (82 - 88)^2 + (78 - 88)^2 + (75 - 87)^2} \\ &= 21,91 \end{aligned}$$

$$\begin{aligned} d(x13, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(83 - 98)^2 + (92 - 94)^2 + (95 - 88)^2 + (84 - 88)^2 + (94 - 87)^2} \\ &= 18,52 \end{aligned}$$

$$\begin{aligned} d(x14, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(82 - 98)^2 + (90 - 94)^2 + (84 - 88)^2 + (94 - 88)^2 + (89 - 87)^2} \\ &= 18,11 \end{aligned}$$

$$\begin{aligned} d(x15, c1) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\ &= \sqrt{(86 - 98)^2 + (81 - 94)^2 + (96 - 88)^2 + (81 - 88)^2 + (89 - 87)^2} \\ &= 20,74 \end{aligned}$$

Cluster – 2

$$\begin{aligned}d(x1, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\&= \sqrt{(84 - 96)^2 + (96 - 80)^2 + (78 - 82)^2 + (79 - 78)^2 + (75 - 75)^2} \\&= 20,42\end{aligned}$$

$$\begin{aligned}d(x2, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\&= \sqrt{(88 - 96)^2 + (96 - 80)^2 + (95 - 82)^2 + (78 - 78)^2 + (82 - 75)^2} \\&= 23,19\end{aligned}$$

$$\begin{aligned}d(x3, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\&= \sqrt{(97 - 96)^2 + (75 - 80)^2 + (77 - 82)^2 + (95 - 78)^2 + (86 - 75)^2} \\&= 21,47\end{aligned}$$

$$\begin{aligned}d(x4, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\&= \sqrt{(76 - 96)^2 + (93 - 80)^2 + (98 - 82)^2 + (77 - 78)^2 + (93 - 75)^2} \\&= 33,91\end{aligned}$$

$$\begin{aligned}d(x5, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\&= \sqrt{(95 - 96)^2 + (94 - 80)^2 + (77 - 82)^2 + (93 - 78)^2 + (92 - 75)^2} \\&= 27,13\end{aligned}$$

$$\begin{aligned}d(x6, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\&= \sqrt{(87 - 96)^2 + (92 - 80)^2 + (85 - 82)^2 + (78 - 78)^2 + (85 - 75)^2} \\&= 18,28\end{aligned}$$

$$\begin{aligned}d(x7, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\&= \sqrt{(80 - 96)^2 + (81 - 80)^2 + (88 - 82)^2 + (75 - 78)^2 + (92 - 75)^2} \\&= 24,31\end{aligned}$$

$$\begin{aligned}d(x8, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\&= \sqrt{(86 - 96)^2 + (79 - 80)^2 + (93 - 82)^2 + (80 - 78)^2 + (86 - 75)^2} \\&= 18,63\end{aligned}$$

$$\begin{aligned}d(x9, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\&= \sqrt{(95 - 96)^2 + (91 - 80)^2 + (85 - 82)^2 + (89 - 78)^2 + (85 - 75)^2} \\&= 18,76\end{aligned}$$

$$\begin{aligned}d(x10, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\&= \sqrt{(98 - 96)^2 + (94 - 80)^2 + (88 - 82)^2 + (88 - 78)^2 + (87 - 75)^2} \\&= 21,91\end{aligned}$$

$$\begin{aligned}d(x11, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\&= \sqrt{(78 - 96)^2 + (97 - 80)^2 + (90 - 82)^2 + (90 - 78)^2 + (95 - 75)^2} \\&= 34,94\end{aligned}$$

$$\begin{aligned}d(x12, c2) &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\&= \sqrt{(96 - 96)^2 + (80 - 80)^2 + (82 - 82)^2 + (78 - 78)^2 + (75 - 75)^2} \\&= 0\end{aligned}$$

$$\begin{aligned}
 & d(x_{13}, c_2) \\
 &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(83 - 96)^2 + (92 - 80)^2 + (95 - 82)^2 + (84 - 78)^2 + (94 - 75)^2} \\
 &= 29,65
 \end{aligned}$$

$$\begin{aligned}
 & d(x_{14}, c_2) \\
 &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(82 - 96)^2 + (90 - 80)^2 + (84 - 82)^2 + (94 - 78)^2 + (89 - 75)^2} \\
 &= 27,42
 \end{aligned}$$

$$\begin{aligned}
 & d(x_{15}, c_2) \\
 &= \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2 + (e_1 - c_{1e})^2} \\
 &= \sqrt{(86 - 96)^2 + (81 - 80)^2 + (96 - 82)^2 + (81 - 78)^2 + (89 - 75)^2} \\
 &= 22,4
 \end{aligned}$$

