

Implementasi Fuzzy C-Means Clustering Dalam Proses Seleksi Penerimaan Kartu Indonesia Pintar Kuliah

Divangga Revansa Arya Pradhana^{1*}, Lola Sekar Arum², Bagas Muttaqi³, Dwi Hartanti⁴

^{1,2,3,4}Teknik Informatika

Universitas Duta Bangsa Surakarta

¹*202021046@mhs.udb.ac.id, ²202030337@mhs.udb.ac.id, ³202021173@mhs.udb.ac.id, ⁴dwihartanti@udb.ac.id

Abstrak— Logika fuzzy merupakan salah satu dari banyaknya logika pembentuk soft computing. Merupakan peningkatan dari logika Boolean yang berhadapan dengan konsep kebenaran sebagian. Saat logika klasik menyatakan bahwa segala hal dapat diekspresikan dalam istilah biner (0 atau 1, hitam atau putih, ya atau tidak), logika fuzzy menggantikan kebenaran boolean dengan tingkat kebenaran. Dalam penelitian filterisasi penerimaan KIP-Kuliah ini logika fuzzy yang digunakan yaitu fuzzy c-means clustering. Dimana untuk menerima beasiswa KIP-K diperlukan beberapa kriteria yang harus terpenuhi. Adapun kriteria yang perlu dipenuhi diantaranya adalah prestasi, nilai ujian, penghasilan orang tua, penerima kartu sejenis KIP, dan status kepemilikan rumah atau aset. Dengan demikian, diperlukan sistem pendukung keputusan yang dapat membantu untuk menentukan penerimaan beasiswa tersebut berdasarkan kriteria-kriteria minimal yang telah ditentukan, guna membantu pengambilan keputusan yang lebih akurat, efektif, efisien dan hemat waktu.

Kata kunci— Fuzzy C-Means, Data Mining, Data Science, KIP-K

Abstract— Fuzzy logic is one of the many logics that build soft computing. It is an improvement of Boolean logic that deals with the concept of partial truth. While classical logic states that everything can be expressed in binary terms (0 or 1, black or white, yes or no), fuzzy logic replaces boolean truth with degrees of truth. In this KIP-K admission filtering research, the fuzzy logic used is fuzzy c-means clustering. Where the way to receive a KIP-K scholarship, several criteria must be fulfilled. Some of the criteria that need to be fulfilled such as is achievement, test scores, parents income, recipients of KIP-like cards, and home or asset ownership status. Thus, a decision support system is needed that can help determine the acceptance of these scholarships based on predetermined minimum criteria, in order to help make more accurate, effective, efficient and time-saving decisions.

Keywords— Fuzzy C-Means, Data Mining, Data Science, KIP-K

I. PENDAHULUAN

Beasiswa KIP-Kuliah adalah suatu bentuk apresiasi kepada calon mahasiswa yang memiliki keterbatasan ekonomi untuk melanjutkan ke jenjang pendidikan yang hendak ditempuh. Pemberian beasiswa kepada mahasiswa dilaksanakan secara selektif, dimana mahasiswa calon penerimanya harus memenuhi beberapa syarat yang telah ditentukan oleh kemendikbud ristekdikti yang telah bekerjasama juga dengan kemensos. Syarat yang diperlukan untuk mendukung pengambilan keputusan atau seleksi diantaranya ialah prestasi, nilai akhir atau ujian, penghasilan orang tua, kepemilikan aset, dan kepemilikan kartu sejenis KIP.

Proses seleksi yang dilakukan sebelumnya yaitu menggunakan metode manual atau konvensional yaitu dengan memasukkan data-data ke dalam file spreadsheet atau excel kemudian dilakukan filtering atau sorting data

mahasiswa secara manual. Hal tersebut seringkali menimbulkan beberapa permasalahan, sebagai contoh yaitu kerancuan data, penerimaan beasiswa tidak akurat, kesalahan penginputan data dikarenakan kekurangtelitian admin, serta membutuhkan waktu seleksi yang cukup lama dan rumit. Selain itu transparansi dalam penerimaan beasiswa juga tidak jelas. Untuk itu dibutuhkan suatu sistem pendukung keputusan dalam melakukan proses seleksi penerimaan beasiswa KIP-K tersebut.

”Referensi[1],[2],[3],[4] menunjukkan bahwa metode fuzzy c-means clustering, dapat digunakan untuk penentuan jurusan, beasiswa, pengelompokan kualitas sekolah, dan rekomendasi bidang keahlian”. Dalam penelitian dengan tema yang serupa dengan referensi tersebut, untuk mengolah data mahasiswa calon penerima beasiswa metode FCM atau Fuzzy C-Means juga diterapkan dalam proses seleksinya. Sistem pendukung keputusan ini diharapkan dapat membantu proses seleksi

supaya lebih efektif, efisien, akurat, dan transparan.

Langkah awal dari Fuzzy C-Means clustering adalah menentukan pusat cluster, yang nantinya digunakan untuk menandai titik rata-rata untuk tiap-tiap cluster. Pada kondisi awal, pusat cluster data yang dimiliki terhitung belum akurat. Tiap-tiap titik data memiliki derajat keanggotaan untuk tiap-tiap cluster. Dengan menentukan pusat cluster baru dan derajat keanggotaan tiap-tiap titik data secara berulang, maka akan dapat diketahui bahwa pusat cluster akan bergerak menuju target titik yang akurat. Output dari Fuzzy C-Means merupakan deretan pusat cluster dan beberapa derajat keanggotaan untuk tiap-tiap titik data. Dalam penelitian akan dibahas bagaimana proses sorting penerimaan beasiswa kepada mahasiswa yang direkomendasikan menerimanya berdasarkan kriteria yang telah ditentukan.

