

Penerapan Data Mining Menggunakan Algoritma K- Means Clustering dalam Pengelompokan Penerima Beasiswa KIP Kuliah

Application of Data Mining Using the K-Means Clustering Algorithm in the Grouping of KIP Lecture Scholarship Recipients

Hendra Di Kesuma¹, Syafi'ul Hamidani²

¹Program Studi Sistem Informasi, Universitas UIGM Palembang

²Program Studi Sistem Informasi, STMIK Bina Nusantara Jaya Lubuk Linggau

¹Jl. Jend. Sudirman KM 4 No. 62, 20 Ilir D IV Kec. Ilir Tim I Kota Palembang

²Jl Yos Sudarso No. 97 A Kec Lubuklinggau Timur II Kota Lubuklinggau

E-mail : hendra.dikesuma@uigm.ac.id¹, hamidanipertama@gmail.com²

Abstract

Data mining is a technique for processing data with large data capacity or known as big data. The process used in this technique is known as the knowledge discovery process in the database. Where there are 4 models that can be applied to data mining, one of which is the cluster. Clustering can be done with several algorithms, one of which is k-means clustering. Where this algorithm is used to assist in the process of grouping data based on the shortest distance from the cluster formed. In the KIP Lecture scholarship selection process, an analysis is needed that is able to classify data based on an assessment of the variables used so as to produce groups of data that have very close similarities. Where by grouping the data, participants can be identified who are eligible to receive KIP college scholarships according to predetermined standard criteria. The criteria used are parents' income, number of parents' dependents, written test results. Testing the results of the k-means clustering analysis uses a test application, namely rapidminer 5. Clustering is done to be able to determine college kip applicants who are eligible to receive scholarships. From the results of the analysis and testing, it was found that the number of participants was based on predetermined clusters, namely in cluster 1 there were 6 people with disqualified status, cluster 2 with 7 people with reserve status and cluster 3 with 7 people passed. From the results of these recommendations the tertiary institution can easily determine applicants who are entitled to get a KIP college scholarship.

Keywords: Data Mining, K-Means Clustering, Scholarship, clustering

Abstrak

Data mining merupakan teknik dalam melakukan pengolahan data dengan kapasitas data yang besar atau dikenal dengan bigdata. Proses yang digunakan pada teknik ini dikenal dengan proses knowledge discovery in database. Dimana terdapat 4 model yang dapat diterapkan pada data mining salah satunya adalah cluster. Pengelompokan dapat dilakukan dengan beberapa algoritma salah satunya adalah k-means clustering. Dimana algoritma ini digunakan untuk membantu dalam proses pengelompokan data berdasarkan jarak terdekat dari cluster yang dibentuk. Pada proses seleksi beasiswa KIP Kuliah dibutuhkan analisa yang mampu mengelompokan data berdasarkan penilaian terhadap variabel yang digunakan sehingga menghasilkan kelompok data yang memiliki kesamaan yang sangat dekat. Dimana dengan dilakukan pengelompokan data maka dapat diketahui peserta yang layak menerima beasiswa KIP kuliah sesuai dengan standar kriteria yang telah ditetapkan. Adapun kriteria yang digunakan adalah penghasilan orang tua, jumlah tanggungan orang tua, hasil tes tertulis. Pengujian hasil analisa k-means clustering menggunakan aplikasi pengujian yaitu rapidminer 5. Pengelompokan dilakukan untuk dapat menentukan pendaftar kip kuliah yang layak menerima beasiswa. Dari hasil analisa dan pengujian didapatkan jumlah peserta berdasarkan kluster yang telah ditentukan yaitu pada cluster 1 berjumlah 6 orang dengan status tidak lulus, kluster 2 berjumlah 7 orang status cadangan dan kluster 3 berjumlah 7 orang dinyatakan lulus. Dari hasil rekomendasi tersebut pihak perguruan tinggi dapat dengan mudah dalam menetapkan pendaftar yang berhak mendapatkan beasiswa KIP kuliah.