4. Mengelompokkan data berdasarkan dengan nilai jarak terdekat yang dimiliki antara data centeroid dengan data yang terdapat pada dataset. Setelah proses perhitungan selesai, maka selanjutnya dilakukan proses pengelompokan data dan untuk melihat data beradap pada kelompok atau Cluster ataupun Cluster 2. Adapun hasil pengelompokan dapat dilihat pada tabel 3 berikut.

Tabel 3. Hasil Pengelompokan Cluster Iterasi 1

Nama Calon Mahasiswa	Nilai 1	Nilai 2	Nilai 3	Nilai 4	Nilai 5	DC1	DC2	Hasil Cluster
Calon Mahasiswa 1	84	96	78	79	75	22,91	20,42	C2
Calon Mahasiswa 2	88	96	95	78	82	16,67	23,19	C1
Calon Mahasiswa 3	97	75	77	95	86	23,09	21,47	C2
Calon Mahasiswa 4	76	93	98	77	93	27,24	33,91	C1
Calon Mahasiswa 5	95	94	77	93	92	13,42	27,13	C1
Calon Mahasiswa 6	87	92	85	78	85	15,43	18,28	C1
Calon Mahasiswa 7	80	81	88	75	92	26,21	24,31	C2
Calon Mahasiswa 8	86	79	93	80	86	21,42	18,63	C2
Calon Mahasiswa 9	95	91	85	89	85	5,66	18,76	C1
Calon Mahasiswa 10	98	94	88	88	87	0,00	21,91	C1
Calon Mahasiswa 11	78	97	90	90	95	21,93	34,94	C1
Calon Mahasiswa 12	96	80	82	78	75	21,91	0,00	C2
Calon Mahasiswa 13	83	92	95	84	94	18,52	29,65	C1
Calon Mahasiswa 14	82	90	84	94	89	18,11	27,42	C1
Calon Mahasiswa 15	86	81	96	81	89	20,74	22,41	C1

Hasil pengelompokan terhadap cluster dapat dilihat pada tabel 3. Pada hasil di tabel tersebut dapat terlihat untuk data berada pada cluster mana. Pada data tersebut dapat terlihat bahwasannya untuk Cluster 1 terdapat 10 (sepuluh) data calon mahasiswa, sedangkan pada cluster 2 terdapat 5 (lima) data calon mahasiswa.

Iterasi – 2

Selanjutnya, melanjutkan proses pada iterasi selanjutnya. Untuk melakukan iterasi selanjutnya terlebih dahulu mencari nilai centeroid baru berdasarkan hasil pembentukan cluster. Adapun rumus yang digunakan untuk mencari nilai centeroid baru dari setiap cluster yaitu:

$$y_i(t + 1) = \frac{1}{N_{sj}} \sum_{jesj} x_j$$

Tabel 4. Nilai Centeroid Baru Cluster 1

Nama Calon Mahasiswa	Nilai 1	Nilai 2	Nilai 3	Nilai 4	Nilai 5
Calon Mahasiswa 2	88	96	95	78	82
Calon Mahasiswa 4	76	93	98	77	93
Calon Mahasiswa 5	95	94	77	93	92
Calon Mahasiswa 6	87	92	85	78	85
Calon Mahasiswa 9	95	91	85	89	85
Calon Mahasiswa 10	98	94	88	88	87
Calon Mahasiswa 11	78	97	90	90	95
Calon Mahasiswa 13	83	92	95	84	94
Calon Mahasiswa 14	82	90	84	94	89
Calon Mahasiswa 15	86	81	96	81	89

Nama Calon Mahasiswa	Nilai 1	Nilai 2	Nilai 3	Nilai 4	Nilai 5
	86,8	92	89,3	85,2	89,1

Setelah didapatkan nilai centeroid baru untuk cluster 1 pada tabel 4, maka dapat dilihat juga untuk nilai centeroid baru untuk cluster 2. Adapun nilai centeroid baru dapat dilihat pada tabel 5.

