

SISTEM PENDUKUNG KEPUTUSAN UNTUK MENENTUKAN KELOMPOK UANG KULIAH TUNGGAL MAHASISWA BARU DENGAN METODE FUZZY C-MEANS

(Studi Kasus Institut Teknologi Sumatera)

Reizky Patrical Sirya¹, Masayu Leylia Khodra², Ahmad Luky Ramdani³

*Program Studi Teknik Informatika, Institut Teknologi Sumatera
Jalan Terusan Ryacudu, Desa Way Hui, Kecamatan Jatiagung, Lampung Selatan, 3536*

E-mail : reizkypatrical@gmail.com

ABSTRAK

Uang Kuliah Tunggal (UKT) merupakan kebijakan mengenai sistem pembayaran biaya kuliah bagi seluruh perguruan tinggi (PTN) di Indonesia, yaitu dari sistem Sumbangan Pengembangan Manajemen Pendidikan (SPMA) dengan uang pangkalnya menjadi sistem pembayaran tanpa uang pangkal. Di Institut Teknologi Sumatera (ITERA) sendiri UKT telah diterapkan sejak tahun 2014, namun masih banyak mahasiswa baru melakukan banding UKT karena hasil pengelompokan UKT dianggap tidak sesuai. Oleh sebab itu, pada tugas akhir ini akan dilakukan metode lain dalam pengelompokan UKT di ITERA. Pada sistem pengelompokan UKT akan dilakukan dengan menggunakan metode clustering Fuzzy C-Means. Untuk memperoleh hasil cluster optimal maka dibutuhkan kombinasi kriteria yang optimal. Untuk menentukan kombinasi kriteria digunakan feature selection dengan algoritma ReliefF. Setelah didapat hasil pemeringkatan kriteria dengan ReliefF, setiap kombinasi kriteria akan diuji dengan pengelompokan Fuzzy C-Means untuk mencari nilai indeks validitas partition coefficient tertinggi. Berdasarkan hasil eksperimen pengujian kombinasi indeks partition coefficient tertinggi diperoleh sebesar 0.6394 yang merupakan kombinasi dari kriteria penghasilan_tg_jwb, kd_pekerjaan_tg_jwb, kd_pekerjaan_ayah, penghasilan_ayah, pengeluaran_makan_transport, biling dan bidikmisi.

Kata Kunci : Uang Kuliah Tunggal; Clustering; Fuzzy C-Means; ReliefF; Partition Coefficient;

ABSTRACT

Uang Kuliah Tunggal (UKT) is a policy regarding the tuition fee payment system for all public universities (PTN) in Indonesia, from the Sumbangan Pengembangan Manajemen Pendidikan (SPMA) system with the starting fee being a payment system without a starting fee. In Institut Teknologi Sumatera (ITERA), UKT has been implemented since 2014, but there are still many new students who do UKT appeal because the results of the UKT clustering are considered unsuitable. Therefore, in this thesis will be carried out another UKT clustering method in ITERA. In the UKT clustering system will be done using the Fuzzy C-Means clustering method. To obtain the optimal cluster results, an optimal combination of criteria is needed. To determine the combination of criteria used feature selection with the ReliefF algorithm. After obtaining the results of ranking criteria with ReliefF, each combination of criteria will be tested by Fuzzy C-Means clustering to find the highest index value of partition coefficient validity. Based on the experimental results of testing the highest partition coefficient index combination obtained by 0.6394 which is a combination of penghasilan_tg_jwb, kd_pekerjaan_tg_jwb, kd_pekerjaan_ayah, penghasilan_ayah, pengeluaran_makan_transport, biling dan bidikmisi.