Kata kunci: Data Mining, K-Means Clustering, Beasiswa, pengelompokan

1. Pendahuluan

Beasiswa merupakan salah satu fasilitas yang di berikan pemerintah untuk membantu siswa ataupun mahasiswa

dalam menempuh pendidikan. Beasiswa diperuntukan untuk semua level pendidikan yang ada di Indonesia. Tujuan adanya beasiswa pendidikan adalah untuk memberikan kesempatan dan kesamarataan baik putra

Jurnal Ilmiah Binary STMIK Bina Nusantara Jaya

Vol. 05 No. 01 Tahun 2023, ISSN : 2657- 2117 | DOI : 10.52303/jb.v5i1.102

putri Indonesia untuk mendapatkan pendidikan yang layak. Saat ini bagi calon mahasiswa yang ingin melanjutkan pendidikan tinggi pada level perguruan tinggi pada program vokasi dan akademik yaitu dapat memanfaatkan peluang untuk memperoleh beasiswa pendidikan atau yang lebih dikenal dengan beasiswa KIP Kuliah. Dimana beasiswa ini diberikan kepada mahasiswa dalam bentuk biaya kuliah dan biaya hidup. Adapun peruntukan beasiswa ini dimulai mahasiswa kuliah sampai dengan selesai masa studi yaitu sebanyak 8 semester. Saat ini di STMIK Bina Nusantara Jaya Lubuk Linggau telah melakukan proses seleksi penerimaan mahasiswa baru melalui jalur beasiswa KIP Kuliah. Tingginya peminat melalui jalur beasiswa menjadi kendala pada bagian pengelola beasiswa KIP kuliah dikarenakan jumlah ataupun kuota yang diberikan kepada masing-masing perguruan tinggi terbatas. Maka perlu dilakukan analisa yang tepat untuk merekomendasikan calon mahasiswa baru yang berhak menerima beasiswa KIP kuliah berdasarkan peraturan dan persyaratan yang berlaku. Saat ini proses seleksi yang dilakukan sebatas verifikasi dan validasi secara manual terhadap berkas pendaftar yang masuk. Kemudian dilihat berdasarkan tingkat kriteria yang paling penting yaitu hasil tes seleksi dan penghasilan orang tua. Proses tersebut belum menggambarkan hasil yang tepat karena belum dilakukan analisa data untuk mengelompokkan data yang tepat untuk mengetahui penerima beasiswa KIP. Oleh karena ini penelitian ini dilakukan bertujuan untuk melakukan analisa dengan menerapkan data mining menggunakan algoritma k-means clustering untuk mengetahui dan menghasilkan kelompok data yang tepat sebagai penerima beasiswa KIP kuliah berdasarkan kriteria dan sesuai dengan jumlah peserta yang mendaftar. Sehingga memberikan kemudahan dan ketepatan bagi pihak perguruan tinggi dalam melakukan proses seleksi penerima beasiswa.

2. Tinjauan Pustaka

2.1 Analisis Penelitian Terdahulu

Berdasarkan analisa penelitian terdahulu menurut mutaqin tentang pengelompokan judul skripsi mahasiswa dimana pengelompokan mahasiswa menggunakan metode clustering dimana Clustering ini bertujuan untuk membagi mahasiswa ke dalam cluster berdasarkan nilai yang diperolehnya dari semester 1 s/d 7, sehingga dapat menghasilkan rekomendasi mahasiswa dalam mengambil topik skripsi[1] sedangkan menurut sulistiyowati tentang pengelompokan siswa kelas unggulan. Dimana pengelompokan kelas unggulan dilakukan untuk memudahkan fasilitasi pendidikan siswa berdasarkan perbedaan kemampuannya dalam belajar dan mengikuti pembelajaran[2]. Selain itu penelitian yang dilakukan semiring yaitu pengelompokan siswa berprestasi dimana penelitian ini bertujuan untuk membangun sistem yang dapat mempermudah sekolah dengan cepat dan tepat dalam pengelompokan prestasi siswa[3]