Tabel 5. Nilai Centeroid Baru Cluster 2

Nama Calon Mahasiswa	Nilai 1	Nilai 2	Nilai 3	Nilai 4	Nilai 5
Calon Mahasiswa 1	84	96	78	79	75
Calon Mahasiswa 3	97	75	77	95	86
Calon Mahasiswa 7	80	81	88	75	92
Calon Mahasiswa 8	86	79	93	80	86
Calon Mahasiswa 12	96	80	82	78	75
	88,6	82,2	83,6	81,4	82,8

Setelah didapatkan nilai centeroid baru pada cluster 2 pada tabel 5, untuk mempermudah dalam proses dan membaca data maka lebih baik data disusun pada tabel 6.

Tabel 6. Nilai Centeroid Baru Cluster 1 dan Cluster 2

Cluster	Nilai 1	Nilai 2	Nilai 3	Nilai 4	Nilai 5
C1	86,8	92	89,3	85,2	89,1
C2	88,6	82,2	83,6	81,4	82,8

Pada tabel 6 dapat dilihat nilai centeroid baru untuk cluster 1 dan cluster 2 pada iterasi 2. Maka selanjutnya dengan proses yang sama seperti iterasi 1. Proses terhadap perhitungan jarak yang dilakukan pada iterasi 2 sama halnya dengan perhitungan jarak pada iterasi 1. Jarak dihitung berdasarkan data dan juga nilai centeroid baru untuk masing – masing cluster. Adapaun untuk hasil dari penentuan cluster pada iterasi 2 dapat dilihat pada tabel 7 berikut ini :

Tabel 7. Hasil Perhitungan Jarak Iterasi 2

Nama Calon Mahasiswa	Nilai 1	Nilai 2	Nilai 3	Nilai 4	Nilai 5	DC1	DC2	Hasil Cluster
Calon Mahasiswa 1	84	96	78	79	75	19,72	17,59	C2
Calon Mahasiswa 2	88	96	95	78	82	12,34	18,25	C1
Calon Mahasiswa 3	97	75	77	95	86	25,49	19,00	C2
Calon Mahasiswa 4	76	93	98	77	93	16,61	24,62	C1
Calon Mahasiswa 5	95	94	77	93	92	17,08	21,05	C1
Calon Mahasiswa 6	87	92	85	78	85	9,34	10,81	C1
Calon Mahasiswa 7	80	81	88	75	92	16,77	14,84	C2
Calon Mahasiswa 8	86	79	93	80	86	14,83	10,84	C2
Calon Mahasiswa 9	95	91	85	89	85	10,86	13,53	C1
Calon Mahasiswa 10	98	94	88	88	87	11,97	17,55	C1
Calon Mahasiswa 11	78	97	90	90	95	12,68	24,40	C1
Calon Mahasiswa 12	96	80	82	78	75	23,08	11,60	C2
Calon Mahasiswa 13	83	92	95	84	94	8,51	19,74	C1
Calon Mahasiswa 14	82	90	84	94	89	11,51	17,37	C1
Calon Mahasiswa 15	86	81	96	81	89	13,57	14,16	C1

Pada tabel 7, dapat dilihat hasil dari perhitungan jarak untuk iterasi 2 dan juga hasil pengelompokan data terhadap masing – masing cluster. Setelah didapatkan hasil pada iterasi 2, maka selanjutnya dilakukan perbandingan terhadap hasil yang didapatkan antara hasil clustering pada iterasi 1 dan hasil clustering pada iterasi 2. Dimana untuk hasil tersebut dapat dilihat pada tabel 8 berikut:

Tabel 8. Hasil Perbandingan Cluster Iterasi 1 dan Iterasi 2

Nama Calon Mahasiswa	Nilai 1	Nilai 2	Nilai 3	Nilai 4	Nilai 5	Hasil Iterasi 1	Hasil Iterasi 2
Calon Mahasiswa 1	84	96	78	79	75	C2	C2
Calon Mahasiswa 2	88	96	95	78	82	C1	C1
Calon Mahasiswa 3	97	75	77	95	86	C2	C2
Calon Mahasiswa 4	76	93	98	77	93	C1	C1
Calon Mahasiswa 5	95	94	77	93	92	C1	C1
Calon Mahasiswa 6	87	92	85	78	85	C1	C1
Calon Mahasiswa 7	80	81	88	75	92	C2	C2
Calon Mahasiswa 8	86	79	93	80	86	C2	C2
Calon Mahasiswa 9	95	91	85	89	85	C1	C1