II. METODOLOGI PENELITIAN

A. Tahapan Penelitian

Dalam melaksanakan suatu penelitian, dibutuhkan beberapa metode atau tahapan penelitian sebagai kerangka kerja agar penelitian yang dilaksanakan berjalan dengan lancar. Berikut tahapan yang dilakukan oleh peneliti :

- 1) Pengumpulan data : Metode ini dilakukan dengan cara mencari sumber data asli atau sumber data (*private/public*), menggunakan situs dataset [10]
- 2) Studi Pustaka : Penulisan dilakukan dengan cara mempelajari berbagai macam referensi pustaka yang berkaitan dengan *Fuzzy C-Means*, terutama *journal*.
- 3) Implementasi (*Implementation*) : Implementasi ini dilakukan setelah data diolah menggunakan *Microsoft Excel*, menggunakan *Software RapidMiner Studio*.

B. Fuzzy C-Means

Fuzzy C-Means merupakan pengembangan dari logika Algoritma fuzzy yang konsepnya ditemukan oleh Lotfi Zadeh[6], dengan dasar pemikiran bahwa tidak ada keadaan yang bersifat *true* atau *off*, setiap hasil keluaran sistem pastinya terdapat nilai gradasi di antara *true* or *off* dengan

cara melakukan pergeseran skala variabel yang dapat diukur sebagai bagian dari '*true*' atau bagian dari '*false*'. Jika dalam[5] algoritma yang diterapkan ialah algoritma dasar dari logika fuzzy itu sendiri, pada tahap penelitian ini digunakan yaitu algoritma *Fuzzy C-Means* yang dimana FCM sendiri merupakan proses untuk menentukan *cluster* berdasarkan pusat *cluster* dan derajat keanggotaan antara rentang 0 hingga 1. *Fuzzy C-Means* adalah suatu teknik *clustering* data dimana tiap datanya ditentukan oleh jarak keanggotaan dari tiap data. Berikut langkah-langkah perhitungan *Fuzzy C-Means* :

- 1) Memasukkan data yang akan di klaster X berupa matriks $n \times m$ (n = Jumlah data, m = Atribut data). X_{ij} = Data sample ke-I ($i = 1, 2, \dots, n$), atribut ke-j ($j = 1, 2, \dots, m$).
- 2) Tentukan : jumlah ***cluster*** = c , pangkat **pembobot** = w , **maksimum iterasi** = **maxIter**, **error terkecil yang diharapkan** = ζ , **fungsi objektif awal** = $P = 0$, dan **iterasi awal** = $t = 0$.
- 3) Bangkitkan nilai acak μ_{ik} sebagai elemen matriks partisi awal u . μ_{ik} adalah nilai derajat keanggotaan yang didapatkan secara random dengan interval 0 sampai dengan 1. Matriks partisi awal ini belum akurat, untuk itu perlu dilakukan perulangan sampai mendapatkan derajat keanggotaan dan pusat cluster yang akurat.
- 4) Hitung titik pusat *cluster* ke-k

$$V_{ki} = \frac{\sum_{i=1}^n ((\mu_{ik})^w * X_{ij})}{\sum_{i=1}^n (\mu_{ik})^w}$$

Keterangan :

V_{ki} : titik pusat *cluster*

X_{ij} : data

μ_{ik} :derajat keanggotaan data pada tiap *cluster*

- 5) Hitung fungsi objektif (P_t)

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\left[\sum_{i=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^w \right)$$

Keterangan :

P_t : titik pusat *cluster*

V_{ki} : titik pusat *cluster*

X_{ij} : data

μ_{ik} :derajat keanggotaan data pada tiap *cluster*

6) Hitung perubahan matriks partisi U

Perubahan matriks partisi U yang akan digunakan sebagai keanggotaan baru pada iterasi seterusnya.

$$\mu_{ik} = \frac{\left[\sum_{j=1}^m (x_{ij} - v_{kj})^2 \right]^{\frac{-1}{w-1}}}{\sum_{k=1}^c \left[\sum_{j=1}^m (x_{ij} - v_{kj})^2 \right]^{\frac{-1}{w-1}}}$$

Dengan $i = 1, 2, \dots, n$; dan $k = 1, 2, \dots, c$.

V_{ki} : titik pusat *cluster*

X_{ij} : data

μ_{ik} :derajat keanggotaan data pada tiap *cluster*

7) Cek kondisi berhenti

Jika $(|P_t - P_{t-1}| < \zeta)$ atau $t > \max Iter$ maka proses iterasi berhenti; Jika tidak: $t = t + 1$, dan ulangi langkah ke-4.

Seperti dalam [7],[8],[9] langkah-langkah dalam proses clustering yang dilakukan hampir sama dengan apa yang telah dijabarkan di atas, meskipun dalam pembahasannya nanti terdapat beberapa perbedaan atau terdapat penambahan proses, hal tersebut dilakukan guna untuk memperjelas darimana nilai-nilai tersebut diperoleh. Sehingga tidak terdapat kerancuan data atau nilai yang menimbulkan kebingungan dan kesalahan perhitungan pada langkah-langkah selanjutnya.

III. HASIL DAN PEMBAHASAN

Dalam pembahasan, peneliti menggunakan dan mengutip dataset yang dikutip dari website penyedia dataset terbesar di internet bernama kaggle.com yang dikutip dari [10].