2.2. Data Mining

Data mining merupakan metode yang digunakan dalam pengolahan data berskala besar. Data yang diolah dengan teknik data mining ini kemudian menghasilkan suatu pengetahuan baru, hasil dari pengolahan data tersebut dapat digunakan dalam menentukan keputusan. Dalam data mining juga terdapat metode - metode yang dapat digunakan seperti Classification, Clustering, Estimation, Prediction, Association [4][5]

2.3. K-Means Clustering

Berbagai algoritma cerdas untuk menganalisa telah banyak diimplementasikan. Salah satunya adalah algoritma K-Means Clustering yang merupakan algoritma pengelompokan iteratif yang sederhana diimplementasikan, relatif cepat, dan mudah beradaptasi [6][7]. K-means clustering merupakan salah satu algoritma machine learning yang sederhana dan populer digunakan untuk memecahkan pengelompokan data. Dalam machine learning, k-means clustering termasuk kedalam jenis algoritma unsupervised learning[8][9]

Adapun tahapan penerapan algoritma K-Means Clustering adalah sebagai berikut [10]

- Menentukan Nilai K sebanyak jumlah cluster atau kelompok yang diinginkan
- Pilih sebanyak K data dari set data sebagai pusat cluster (centroid) secara acak
- Menghitung jarak antara objek dengan masing masing centroid. Bisa menggunakan persamaan Euclidean Distance

$$d(x_i - x_j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + (x_{i3} - x_{j3})^2}$$

- Mengelompokkan objek berdasarkan jarak terdekat dengan centroid
- Menentukan centroid baru dengan menggunakan persamaan

$$c_{m(q)} = \frac{1}{n_m} \sum_{i=1}^{n_m} X_{i(q)}$$

- Ulangi langka 3 dan 4 hingga tidak ada lagi objek yang berpindah cluster

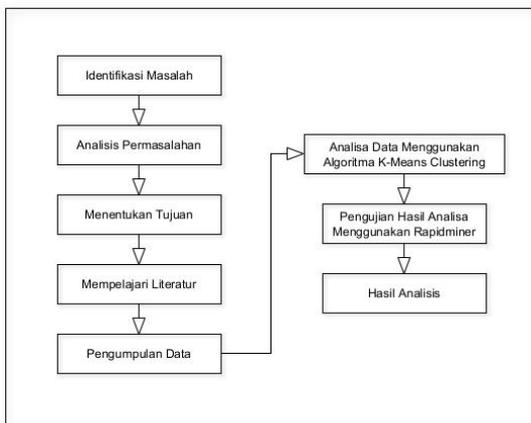
2.3. Rapidminer

RapidMiner adalah platform perangkat lunak data ilmu pengetahuan yang dikembangkan oleh perusahaan yang bernama yang sama, yang menyediakan lingkungan terpadu untuk pembelajaran mesin (machine learning), pembelajaran mendalam (deep learning), penambangan teks (text mining), dan analisis prediktif (predictive analytics)[11]

3. Metodologi Penelitian

3.1 Kerangka Penelitian

Rangkaian kegiatan penelitian dapat dilakukan melalui tahapan penelitian yang dilakukan. Adapun tahapan penelitian ini adalah seperti pada gambar 1 sebagai berikut[12].