Nama Calon Mahasiswa	Nilai 1	Nilai 2	Nilai 3	Nilai 4	Nilai 5	Hasil Iterasi 1	Hasil Iterasi 2
Calon Mahasiswa 10	98	94	88	88	87	C1	C1
Calon Mahasiswa 11	78	97	90	90	95	C1	C1
Calon Mahasiswa 12	96	80	82	78	75	C2	C2
Calon Mahasiswa 13	83	92	95	84	94	C1	C1
Calon Mahasiswa 14	82	90	84	94	89	C1	C1
Calon Mahasiswa 15	86	81	96	81	89	C1	C1

Dari tabel 8 diatas dapat dilihat hasil perbandingan clustering dapat pada iterasi 1 dan iterasi 2. Pada tabel tersebut dapat dilihat bahwasannya tidak terdapat perbedaan hasil clustering data, maka dari itu proses pada algoritma K-Means dihentikan dan berakhir pada iterasi 2. Dari hasil tersebut, maka didapatkan hasil pengelompokan calon mahasiswa seperti tabel 9 berikut:

Tabel 9. Hasil Pengelompokan Cluster Penerima Beasiswa

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

Tabel 9 merupakan hasil akhir dari proses algoritma K-Means, pada tabel 9 menjelaskan secara detail siapa saja yang termasuk dalam cluster. Dari data tersebut dapat dilihat bahwasannya terdapat 10 (sepuluh) data calon mahasiswa yang masuk dalam cluster 1 yaitu Calon Mahasiswa 2, Calon Mahasiswa 4, Calon Mahasiswa 5, Calon Mahasiswa 6, Calon Mahasiswa 9, Calon Mahasiswa 10, Calon Mahasiswa 11, Calon Mahasiswa 13, Calon Mahasiswa 14, dan Calon Mahasiswa 15. Sedangkan untuk cluster 2 terdapat 5 (lima) data calon mahasiswa yaitu Calon Mahasiswa 1, Calon Mahasiswa 3, Calon Mahasiswa 7, Calon Mahasiswa 8 dan Calon Mahasiswa 12.

4. KESIMPULAN

Tahapan akhir dari sebuah penelitian merupakan penarikan kesimpulan. Dari proses yang telah dilaksanakan, maka dapat ditarik kesimpulan bahwasannya Data mining merupakan cara yang tepat digunakan untuk membantu dalam pengambilan keputusan, dimana hal tersebut dikarenakan data mining dapat digunakan untuk mengolah data yang cukup besar serta mendapatkan sebuah informasi terhadap proses penggalian data. Proses pada data mining yang dilakukan dengan menggunakan clustering dan algoritma K-Means. Terdapat 2 hasil clustering yang didapatkan serta hasil dari penerapan algoritma K-Means bahwasannya terdapat 10 (sepuluh) data calon mahasiswa yang masuk dalam cluster 1 yaitu Calon Mahasiswa 2, Calon Mahasiswa 4, Calon Mahasiswa 5, Calon Mahasiswa 6, Calon Mahasiswa 9, Calon Mahasiswa 10, Calon Mahasiswa 11, Calon Mahasiswa 13, Calon Mahasiswa 14, dan Calon Mahasiswa 15. Sedangkan untuk cluster 2 terdapat 5 (lima) data calon mahasiswa yaitu Calon Mahasiswa 1, Calon Mahasiswa 3, Calon Mahasiswa 7, Calon Mahasiswa 8 dan Calon Mahasiswa 12.

REFERENCES

- [1] L. Marlia, P. Ngatmini, and P. E. Tobing, "Analisis Segmentasi Pengunjung Mall Menggunakan Algoritma K-means dan Fuzzy C-means," *SENTIMAS Semin. Nas. Penelit. dan Pengabd. Masy.*, pp. 240–247, 2023, [Online]. Available: <https://journal.irpi.or.id/index.php/sentimas%0D>.
- [2] A. S. Wahyuni, E. Haerani, E. Budianita, and L. Afrianti, "Pemanfaatan Algoritma K-Means Dalam Menentukan Potensi Hasil Produksi Kelapa Sawit," *J. Sist. Komput. dan Inform.*, vol. 5, no. 2, pp. 0–7, 2023, doi: 10.30865/json.v5i2.7226.
- [3] E. Desi, S. Aliyah, S. Lestari, and W. Dari, "Implementasi Algoritma K-Means Untuk Penerimaan Siswa Baru Di SMANPAS Berdasarkan Nilai Rapot dan Hasil Tes," *It (Informatic Tech. J.)*, vol. 10, no. 1, p. 01, 2022, doi: 10.22303/it.10.1.2022.01-10.
- [4] G. Triyandana, L. A. Putri, and Y. Umaidah, "Penerapan Data Mining Pengelompokan Menu Makanan dan Minuman Berdasarkan Tingkat Penjualan Menggunakan Metode K-Means," *J. Appl. Informatics Comput.*, vol. 6, no. 1, pp. 40–46, 2022, doi: 10.30871/jaic.v6i1.3824.
- [5] D. K. Utami, N. Irawati, and Sumantri, "Analisis Metode k-Means pada Clustering Penerimaan Bantuan