A. Penghitungan Manual Microsoft Excel

Alias	a1	a2	a3	a4	a5	a6
Mahasiswa1	1	55	3.5	1	1	0
Mahasiswa2	0	75	2.5	0	2	0
Mahasiswa3	1	77	2.2	1	1	1
Mahasiswa4	1	50	2.3	0	2	0
Mahasiswa5	1	91	2	1	2	1
Mahasiswa6	1	80	2.5	1	2	1
Mahasiswa7	0	45	1.2	0	2	0
Mahasiswa8	0	60	2.9	0	1	0
Mahasiswa9	0	65	3.1	0	2	0
Mahasiswa10	0	80	3.4	0	2	0
Mahasiswa11	1	78	2.8	0	2	0
Mahasiswa12	0	40	2.1	1	2	0
Mahasiswa13	0	75	1.1	1	2	1
Mahasiswa14	0	95	1.5	1	2	1
Mahasiswa15	1	80	2.2	1	2	1
Mahasiswa16	0	58	2.5	0	2	0
Mahasiswa17	0	70	2.6	0	2	0
Mahasiswa18	0	92	1.8	1	2	1
Mahasiswa19	0	47	1.7	1	2	0
Mahasiswa20	0	59	2.9	0	2	0
Mahasiswa21	0	88	3.8	0	2	0
Mahasiswa22	0	87	3.9	0	2	0

Tabel 1. Data Calon Mahasiswa Penerima Beasiswa KIP-Kuliah

Data tersebut di atas merupakan dataset yang peneliti gunakan dalam melakukan studi kasus perhitungan manual menggunakan microsoft excel. Dimana di dalam dataset tersebut diklasifikasikan beberapa atribut dan class yang menganalogikan setiap syarat-syarat yang diperlukan untuk melakukan seleksi penerimaan beasiswa KIP-K. Berikut klasifikasinya:

1. Header a1 mewakili atribut prestasi,
2. Header a2 mewakili atribut nilai ujian,
3. Header a3 mewakili atribut penghasilan orang tua,
4. Header a4 mewakili atribut kepemilikan kartu sejenis,
5. Header a5 mewakili atribut kepemilikan asset atau rumah,
6. Dan header a6 mewakili class target hak penerima beasiswa dimana angka 0 artinya tidak berhak dan angka 1 artinya berhak

menerima.

Setelah memiliki dataset di atas hal yang dilakukan berikutnya untuk menghitung dan mengolah datanya yaitu:

- 1) Menentukan nilai parameter awal
 1. Jumlah cluster = 2
 2. Pangkat pembobot = 2
 3. Maksimum iterasi = max iterasi
 4. Error terkecil diharapkan = 1
 5. Fungsi objektif awal = 0
 6. Iterasi awal = 1
- 2) Menentukan bilangan random (U) sesuai dengan jumlah cluster, dan dengan syarat jumlah dari kedua bilangan tersebut adalah samadengan 1

u1	u2
0.1	0.9
0.2	0.8
0.3	0.7
0.4	0.6
0.5	0.5
0.6	0.4
0.7	0.3
0.2	0.8
0.4	0.6
0.6	0.4
0.8	0.2
0.1	0.9
0.2	0.8
0.4	0.6
0.2	0.8
0.4	0.6
0.6	0.4
0.8	0.2
0.1	0.9
0.7	0.3
0.3	0.7
0.2	0.8

Tabel 2. Tabel Penentuan Bilangan Random U

Jumlah suku U sesuai dengan jumlah

cluster yang kita tentukan. Untuk menentukan nilai random U di atas, harus dipastikan bahwa jumlah dari kedua suku U adalah samadengan 1. Dapat dilihat pada tabel di atas, bahwa kedua suku U tersebut telah memenuhi jumlah syarat yaitu = 1.

- 3) Menentukan centroid atau pusat cluster. Dikarenakan cluster yang telah ditentukan adalah berjumlah 2 cluster, maka kedua pusat cluster tersebutlah yang akan ditentukan nilai centroid atau pusat clusternya.

Untuk menentukan nilai centroid, berikut langkah-langkahnya:

- a) Untuk header ke-1 pada excel. Diisi dengan nilai hasil pemangkatan U pada setiap cluster dengan w atau pangkat pembobot. Pangkat pembobot yang telah ditentukan adalah 2. Maka, setiap U dipangkatkan dengan 2.
- b) Untuk header ke-2 hingga header terakhir (header ke-7), kalikan hasil pemangkatan poin pertama tadi dengan setiap data dari atribut a1 hingga atribut a6.
- c) Setelah semua hasil perhitungan dari header 1 hingga 7. Hitung jumlah (sum of product) dari setiap kolom ke bawah.

Langkah-langkah di atas dilakukan di setiap menentukan nilai centroid setiap clusternya. Yaitu cluster pertama dan kedua. Jadi akan diperlukan 2 tabel untuk menghitungnya.

data ke-	u1^w	u1^w*a1	u1^w*a2
1	0.01	0.01	0.55
2	0.04	0	3
3	0.09	0.09	6.93
4	0.16	0.16	8
5	0.25	0.25	22.75
6	0.36	0.36	28.8
7	0.49	0	22.05
8	0.04	0	2.4
9	0.16	0	10.4
10	0.36	0	28.8
11	0.64	0.64	49.92
12	0.01	0	0.4
13	0.04	0	3
14	0.16	0	15.2
15	0.04	0.04	3.2
16	0.16	0	9.28
17	0.36	0	25.2
18	0.64	0	58.88
19	0.01	0	0.47
20	0.49	0	28.91
21	0.09	0	7.92
22	0.04	0	3.48
Jumlah	4.64	1.55	339.54

u1^w*a3	u1^w*a4	u1^w*a5	u1^w*a6
0.035	0.01	0.01	0
0.1	0	0.08	0
0.198	0.09	0.09	0.09
0.368	0	0.32	0
0.5	0.25	0.5	0.25
0.9	0.36	0.72	0.36
0.588	0	0.98	0
0.116	0	0.04	0
0.496	0	0.32	0
1.224	0	0.72	0
1.792	0	1.28	0
0.021	0.01	0.02	0
0.044	0.04	0.08	0.04
0.24	0.16	0.32	0.16
0.088	0.04	0.08	0.04
0.4	0	0.32	0
0.936	0	0.72	0
1.152	0.64	1.28	0.64
0.017	0.01	0.02	0
1.421	0	0.98	0
0.342	0	0.18	0
0.156	0	0.08	0
11.134	1.61	9.14	1.58