Gambar 1. Tahapan Penelitian

- Pada tahapan awal dilakukan identifikasi terhadap masalah yang ada dan melakukan proses studi literatur terhadap beberapa sumber diantaranya artikel, buku dan karya ilmiah lainnya. Adapun tujuan ini dilakukan untuk mendapatkan informasi yang dibutuhkan.
- Tahapan selanjutnya adalah menganalisa permasalahan yaitu melakukan pemahaman terhadap permasalahan pada objek penelitian sesuai dengan batasan yang telah ditetapkan sehingga hasil penelitian dapat sesuai dengan yang diharapkan
- Menentukan tujuan penelitian yaitu menetapkan arah penelitian yang jelas agar hasil yang diperoleh dari penelitian dapat bermanfaat bagi pengguna.
- Selanjutnya adalah mempelajari studi literatur berkenaan dengan data mining, KDD, algoritma K-Means Clustering dan sumber data lainnya sesuai dengan kebutuhan penelitian. Agar penelitian dapat dilakukan sesuai dengan target dan tepat sasaran.
- Pengumpulan data dilakukan untuk memperoleh data utama dengan cara melakukan observasi dan wawan cara pada bagian kemahasiswaan adapun data utama yang diperlukan adalah data berkenaan dengan seleksi penerima beasiswa KIP Kuliah.
- Tahap analisa data dilakukan dengan menggunakan algoritma yang dipilih yaitu K-Means-Clustering dengan menerapkan 6 tahapan pada algoritma tersebut.
- Dari hasil analisa data menggunakan algoritma K-Means Clustering selanjutnya dilakukan pengujian terhadap hasil analisa menggunakan algoritma dengan menerapkan aplikasi pengujian yaitu aplikasi rapidminer 5

- Tahapan akhir adalah hasil analisis adalah melihat hasil analisis menggunakan algoritma dan menggunakan aplikasi pengujian untuk melihat kecocokan hasil yang diperoleh sebagai dasar yang dapat digunakan top manajemen memilih peserta yang layak menerima beasiswa KIP Kuliah berdasarkan hasil pengelompokan yang terbentuk.

4. Hasil dan Pembahasan

Tahapan awal dilakukan dengan memilih dan menetapkan data yang diperoleh dari bagian kemahasiswaan STMIK BNJ Lubuklinggau, kemudian dilakukan proses pre-processing agar dapat dilakukan proses transformasi data pendaftaran KIP Kuliah

Tabel 1. Data Pendaftar KIP Kuliah

NoDaftar	Penghasilan	Tanggungan	tes tertulis
DF01	1 Juta	2 Orang	40
DF02	2 Juta	5 Orang	70
DF03	2 Juta	4 Orang	50
DF04	3 Juta	2 Orang	60
DF05	1 Juta	2 Orang	40
DF06	1 Juta	3 Orang	70
DF07	2 Juta	3 Orang	70
DF08	2 Juta	1 Orang	70
DF09	3 Juta	2 Orang	80
DF10	3 Juta	2 Orang	60
DF11	2 Juta	5 Orang	60
DF12	2 Juta	5 Orang	60
DF13	2 Juta	5 Orang	50
DF14	2 Juta	3 Orang	90
DF15	1 Juta	1 Orang	70
DF16	5 Juta	1 Orang	50
DF17	2 Juta	3 Orang	40
DF18	2 Juta	4 Orang	40
DF19	2 Juta	3 Orang	70
DF20	1 Juta	2 Orang	60

Kemudian data diolah menggunakan algoritma k-means clustering.

4.1 Memasukan data yang telah ditransformasi

Adapun data sampel yang digunakan dari hasil transformasi data adalah seperti pada tabel 1 sebagai berikut :

Tabel 2. Data Set yang telah ditransformasi

NoDaftar	Penghasilan	Tanggungan	tes tertulis
DF01	1	2	4
DF02	2	5	7
DF03	2	4	5
DF04	3	2	6
DF05	1	2	4
DF06	1	3	7
DF07	2	3	7
DF08	2	1	7
DF09	3	2	8
DF10	3	2	6
DF11	2	5	6
DF12	2	5	6

DF13	2	5	5
DF14	2	3	9
DF15	1	1	7
DF16	5	1	5
DF17	2	3	4
DF18	2	4	4
DF19	2	3	7
DF20	1	2	6

4.2 Penentuan Jumlah Kluster

Pada penelitian ini penulis penentuan 3 cluster yang digunakan untuk pengelompokan peserta seleksi beasiswa KIP Kuliah yaitu Tidak Lulus, Cadangan dan Lulus seperti pada tabel 3 sebagai berikut.