- PKH Desa Pulau Rakyat Tua,” *Sist. J. Sist. Inf.*, vol. 12, no. 3, pp. 953–961, 2023.
- [6] S. Amaliyah, Jasmir, and S. Rianti, “Penerapan Data Mining Untuk Menentukan Kelompok Prioritas Penerima Bantuan PKH Menggunakan Metode Clustering K-Means Pada Desa Kuala Dendang,” *J. Inform. Dan Rekayasa Komputer(JAKAKOM)*, vol. 3, no. 1, pp. 453–458, 2023, doi: 10.33998/jakakom.2023.3.1.802.
- [7] W. Febrianti and Mustopa Husein Lubis, “Data Mining dalam Mengidentifikasi Calon Penerima Bantuan Satu Keluarga Satu Sarjana (SKSS) dengan Menggunakan Algoritma K-Means,” *J. Inform. Ekon. Bisnis*, vol. 4, pp. 53–58, 2022, doi: 10.37034/infec.v4i2.117.
- [8] R. F. N. Alifah and A. C. Fauzan, “Implementasi Algoritma K-Means Clustering Berbasis Jarak Manhattan untuk Klasterisasi Konsentrasi Bidang Mahasiswa,” *Ilk. J. Comput. Sci. Appl. Informatics*, vol. 5, no. 1, pp. 31–41, 2023, doi: 10.28926/ilkomnika.v5i1.542.
- [9] C. C. Hutagalung, F. A. Syam, Afriansyah, and A. Rozaq, “IMPLEMENTASI ALGORITMA K-MEANS UNTUK MENGENALI POLA,” in *Seminar Nasional Teknologi Informasi & Ilmu Komputer (SEMATER)*, 2023, vol. 2, no. 1, pp. 171–182.
- [10] F. Salsabila, I. Fitrianti, Y. Umaidah, and N. Heryana, “PENERAPAN METODE CRISP-DM UNTUK ANALISA PENDAPATAN BERSIH BULANAN PEKERJA INFORMAL DI PROVINSI JAWA BARAT DENGAN ALGORITMA K-MEANS,” *DINAMIK*, vol. 28, no. September, pp. 64–66, 2023.
- [11] D. P. Utomo, P. Sirait, and R. Yunis, “Reduksi Atribut Pada Dataset Penyakit Jantung dan Klasifikasi Menggunakan Algoritma C5. 0,” *J. Media Inform. Budidarma*, vol. 4, no. 4, pp. 994–1006, 2020, doi: 10.30865/mib.v4i4.2355.
- [12] F. Telaumbanua, J. M. Purba, and D. P. Utomo, “Analysis of Online Learning Understanding Patterns at Budi Darma University Using the C5 . 0 Algorithm,” *IJICS (International J. Informatics Comput. Sci.)*, vol. 5, no. 2, pp. 118–122, 2021, doi: 10.30865/ijics.v5i2.3129.
- [13] D. P. Utomo and S. Aripin, “Penerapan Algoritma C5. 0 Untuk Mengetahui Pola Kepuasan Mahasiswa di Masa Pembelajaran Daring,” 2021.
- [14] N. Lusianah, A. Irma Purnamasari, and B. Nurhakim, “Implementasi Algoritma K-Means Dalam Pengelompokan Jumlah Wisatawan Akomodasi Di Jawa Barat,” *J. Ekonomi, Bisnis dan Manaj.*, vol. 2, no. 1, pp. 254–268, 2023.
- [15] F. Nuraeni, D. Kurniadi, and G. Fauzian Dermawan, “Pemetaan Karakteristik Mahasiswa Penerima Kartu Indonesia Pintar Kuliah (KIP-K) menggunakan Algoritma K-Means++,” *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 11, no. 3, pp. 437–443, 2023, doi: 10.32736/sisfokom.v11i3.1439.
- [16] J. Faran and R. T. Aldisa, “Perbandingan Algoritma K-Means dan K-Medoids Dalam Pengelompokan Kelas Untuk Mahasiswa Baru Program Magister,” *J. Inf. Syst. Res.*, vol. 5, no. 2, pp. 509–519, 2024, doi: 10.47065/josh.v5i2.4753.