Tabel 3. Tabel Perhitungan Nilai Centroid Cluster 1

data ke-	u1^w	u1^w*a1	u1^w*a2
1	0.81	0.81	44.55
2	0.64	0	48
3	0.49	0.49	37.73
4	0.36	0.36	18
5	0.25	0.25	22.75
6	0.16	0.16	12.8
7	0.09	0	4.05
8	0.64	0	38.4
9	0.36	0	23.4
10	0.16	0	12.8
11	0.04	0.04	3.12
12	0.81	0	32.4
13	0.64	0	48
14	0.36	0	34.2
15	0.64	0.64	51.2
16	0.36	0	20.88
17	0.16	0	11.2
18	0.04	0	3.68
19	0.81	0	38.07
20	0.09	0	5.31
21	0.49	0	43.12
22	0.64	0	55.68
Jumlah	9.04	2.75	609.34

u1^w*a3	u1^w*a4	u1^w*a5	u1^w*a6
2.835	0.81	0.81	0
1.6	0	1.28	0
1.078	0.49	0.49	0.49
0.828	0	0.72	0
0.5	0.25	0.5	0.25
0.4	0.16	0.32	0.16
0.108	0	0.18	0
1.856	0	0.64	0
1.116	0	0.72	0
0.544	0	0.32	0
0.112	0	0.08	0
1.701	0.81	1.62	0
0.704	0.64	1.28	0.64
0.54	0.36	0.72	0.36
1.408	0.64	1.28	0.64
0.9	0	0.72	0
0.416	0	0.32	0
0.072	0.04	0.08	0.04
1.377	0.81	1.62	0
0.261	0	0.18	0
1.862	0	0.98	0
2.496	0	1.28	0
22.714	5.01	16.14	2.58

Tabel3. Tabel Perhitungan Nilai Centroid Cluster 2

- 4) Menghitung nilai cluster baru pada setiap atribut

Setelah mendapatkan nilai centroid. Nilai-nilai yang telah didapatkan tadi digunakan untuk mencari nilai cluster baru.

Langkah-langkah :

- Untuk Cluster pertama, atribut a1 hingga atribut a6 yaitu dengan rumus $= \frac{\sum Centroid a_n}{U1^w}$, dilakukan dari a1 hingga a6 dengan rumus yang sama, $a_n = a1, a2, a3, a4, a5, a6$
- Untuk Cluster kedua dihitung dengan ketentuan dan rumus yang sama seperti poin pertama (poin a)

a1	a2	a3
0.334051724	73.17672414	2.399568966
0.30420354	67.40486726	2.512610619

a4	a5	a6
0.346982759	1.969827586	0.340517241
0.55420354	1.78539823	0.28539823

Tabel 4. Tabel Perhitungan Nilai Cluster Baru

- 5) Menghitung nilai fungsi objektifitas
Untuk menentukan nilai fungsi

objektifitas diperlukan nilai cluster baru yang telah dihitung di Poin sebelumnya tadi. Langkah-langkahnya adalah sebagai berikut :

- a) Hitung nilai fungsi objektifitas dengan rumus excel =

$$\frac{(\text{Nilai atribut data awal } (a_n) - \text{Nilai clustering baru } (a_n)^2}{}$$

- b) Langkah poin a dilakukan terus menerus dari a1 hingga a6
c) Setelah semua nilai objektif pada atribut telah didapatkan lalu jumlahkan nilai objektif (Σ) per barisnya. Dari kiri(a1) hingga ke kanan(a6)

Langkah-langkah di atas dilakukan untuk tabel cluster pertama dan cluster kedua.

[a1Data-a1NewClust1]^2	[a2Data-a2NewClust1]^2	[a3Data-a3NewClust1]^2
0.443487106	330.3933004	1.210948462
0.111590554	3.324334869	0.010086393
0.443487106	14.61743832	0.039827772
0.443487106	537.1605418	0.009913979
0.443487106	317.6691625	0.159655358
0.443487106	46.55709349	0.010086393
0.111590554	793.9277831	1.438965703
0.111590554	173.626059	0.25043122
0.111590554	66.85881763	0.490603634
0.111590554	46.55709349	1.000862255
0.443487106	23.26399004	0.160345013
0.111590554	1100.695025	0.089741565
0.111590554	3.324334869	1.688879496
0.111590554	476.2553694	0.809224324
0.443487106	46.55709349	0.039827772
0.111590554	230.3329556	0.010086393
0.111590554	10.09157625	0.0401726
0.111590554	354.3157142	0.359482944
0.111590554	685.2208866	0.489396738
0.111590554	200.9795073	0.25043122
0.111590554	219.7295073	1.961207082
0.111590554	191.0829556	2.251293289

[a4Data-a4NewClust1]^2	[a5Data-a5NewClust1]^2	[a6Data-a6NewClust1]^2	Total
0.426431518	0.940565547	0.115951992	333.5307
0.120397035	0.000910375	0.115951992	3.683271
0.426431518	0.940565547	0.434917509	16.90267
0.120397035	0.000910375	0.115951992	537.8512
0.426431518	0.000910375	0.434917509	319.1346
0.426431518	0.000910375	0.434917509	47.87293
0.120397035	0.000910375	0.115951992	795.7156
0.120397035	0.940565547	0.115951992	175.165
0.120397035	0.000910375	0.115951992	67.69827
0.120397035	0.000910375	0.115951992	47.90681
0.120397035	0.000910375	0.115951992	24.10508
0.426431518	0.000910375	0.115951992	1101.44
0.426431518	0.000910375	0.434917509	5.987064
0.426431518	0.000910375	0.434917509	478.0384
0.426431518	0.000910375	0.434917509	47.90267
0.120397035	0.000910375	0.115951992	230.6919
0.120397035	0.000910375	0.115951992	10.4806
0.426431518	0.000910375	0.434917509	355.649
0.426431518	0.000910375	0.115951992	686.3652
0.120397035	0.000910375	0.115951992	201.5788
0.120397035	0.000910375	0.115951992	222.0396
0.120397035	0.000910375	0.115951992	193.6831