Tabel 3. Data Kluster

Cluster	Min	Max	Cluster
1	0	7	Tidak Lulus
2	8	13	Cadangan
3	14	20	Lulus

4.3 Menentukan Centroid dari data set

Kemudian tahapan selanjutnya adalah menentukan centroid atau titik pusat yang diperoleh secara acak dari data set yang ada seperti pada tabel 3,4,5 sebagai berikut.

Tabel 4. Data Centroid 1

NoDaftar	Penghasilan	Tanggung	tes tertulis
DF01	1	2	4
DF05	1	2	4
Cluster 1	1	2	4

Tabel 5. Data Centroid 2

NODaftar	Penghasilan	Tanggung	Tes tertulis
DF03	2	4	5
DF04	3	2	6
DF06	1	3	7
DF07	2	3	7
DF08	2	1	7
DF09	3	2	8
DF10	3	2	6
DF11	2	5	6
DF12	2	5	6
DF13	2	5	5
DF15	1	1	7
DF16	5	1	5
DF17	2	3	4
DF18	2	4	4
DF19	2	3	7
DF20	1	2	6
Cluster 2	2,1875	2,875	6

Tabel 6. Data Centroid 3

NoDaftar	Penghasilan	Tanggung	tes tertulis
DF02	2	5	7
DF14	2	3	9
Cluster 3	2	4	8

Kemudian menghitung jarak terdekat atau minimum sehingga dikelompokan berdasarkan jarak terdekat terhadap centroid. Adapun nilai centroid yang diperoleh seperti pada tabel 2 sebagai berikut

Tabel 7. Hasil Perhitungan jarak minimum (nilai centroid)

NoDaftar	Penghasilan Ortu	Tanggung Ortu	nilai tes tertulis	C1	C2	C3	Min	Cluster
DF01	1	2	4	0	2,48511191 9	4,58257 6	0	1
DF02	2	5	7	4,3588989 4	2,35600960 3	1,41421 4	1,41421356 2	3
DF03	2	4	5	2,4494897 4	1,51683263 7	3	1,51683263 7	2
DF04	3	2	6	2,8284271 2	1,19406082 3	3	1,19406082 3	2
DF05	1	2	4	0	2,48511191 9	4,58257 6	0	1
DF06	1	3	7	3,1622776 6	1,55749197 4	1,73205 1	1,55749197 4	2
DF07	2	3	7	3,3166247 9	1,02507621 7	1,41421 4	1,02507621 7	2
DF08	2	1	7	3,3166247 9	2,13325602 1	3,16227 8	2,13325602 1	2
DF09	3	2	8	4,4721359 5	2,32933064 4	2,23606 8	2,23606797 7	2
DF10	3	2	6	2,8284271 2	1,19406082 3	3	1,19406082 3	2

DF11	2	5	6	3,7416573 9	2,13325602 1	2,23606 8	2,13325602 1	2
DF12	2	5	6	3,7416573 9	2,13325602 1	2,23606 8	2,13325602 1	2
DF13	2	5	5	3,3166247 9	2,35600960 3	3,16227 8	2,35600960 3	2
DF14	2	3	9	5,1961524 2	3,00845163 7	1,41421 4	1,41421356 2	3
DF15	1	1	7	3,1622776 6	2,43429276 2	3,31662 5	2,43429276 2	2
DF16	5	1	5	4,2426406 9	3,52502216 3	5,19615 2	3,52502216 3	2
DF17	2	3	4	1,4142135 6	2,01265527 4	4,12310 6	1,41421356 2	1
DF18	2	4	4	2,2360679 8	2,30234255 7	4 7	2,23606797 7	1
DF19	2	3	7	3,3166247 9	1,02507621 7	1,41421 4	1,02507621 7	2
DF20	1	2	6	2 5	1,47505296 5	3	1,47505296 5	2

4.4 Menghitung Jarak Terdekat

Kemudian lakukan proses perhitungan dengan nilai centroid yang baru dengan cara menjumlahkan data peserta dibagi dengan total peserta disetiap centroid sehingga diperoleh jarak terdekat seperti pada tabel 8 sebagai berikut :