- [17] A. S. Ilmananda and H. D. Ranglalin, “Klasterisasi Negara Pengekspor Beras ke Indonesia Menggunakan Algoritma K-Means Clustering,” *J. SISFOTENIKA*, vol. 13, no. 2, pp. 139–150, 2023, [Online]. Available: <http://sisfotenika.stmikpontianak.ac.id/index.php/ST>.
- [18] R. Watrionthos, R. Handayani, A. F. P. Akhir, A. Ambiyar, and U. Verawardina, “Penerapan Algoritma K-Means Pada Pemetaan Kemampuan Penggunaan Teknologi Informasi Remaja dan Dewasa di Indonesia,” *J. Comput. Syst. Informatics*, vol. 4, no. 1, pp. 45–50, 2022, doi: 10.47065/josyc.v4i1.2264.
- [19] T. B. Pamungkas, S. Maesaroh, and P. Ardiansyah, “Implementasi Data Mining Pada Stok Penggunaan Barang Di Gmf Aeroasia Menggunakan Algoritma K-Means Clustering,” *J. Ilm. Sains dan Teknol.*, vol. 7, no. 2, pp. 112–123, 2023, doi: 10.47080/saintek.v7i2.2697.
- [20] Y. Mayona, R. Buaton, and M. Simanjutak, “Data Mining Clustering Tingkat Kejahatan Dengan Metode Algoritma K-Means (Studi Kasus: Kejaksaan Negeri Binjai),” *J. Inform. Kaputama*, vol. 6, no. 3, pp. 2548–9739, 2022.
- [21] R. A. Kurniawan, M. S. Hasibuan, P. Piramida, and R. S. Ramadhan, “Penerapan Algoritma K-Means Untuk Clustering Tempat Makan Di Batubara,” *J. Comput. Sci. Informatics Eng.*, vol. 01, no. 1, pp. 10–18, 2022, doi: 10.55537/cosie.v1i1.27.
- [22] L. Awaliyah, N. Rahaningsih, and R. D. Dana, “IMPLEMENTASI ALGORITMA K-MEANS DALAM ANALISIS CLUSTER KORBAN KEKERASAN DI PROVINSI JAWA BARAT,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 1, pp. 188–195, 2024.
- [23] A. Situmorang, Yusfrizal, and I. Ambarita, “Pengelompokan Data Menggunakan Algoritma K-Means Untuk Meningkatkan Motivasi Belajar Siswa MTS Alwashliyah 48 Kebun Lada,” 2023.
- [24] D. T. Yuliana, M. I. A. Fathoni, and N. Kurniawati, “Penentuan Penerima Kartu Indonesia Pintar KIP Kuliah Dengan Menggunakan Metode K-Means Clustering,” *J. Focus Action Res. Math. (Factor M)*, vol. 5, no. 1, pp. 127–141, 2022, doi: 10.30762/f_m.v5i1.570.
- [25] D. Kurniadi, Y. H. Agustin, H. I. N. Akbar, and I. Farida, “Penerapan Algoritma k-Means Clustering untuk Pengelompokan Pembangunan Jalan pada Dinas Pekerjaan Umum dan Penataan Ruang,” *AITI J. Teknol. Inf.*, vol. 20, no. 1, pp. 64–77, 2023, doi: 10.24246/aiti.v20i1.64-77.
- [26] A. Asmana, Y. Arie Wijaya, and M. Martanto, “Clustering Data Calon Siswa Baru Menggunakan Metode K-Means Di Sekolah Menengah Kejuruan Wahidin Kota Cirebon,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 6,

no. 2, pp. 552–559, 2022, doi: 10.36040/jati.v6i2.5236.

- [27] I. Irawan, U. Rizki, P. M. Jakak, M. B. Prayogi, and M. Rahman, “Penerapan Metode K-Means Clustering Dalam Pengembangan Strategi Promosi Berbasis Data Penerimaan Mahasiswa Baru (Studi Kasus :Universitas Nurul Huda),” *J. Nas. Ilmu Komput.*, vol. 5, no. 1, 2024.