Tabel 5. Tabel Perhitungan Nilai Fungsi Objektifitas Cluster 1

[a1Data-a1NewClust2]^2	[a2Data-a2NewClust2]^2	[a3Data-a3NewClust2]^2
0.484132714	153.8807317	0.974937789
0.092539794	57.68604139	0.000159028
0.484132714	92.06657236	0.097725399
0.484132714	302.9294042	0.045203276
0.484132714	556.7302892	0.262769647
0.484132714	158.6373688	0.000159028
0.092539794	501.9780768	1.722946638
0.092539794	54.83205909	0.150070532
0.092539794	5.783386522	0.345026284
0.092539794	158.6373688	0.787459913
0.484132714	112.2568378	0.082592656
0.092539794	751.0267494	0.170247523
0.092539794	57.68604139	1.995468762
0.092539794	761.4913511	1.025380267
0.484132714	158.6373688	0.097725399
0.092539794	88.45152811	0.000159028
0.092539794	6.734713956	0.007636904
0.092539794	604.9205547	0.507813895
0.092539794	416.3586078	0.660336019
0.092539794	70.6417936	0.150070532
0.092539794	424.1594927	1.657371417
0.092539794	383.9692272	1.924849293

[a4Data-a4NewClust2]^2	[a5Data-a5NewClust2]^2	[a6Data-a6NewClust2]^2	Total
0.198734484	0.61685038	0.08145215	156.2368
0.307141564	0.04605392	0.08145215	58.21339
0.198734484	0.61685038	0.51065569	93.97467
0.307141564	0.04605392	0.08145215	303.8934
0.198734484	0.04605392	0.51065569	558.2326
0.198734484	0.04605392	0.51065569	159.8771
0.307141564	0.04605392	0.08145215	504.2282
0.307141564	0.61685038	0.08145215	56.08011
0.307141564	0.04605392	0.08145215	6.6556
0.307141564	0.04605392	0.08145215	159.952
0.307141564	0.04605392	0.08145215	113.2582
0.198734484	0.04605392	0.08145215	751.6158
0.198734484	0.04605392	0.51065569	60.52949
0.198734484	0.04605392	0.51065569	763.3647
0.198734484	0.04605392	0.51065569	159.9747
0.307141564	0.04605392	0.08145215	88.97887
0.307141564	0.04605392	0.08145215	7.269538
0.198734484	0.04605392	0.51065569	606.2764
0.198734484	0.04605392	0.08145215	417.4377
0.307141564	0.04605392	0.08145215	71.31905
0.307141564	0.04605392	0.08145215	426.3441
0.307141564	0.04605392	0.08145215	386.4213

Tabel 6. Tabel Perhitungan Nilai Fungsi Objektifitas Cluster 1

- 6) Menghitung P Total

Perhitungan P Total digunakan untuk menentukan sum of error guna menentukan kapan iterasi harus berhenti. Berikut langkah-langkahnya :

- a) Hitung nilai P setiap clusternya.

Didapatkan dengan rumus =

$$\text{Jumlah NFO perbaris} \times U_{n \text{ cluster}}^2$$

- b) Poin a dilakukan dari data 1 hingga data terakhir (1 sampai 22)
c) Langkah a,b dilakukan untuk setiap cluster
d) Cari P Total. Setelah didapatkan nilai P pada setiap baris data di setiap clusternya lalu jumlahkan P cluster 1 dengan P cluster 2. (dari

- baris 1 hingga 22)
- Setelah didapatkan nilai P Total setiap barisnya. Lalu jumlahkan (sum) semua hasil P Total dari data 1 hingga 22)
 - Jumlah P Total tersebutlah yang digunakan untuk mencari sum of error di setiap iterasinya untuk menentukan kapan iterasi harus berhenti.

Cluster1		
Total	$u1^w$	$P = (total * (u1^w))$
333.530685	0.01	3.33530685
3.683271217	0.04	0.147330849
16.90266777	0.09	1.521240099
537.8512023	0.16	86.05619236
319.1345643	0.25	79.78364108
47.87292639	0.36	17.2342535
795.7155988	0.49	389.9006434
175.1649954	0.04	7.006599814
67.69827122	0.16	10.83172339
47.9068057	0.36	17.24645005
24.10508156	0.64	15.4272522
1101.439651	0.01	11.01439651
5.987064321	0.04	0.239482573
478.0384436	0.16	76.48615098
47.90266777	0.04	1.916106711
230.6918919	0.16	36.91070271
10.4805988	0.36	3.773015569
355.6490471	0.64	227.6153901
686.3651678	0.01	6.863651678
201.5787885	0.49	98.77360634
222.0395643	0.09	19.98356079
193.6830988	0.04	7.747323952

Cluster2			P Total
Total	$u2^w$	$P = (total * (u2^w))$	
156.2368	0.81	126.5518397	129.8871
58.21339	0.64	37.25656822	37.4039
93.97467	0.49	46.0475888	47.56883
303.8934	0.36	109.4016196	195.4578
558.2326	0.25	139.5581589	219.3418
159.8771	0.16	25.58033675	42.81459
504.2282	0.09	45.38053898	435.2812
56.08011	0.64	35.89127264	42.89787
6.6556	0.36	2.396016084	13.22774
159.952	0.16	25.59232259	42.83877
113.2582	0.04	4.530328434	19.95758
751.6158	0.81	608.8087796	619.8232
60.52949	0.64	38.73887618	38.97836
763.3647	0.36	274.8112975	351.2974
159.9747	0.64	102.3837895	104.2999
88.97887	0.36	32.03239484	68.9431
7.269538	0.16	1.163126126	4.936142
606.2764	0.04	24.2510541	251.8664
417.4377	0.81	338.1245565	344.9882
71.31905	0.09	6.41871464	105.1923
426.3441	0.49	208.9085853	228.8921
386.4213	0.64	247.3096089	255.0569
		Jumlah P Total	3600.95