Tabel 8. Data Centroid Baru

	X1	X2	X3
Cluster 1	1,5	2,75	4
Cluster 2	2,2142857	2,7857143	6,285714286
Cluster 3	2	4	8

4.5 Menghitung kembali pusat kluster dengan keanggotaan kluster yang baru

Pada tahapan ini dilakukan untuk mengetahui nilai pusat kluster apakah terjadi perubahan atau tidak. Jika terjadi perubahan maka dilakukan perhitungan kembali pusat kluster yang baru dan jika nilainya tetap maka proses kluster selesai dilakukan sehingga menghasilkan nilai kluster seperti pada tabel 9 sebagai berikut :

Tabel 9. Hasil Perhitungan Jarak Minimum (Nilai Centroid)

Nim	Penghasilan Ortu	Tanggung n Ortu	nilai tes tertulis	C1	C2	C3	Min	Clust er	keterangan
DF01	1	2	4	2,70487089	0,90138781 9	4,58257 6	0,9013 88	2	tetap
DF02	2	5	7	1,414214	2,33648960 5	3,78318 6488	1,4142 14	1	tetap
DF03	2	4	5	1,677050983	1,78142341 6	3 6	1,6770 51	1	berubah
DF04	3	2	6	2,610076627	1,14731274 3	3	1,1473 13	2	tetap
DF05	1	2	4	2,70487089	0,90138781 9	4,58257 6	0,9013 88	2	tetap
DF06	1	3	7	3,051638904	1,732051	1,42499 5525	1,4249 96	3	tetap
DF07	2	3	7	3,051638904	1,414214	0,77591 2892	0,7759 13	3	tetap
DF08	2	1	7	3,508917212	1,93517388 3	3,16227 8	1,9351 74	3	tetap
DF09	3	2	8	4,337337893	2,04290709 2	2,23606 8	2,0429 07	3	tetap
DF10	3	2	6	2,610076627	1,14731274 3	3	1,1473 13	2	berubah
DF11	2	5	6	2,236068	2,24290263 8	3,05163 8904	2,2360 68	1	berubah
DF12	2	5	6	2,236068	2,24290263 8	3,05163 8904	2,2360 68	1	berubah
DF13	2	5	5	2,512468905	2,56944367 8	3,16227 8	2,5124 69	1	berubah

DF14	2	3	9	5,031152949	2,731150613	1,414214	1,414214	3	tetap
DF15	1	1	7	3,508917212	3,316625	2,27452619	2,274526	3	tetap
DF16	5	1	5	4,038873605	3,549935326	5,196152	3,549935	2	tetap
DF17	2	3	4	2,305716056	0,559016994	4,123106	0,559017	2	tetap
DF18	2	4	4	1,346291202	2,597094138	4	1,346291	1	tetap
DF19	2	3	7	3,051638904	1,414214	0,775912892	0,775913	3	tetap
DF20	1	2	6	2,193741097	1,474269103	3	1,474269	2	tetap

Sampai dengan iterasi ke lima dengan proses nilai kluster pada iterasi keempat dengan menghitung perhitungan nilai centroid baru sehingga didapatkan nilai centroid baru sehingga didapatkan hasil hasil pengelompokan data dan tidak ada perubahan pada pengelompokan seperti pada tabel 10 sebagai berikut.