Tabel 7. Tabel Perhitungan Nilai P Total

- Menghitung perubahan matriks partisi U
Sebelum menghitung perubahan matriks partisi U.
 - hitung terlebih dahulu nilai X Total, yang didapatkan dari menjumlahkan Semua Total nilai fungsi objektifitas yang dihasilkan di pencarian nilai fungsi objektifitas tadi.

$$(\text{Total NFO suku-1 (cluster1)} + \text{Total NFO suku-1 (cluster2)})$$

Seterusnya hingga suku atau data ke-22

- Lalu untuk menentukan U1 yaitu dengan rumus

$$\frac{\text{Jumlah X Cluster 1}}{X \text{ Total}}$$

- Lalu untuk menentukan U2 yaitu dengan rumus

$$\frac{\text{Jumlah X Cluster 2}}{X \text{ Total}}$$

- Pastikan jumlah kedua U adalah samadengan 1
- Simpan nilai u1 dan u2. Dan masukkan kembali nilai u1 dan u2 ke langkah pertama apabila iterasi belum berhak berhenti

Cluster 1					
x1	x2	x3	x4	x5	x6
0.443487106	330.3933004	1.210948462	0.426432	0.940566	0.115951992
0.111590554	3.324334869	0.010086393	0.120397	0.00091	0.115951992
0.443487106	14.61743832	0.039827772	0.426432	0.940566	0.434917509
0.443487106	537.1605418	0.009913979	0.120397	0.00091	0.115951992
0.443487106	317.6691625	0.159655358	0.426432	0.00091	0.434917509
0.443487106	46.55709349	0.010086393	0.426432	0.00091	0.434917509
0.111590554	793.9277831	1.438965703	0.120397	0.00091	0.115951992
0.111590554	173.626059	0.25043122	0.120397	0.940566	0.115951992
0.111590554	66.85881763	0.490603634	0.120397	0.00091	0.115951992
0.111590554	46.55709349	1.000862255	0.120397	0.00091	0.115951992
0.443487106	23.26399004	0.160345013	0.120397	0.00091	0.115951992
0.111590554	1100.695025	0.089741565	0.426432	0.00091	0.115951992
0.111590554	3.324334869	1.688879496	0.426432	0.00091	0.434917509
0.111590554	476.2553694	0.809224324	0.426432	0.00091	0.434917509
0.443487106	46.55709349	0.039827772	0.426432	0.00091	0.434917509
0.111590554	230.3329556	0.010086393	0.120397	0.00091	0.115951992
0.111590554	10.09157625	0.0401726	0.120397	0.00091	0.115951992
0.111590554	354.3157142	0.359482944	0.426432	0.00091	0.434917509
0.111590554	685.2208866	0.489396738	0.426432	0.00091	0.115951992
0.111590554	200.9795073	0.25043122	0.120397	0.00091	0.115951992
0.111590554	219.7295073	1.961207082	0.120397	0.00091	0.115951992
0.111590554	191.0829556	2.251293289	0.120397	0.00091	0.115951992

Cluster 2					
x1	x2	x3	x4	x5	x6
0.484133	153.8807	0.974938	0.198734484	0.61685	0.081452
0.09254	57.68604	0.000159	0.307141564	0.046054	0.081452
0.484133	92.06657	0.097725	0.198734484	0.61685	0.510656
0.484133	302.9294	0.045203	0.307141564	0.046054	0.081452
0.484133	556.7303	0.26277	0.198734484	0.046054	0.510656
0.484133	158.6374	0.000159	0.198734484	0.046054	0.510656
0.09254	501.9781	1.722947	0.307141564	0.046054	0.081452
0.09254	54.83206	0.150071	0.307141564	0.61685	0.081452
0.09254	5.783387	0.345026	0.307141564	0.046054	0.081452
0.09254	158.6374	0.78746	0.307141564	0.046054	0.081452
0.484133	112.2568	0.082593	0.307141564	0.046054	0.081452
0.09254	751.0267	0.170248	0.198734484	0.046054	0.081452
0.09254	57.68604	1.995469	0.198734484	0.046054	0.510656
0.09254	761.4914	1.02538	0.198734484	0.046054	0.510656
0.484133	158.6374	0.097725	0.198734484	0.046054	0.510656
0.09254	88.45153	0.000159	0.307141564	0.046054	0.081452
0.09254	6.734714	0.007637	0.307141564	0.046054	0.081452
0.09254	604.9206	0.507814	0.198734484	0.046054	0.510656
0.09254	416.3586	0.660336	0.198734484	0.046054	0.081452
0.09254	70.64179	0.150071	0.307141564	0.046054	0.081452
0.09254	424.1595	1.657371	0.307141564	0.046054	0.081452
0.09254	383.9692	1.924849	0.307141564	0.046054	0.081452

X Total	U1	U2
489.768	0.681	0.319
61.8967	0.05951	0.94049
110.877	0.15244	0.84756
841.745	0.63897	0.36103
877.367	0.36374	0.63626
207.75	0.23044	0.76956
1299.94	0.61212	0.38788
231.245	0.75749	0.24251
74.3539	0.91049	0.08951
207.859	0.23048	0.76952
137.363	0.17548	0.82452
1853.06	0.59439	0.40561
66.5166	0.09001	0.90999
1241.4	0.38508	0.61492
207.877	0.23044	0.76956
319.671	0.72165	0.27835
17.7501	0.59045	0.40955
961.925	0.36973	0.63027
1103.8	0.62182	0.37818
272.898	0.73866	0.26134
648.384	0.34245	0.65755
580.104	0.33388	0.66612

Tabel 8. Tabel Penentuan Nilai Partisi U Baru Iterasi 1

Cluster 1					
x1	x2	x3	x4	x5	x6
0.587136878	3.051534918	1.139805793	0.4224259	0.59414741	6.37084E-05
0.054639729	472.9261073	0.004572042	0.1225399	0.05252825	6.37084E-05
0.587136878	563.9135646	0.054001916	0.4224259	0.59414741	0.984100196
0.587136878	10.58289182	0.017525292	0.1225399	0.05252825	6.37084E-05
0.587136878	1424.825765	0.186955166	0.4224259	0.05252825	0.984100196
0.587136878	715.3947504	0.004572042	0.4224259	0.05252825	0.984100196
0.054639729	68.11424872	1.518768165	0.1225399	0.05252825	6.37084E-05
0.054639729	45.52017802	0.218665542	0.1225399	0.59414741	6.37084E-05
0.054639729	137.9882111	0.445712293	0.1225399	0.05252825	6.37084E-05
0.054639729	715.3947504	0.936282418	0.1225399	0.05252825	6.37084E-05
0.587136878	612.4072932	0.135142167	0.1225399	0.05252825	6.37084E-05
0.054639729	175.6456056	0.110478541	0.4224259	0.05252825	6.37084E-05
0.054639729	472.9261073	1.77524479	0.4224259	0.05252825	0.984100196
0.054639729	1742.80068	0.86933829	0.4224259	0.05252825	0.984100196
0.587136878	715.3947504	0.054001916	0.4224259	0.05252825	0.984100196
0.054639729	22.53272078	0.004572042	0.1225399	0.05252825	6.37084E-05
0.054639729	280.4574642	0.028095417	0.1225399	0.05252825	6.37084E-05
0.054639729	1501.319494	0.399908416	0.4224259	0.05252825	0.984100196
0.054639729	39.10170596	0.536385041	0.4224259	0.05252825	6.37084E-05
0.054639729	33.0264494	0.218665542	0.1225399	0.05252825	6.37084E-05
0.054639729	1207.344579	1.870375919	0.1225399	0.05252825	6.37084E-05
0.054639729	1138.850851	2.153899294	0.1225399	0.05252825	6.37084E-05

Cluster 2					
x1	x2	x3	x4	x5	x6
0.36213	758.442	1.03031	0.195786671	0.84705	0.31083
0.15859	56.8489	0.00023	0.310830919	0.00634	0.31083
0.36213	30.6896	0.0812	0.195786671	0.84705	0.19579
0.36213	1058.84	0.03421	0.310830919	0.00634	0.31083
0.36213	71.5746	0.23518	0.195786671	0.00634	0.19579
0.36213	6.4507	0.00023	0.195786671	0.00634	0.19579
0.15859	1409.24	1.65111	0.310830919	0.00634	0.31083
0.15859	508.044	0.17226	0.310830919	0.84705	0.31083
0.15859	307.645	0.37828	0.310830919	0.00634	0.31083
0.15859	6.4507	0.83731	0.310830919	0.00634	0.31083
0.36213	20.61	0.09925	0.310830919	0.00634	0.31083
0.15859	1809.64	0.14819	0.195786671	0.00634	0.31083
0.15859	56.8489	1.9181	0.195786671	0.00634	0.19579
0.15859	155.256	0.97014	0.195786671	0.00634	0.19579
0.36213	6.4507	0.0812	0.195786671	0.00634	0.19579
0.15859	602.203	0.00023	0.310830919	0.00634	0.31083
0.15859	157.247	0.01324	0.310830919	0.00634	0.31083
0.15859	89.4949	0.46916	0.195786671	0.00634	0.19579
0.15859	1263.08	0.61616	0.195786671	0.00634	0.31083
0.15859	554.123	0.17226	0.310830919	0.00634	0.31083
0.15859	29.8135	1.72934	0.310830919	0.00634	0.31083
0.15859	19.8932	2.00235	0.310830919	0.00634	0.31083

X Total	U1	U2
766.983	0.00756	0.99244
530.796	0.89142	0.10858
598.927	0.94595	0.05405
1071.23	0.01061	0.98939
1499.63	0.95161	0.04839
724.656	0.99005	0.00995
1481.54	0.04716	0.95284
556.353	0.0836	0.9164
447.475	0.30988	0.69012
724.635	0.98886	0.01114
635.004	0.96583	0.03417
1986.74	0.08873	0.91127
535.539	0.88923	0.11077
1901.97	0.91757	0.08243
724.787	0.98994	0.01006
625.757	0.03638	0.96362
438.762	0.63979	0.36021
1593.75	0.9432	0.0568
1304.53	0.03079	0.96921
588.557	0.05688	0.94312
1241.77	0.97397	0.02603
1163.92	0.98051	0.01949

Tabel 9. Tabel Penentuan Nilai Partisi U Baru Iterasi Terakhir

8) Mengecek kondisi berhenti iterasi

Lakukan perhitungan iterasi ke 2 (dari langkah 1 hingga 7) lagi untuk mendapatkan nilai jumlah P Total yang kedua

Pengecekan kondisi iterasi berhenti yaitu dengan mengurangi Jumlah Nilai P Total iterasi sebelumnya dengan Jumlah Nilai P Total iterasi selanjutnya

$$\text{Error} = \sum P \text{ Total}_{\text{sebelumnya}} - \sum P \text{ Total}_{\text{sesudahnya}}$$

Iterasi dihentikan apabila hasil error terakhir ≤ 1 .

PO Error

iterasi1	3600.95	
iterasi2	2543	1057.95
iterasi3	1430	1113
iterasi4	1136	294
iterasi5	1132	4
iterasi6	1177	-45

Tabel 10. Tabel Sum Of Error

0.1	0.9	0.9	0
0.6	0.4	0.6	1
1	0	1	1
0.1	0.9	0.9	0
0.1	0.9	0.9	0
1	0	1	1
1	0	1	1

Tabel 11. Tabel Penentuan Kondisi/Kategori Cluster Baru

9) Menentukan kategori/kondisi cluster baru
Langkah-langkah :

1. Kutip nilai matriks partisi U terakhir yang didapatkan di iterasi terakhir.
2. Cari nilai maksimal/MAX dari u1 dan u2 suku/data pertama. Dengan rumus excel : **"=MAX(kolomU1:KolomU2)"**
Copy rumus dengan menariknya hingga suku/data(kolom) terakhir
3. Tentukan kategori/kondisi kluster baru dengan rumus excel : **"=IF(kolomPosisiNilaiMaxU=KolomMAX suku pertama;0;1)"**
Copy rumus dengan menariknya hingga suku/data(kolom) terakhir
4. Nilai kondisi/kategori cluster baru telah ditemukan