Tabel 10. Hasil Perhitungan Jarak Minimum (nilai centroid) Iterasi 5

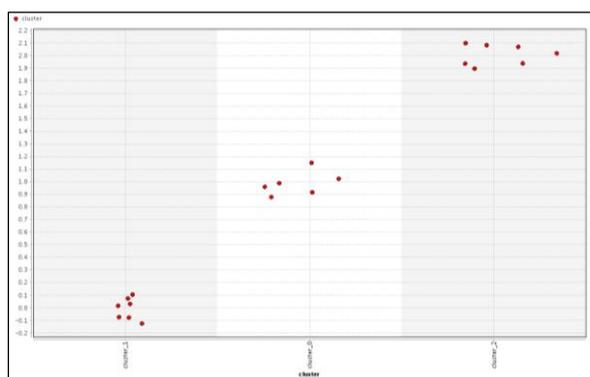
NoDaftar	Penghasil an Ortu	Tanggung an Ortu	nilai tes tertulis	C1	C2	C3	Min	Clust er	keterang an
DF01	1	2	4	3,093581732	1,183215957	3,741657	1,183216	2	tetap
DF02	2	5	7	1	2,92749858	3,464101615	1	1	tetap
DF03	2	4	5	1,341640786	2,650432679	1,414214	1,341641	1	tetap
DF04	3	2	6	2,48997992	1,098458725	3,162278	1,098459	2	tetap
DF05	1	2	4	3,093581732	1,183215957	3,741657	1,183216	2	tetap
DF06	1	3	7	2,863564213	1,574591643	2,44949	1,574592	3	tetap
DF07	2	3	7	2,828427125	2,236068	0,966376892	0,966377	3	tetap
DF08	2	1	7	3,464101615	1,139087644	4,123106	1,139088	3	tetap
DF09	3	2	8	4,171330723	3,741657	1,390641686	1,390642	3	tetap
DF10	3	2	6	2,48997992	1,098458725	3,162278	1,098459	2	tetap
DF11	2	5	6	0	3,034239867	2,720294102	0	1	tetap
DF12	2	5	6	0	3,034239867	2,720294102	0	1	tetap
DF13	2	5	5	1	3,441362406	2,19089023	1	1	tetap
DF14	2	3	9	4,816637832	3,605551	2,379318605	2,379319	3	tetap
DF15	1	1	7	3,492849839	1,686112454	4,242641	1,686112	3	tetap
DF16	5	1	5	4,024922359	3,454545455	5,09902	3,454545	2	tetap
DF17	2	3	4	2,973714043	0,447213595	2,828427	0,447214	2	tetap
DF18	2	4	4	1,095445115	3,414843631	2,236068	1,095445	1	tetap
DF19	2	3	7	2,828427125	2,236068	0,966376892	0,966377	3	tetap
DF20	1	2	6	2,144761059	1,515757455	3,162278	1,515757	2	tetap

4.6. Pengujian Hasil Analisa

Dari hasil analisa perhitungan menggunakan algoritma K-Means Clustering kemudian dilakukan pengujian hasil menggunakan aplikasi rapidminer 5.0. dimana data yang digunakan adalah data pendaftar beasiswa KIP Kuliah pada tahun 2022. Kemudian diperoleh hasil pengujian seperti pada Gambar 2 dan 3 sebagai berikut

Row No.	cluster	Penghasilan...	Tanggunga...	nilai tes terf...
1	cluster_1	1	2	4
2	cluster_0	2	5	7
3	cluster_0	2	4	5
4	cluster_1	3	2	6
5	cluster_1	1	2	4
6	cluster_2	1	3	7
7	cluster_2	2	3	7
8	cluster_2	2	1	7
9	cluster_2	3	2	8
10	cluster_1	3	2	6
11	cluster_0	2	5	6
12	cluster_0	2	5	6
13	cluster_0	2	5	5
14	cluster_2	2	3	9
15	cluster_2	1	1	7
16	cluster_1	5	1	5
17	cluster_1	2	3	4
18	cluster_0	2	4	4
19	cluster_2	2	3	7
20	cluster_1	1	2	6

Gambar 2. Data View 3 Cluster



Gambar 3. Hasil Cluster Menggunakan Rapidminer

Dari hasil cluster pada gambar menunjukkan bahwa jumlah pendaftar yang tergabung pada cluster 1 berjumlah 6 orang, cluster 2 berjumlah 7 orang dan cluster 3 berjumlah 7 orang. Berdasarkan cluster yang terbentuk bahwa pendaftar yang direkomendasikan menerima beasiswa KIP Kuliah adalah pendaftar yang tergabung pada cluster 3 yaitu berjumlah 7 orang.

5. Kesimpulan

5.1 Simpulan

Pada penelitian ini dapat ditarik kesimpulan bahwa algoritma K-Means Clustering dapat digunakan dalam melakukan analisa pengelompokan data penerima beasiswa KIP Kuliah dimana dari data set diperoleh hasil analisa dan pengujian menggunakan aplikasi rapidminer bahwa mahasiswa yang dinyatakan lulus seleksi beasiswa berjumlah 7 orang yaitu pada cluster 3.

Sedangkan mahasiswa yang tidak lulus seleksi berjumlah 6 orang pada cluster 1 sedangkan pada cluster 2 mahasiswa dengan status cadangan sebagai penerima beasiswa KIP Kuliah.

5.2 Saran

Penulis menyarankan agar penelitian ini dapat perlu dilakukan pembuatan sistem dalam bentuk aplikasi pengelompokan seleksi penerima beasiswa KIP Kuliah menggunakan bahasa pemrograman berbasis web dan untuk variabel yang digunakan perlu dilakukan uji validitas sehingga variabel yang dipilih benar benar tepat dan sesuai dengan kebutuhan pengguna.

Daftar Rujukan

- [1] M. R. Muttaqin and M. Defriani, "Algoritma K-Means untuk Pengelompokan Topik Skripsi Mahasiswa," *Ilk. J. Ilm.*, vol. 12, no. 2, pp. 121–129, 2020, doi: 10.33096/ilkom.v12i2.542.121-129.
- [2] A. Sulistiyawati and E. Supriyanto, "Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan," *J. Tekno Kompak*, vol. 15, no. 2, p. 25, 2021, doi: 10.33365/jtk.v15i2.1162.
- [3] S. N. Br Sembiring, H. Winata, and S. Kusnari, "Pengelompokan Prestasi Siswa Menggunakan Algoritma K-Means," *J. Sist. Inf. Triguna Dharma (JURSI TGD)*, vol. 1, no. 1, p. 31, 2022, doi: 10.53513/jursi.v1i1.4784.
- [4] N. Mirantika, "Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Penyebaran Covid-19 di Provinsi Jawa Barat," *Nuansa Inform.*, vol. 15, no. 2, pp. 92–98, 2021, doi: 10.25134/nuansa.v15i2.4321.
- [5] E. Muningsih, I. Maryani, and V. R. Handayani, "Penerapan Metode K-Means dan Optimasi Jumlah Cluster dengan Index Davies Bouldin untuk Clustering Propinsi Berdasarkan Potensi Desa," *J. Sains dan Manaj.*, vol. 9, no. 1, pp. 95–100, 2021, [Online]. Available: <https://ejournal.bsi.ac.id/ejurnal/index.php/evolusi/article/view/10428/4839>
- [6] H. Haviluddin, S. J. Patandianan, G. M. Putra, N. Puspitasari, and H. S. Pakpahan, "Implementasi Metode K-Means Untuk Pengelompokan Rekomendasi Tugas Akhir," *Inform. Mulawarman J. Ilm. Ilmu Komput.*, vol. 16, no. 1, p. 13, 2021, doi: 10.30872/jim.v16i1.5182.
- [7] 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.
- [8] C. A. Sugianto, A. H. Rahayu, and A. Gusman, "Algoritma K-Means untuk Pengelompokan Penyakit Pasien pada Puskesmas Cigugur Tengah," *J. Inf. Technol.*, vol. 2, no. 2, pp. 39–44, 2020, doi: 10.47292/joint.v2i2.30.
- [9] 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.
- [10] 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.
- [11] N. Nur, "Penerapan Metode Clustering dengan Algoritma K-Means untuk Pengelompokan Data Migrasi Penduduk Tiap Kecamatan di Kabupaten Rembang," vol. 6, pp. 729–738, 2023.
- [12] R. Yanto and R. Khoiriah, "Implementasi Data Mining dengan Metode Algoritma Apriori dalam Menentukan Pola Pembelian Obat," *Citec J.*, 2015, doi: 2354-5771.