U1	U2	Max	Kluster Baru
0.1	0.9	0.9	0
0.9	0.1	0.9	1
1	0	1	1
0.1	0.9	0.9	0
1	0	1	1
1	0	1	1
0.1	0.9	0.9	0
0.1	0.9	0.9	0
0.3	0.7	0.7	0
1	0	1	1
1	0	1	1
0.1	0.9	0.9	0
0.9	0.1	0.9	1
1	0	1	1
1	0	1	1

10) Bandingkan kategori/kondisi cluster baru dengan hasil Rapid Miner

id	a6	cluster
1	0	cluster_0
2	0	cluster_1
3	1	cluster_1
4	0	cluster_0
5	1	cluster_1
6	1	cluster_1
7	0	cluster_0
8	0	cluster_0
9	0	cluster_0
10	0	cluster_1
11	0	cluster_1
12	0	cluster_0
13	1	cluster_1
14	1	cluster_1
15	1	cluster_1
16	0	cluster_0
17	0	cluster_1
18	1	cluster_1
19	0	cluster_0
20	0	cluster_0
21	0	cluster_1
22	0	cluster_1

Gambar1. Hasil Perhitungan Otomatis Dengan Rapid Miner

Dengan memasukkan dataset mentah yang didapatkan dari [10] ke dalam aplikasi Rapid Miner Studio lalu memprosesnya dengan mendeklarasikan peran header a6 sebagai class target. Maka didapatkanlah nilai cluster baru hasil perhitungan otomatis, yang dapat dibandingkan dengan nilai cluster hasil perhitungan manual dengan microsoft excel.

IV. KESIMPULAN

Dari hasil penelitian yang telah dilakukan, peneliti melakukan studi kasus terhadap 22 calon penerima beasiswa KIP-K yang dimana datanya diolah menggunakan perhitungan manual dan perhitungan otomatis. Dari data asli yang menunjukkan bahwa terdapat 7 mahasiswa yang mendapatkan angka biner 1 yang berarti mahasiswa tersebut berhak mendapatkan beasiswa, setelah diolah menggunakan excel, berubah menjadi 13 Mahasiswa yang mendapatkan nilai biner 1 atau berhak mendapatkan beasiswa. Selain dengan perhitungan manual menggunakan microsoft excel, perhitungan otomatis juga dilakukan guna untuk membuktikan kebenarannya. Dan setelah dilakukan perhitungan otomatis, ternyata hasil yang didapatkan adalah sama. Hal tersebut membuktikan bahwa perhitungan secara manual sudah benar.

DAFTAR PUSTAKA

- [1] Anjani, D., Mustari, D., & Farkhatin, N. (2019, July). KAJIAN PENERAPAN ALGORITMA FUZZY C-MEANS DALAM PENENTUAN JURUSAN BAGI SISWA KELAS X SMA: STUDI KASUS SMA I BARUNAWATI JAKARTA BARAT. *ORBITH*, 15, 56-63. Retrieved June 24, 2023
- [2] D.L., R., V.Y.I., I., & M.H., H. (2017, Maret). Implementasi Fuzzy C-Means Clustering Dalam Penentuan Beasiswa. *Jurnal Ilmu Matematika dan Terapan*, 11, 1-11. Retrieved June 24, 2023
- [3] Irawati, I., & Sitorus, S. H. (2019). PENENTUAN KELOMPOK KUALITAS SEKOLAH DENGAN METODE FUZZY C-MEANS BERDASARKAN NILAI UJIAN DAN RAPOR. *Jurnal Komputer dan Aplikasi*, 7, 11-22. Retrieved Juni 24, 2023
- [4] Mirza, M. F. (n.d.). METODE CLUSTERING DENGAN ALGORITMA FUZZY C-MEANS UNTUK REKOMENDASI PEMILIHAN BIDANG KEAHLIAN PADA PROGRAM STUDI TEKNIK INFORMATIKA. Retrieved June 25, 2023, from http://eprints.dinus.ac.id/12282/1/jurnal_12191.pdf
- [5] Mulyadi, D., Pratama, E. P., Sarifudin, A., & Yulianti. (2021, September). Penerapan Fuzzy Logic Untuk Sistem Pengendali Lalu Lintas. *Jurnal Informatika Universitas Pamulang*, 6, 540-544. Retrieved June 25, 2023, from <http://openjournal.unpam.ac.id/index.php/informatika/article/view/11575/pdf>
- [6] L. A. Zadeh, "Fuzzy logic," *Computer* (Long. Beach. Calif.), vol. 21, no. 4, pp. 83-93, 1988.
- [7] Ristanto, I. B., Utami, Y. R., & Susyanto, T. (2021, July). Penerapan Metode Clustering dengan Fuzzy C-Means untuk Memetakan Daerah Rawan Kecelakaan Lalu Lintas di Surakarta. *Jurnal Ilmiah Sinus (JIS)*, 19, 27-36. doi:10.30646/sinus.v19i2.562
- [8] Sanusi, W., & Ahmad Zaky, B. N. (n.d.). Analisis Fuzzy C-Means dan Penerapannya Dalam Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Faktor-Faktor Penyebab Gizi Buruk. Retrieved June 24, 2023
- [9] Sari, H. L., & Suranti, D. (n.d.). Perbandingan Algoritma Fuzzy C-Means (FCM) Dan Algoritma Mixture Dalam Penclusteran Data Curah Hujan Kota Bengkulu. Retrieved June 25, 2023
- [10] Purnama, G. W. (2023). *Student-Data-KIP*. Kaggle.com. Retrieved June 24, 2023, from <https://www.kaggle.com/datasets/gedewahyupurnama/student-data-kip?resource=download>