Keywords : Uang Kuliah Tunggal; Clustering; Fuzzy C-Means; ReliefF; Partition Coefficient;

1. PENDAHULUAN

1.1 Latar Belakang

Untuk mengejar ketertinggalan pendidikan dan penyetaraan pendidikan di perguruan tinggi di Indonesia, mulai tahun ajaran 2013/2014, Pemerintah menerapkan kebijakan sistem Uang Kuliah Tunggal (UKT). UKT diterapkan pada Perguruan Tinggi Negeri (PTN) yang tertuang dalam Peraturan Menteri Pendidikan dan Kebudayaan No.55 Tahun 2013. Kebijakan ini dilakukan berdasarkan pertimbangan untuk meringankan beban mahasiswa terhadap biaya Pendidikan.

Secara ringkas UKT merupakan beban biaya yang harus dibayarkan oleh mahasiswa per semester selama masa kuliah di kampus. Kalkulasi dana UKT itu berasal dari Biaya Kuliah Tunggal (BKT) yang dikurangi dengan Bantuan Operasional Perguruan Tinggi Negeri (BOPTN). UKT berprinsip dasar subsidi silang, artinya orang yang mampu secara ekonomi memberikan subsidi kepada pihak yang kurang mampu. Dengan adanya sistem UKT, diharapkan tidak ada lagi mahasiswa yang tidak dapat melanjutkan pendidikan PTN karena biaya uang pangkal kuliah yang dianggap mahal dan tidak sesuai dengan kondisi ekonomi mahasiswa. Setiap Perguruan Tinggi Negeri memiliki tarif UKT yang berbeda beda, perbedaan ini dipengaruhi dari tingkat wilayah dan program studinya.

Institut Teknologi Sumatera (ITERA) sudah menerapkan sistem UKT pada tahun ajaran 2015/2016. Penentuan kelompok UKT ditetapkan oleh pemimpin Perguruan Tinggi, hal ini tertuang dalam Peraturan Menteri Pendidikan dan Kebudayaan Nomor 73 tahun 2014. Termuat pada lampiran Peraturan Menteri No.39 tahun 2016, pengelompokan UKT Institut Teknologi Sumatera dibagi menjadi delapan kelompok, dari kelompok yang terendah (kelompok I) sampai kelompok yang tertinggi (kelompok VIII).

Penentuan UKT di ITERA dilakukan oleh panitia Penerimaan Mahasiswa Baru (PMB). Dalam menentukan golongan UKT dilakukan dengan menghitung bobot dari keterkaitan antara kriteria penghasilan orang tua, tanggungan anak, listrik, dan Pajak Bumi dan Bangunan (PBB). Penentuan bobot kriteria yang dilakukan subjektif karena dilakukan dengan cara skema percobaan pada bobot kriteria. Terdapat kesulitan dalam mengolah data mahasiswa

baru selama pengelompokan, salah satunya yaitu dalam penyesuaian bobot kriteria untuk mencapai hasil yang optimal. Kondisi seperti ini menjadikan dilema dalam pengelompokan, karena kondisi ekonomi mahasiswa baru setiap tahunnya akan berbeda. Hal ini berdampak pada setiap tahunnya mahasiswa baru banyak melakukan banding karena pengelompokan UKT yang dihasilkan masih dianggap tidak sesuai.

Dengan mempertimbangkan hal tersebut maka akan dilakukan metode lain dalam penentuan UKT. Pada tugas akhir ini menggunakan metode clustering Fuzzy C-Means (FCM) dalam mengelompokan UKT mahasiswa baru. Fuzzy C-Means pada dasarnya didasari logika fuzzy, teori ini diperkenalkan oleh Lotfi Zadeh (1965). Fuzzy C-Means merupakan teknik pengelompokan yang cocok untuk kemampuan dalam pengenalan pola, data yang tidak lengkap, informasi campuran, dan dapat memberikan solusi perkiraan lebih cepat [16]

Fuzzy C-Means mampu menampilkan hasil yang detail dan pada implementasi Fuzzy C-Means jumlah cluster berpengaruh pada kualitas segmen yang dihasilkan. Semakin banyak jumlah cluster yang digunakan akan berdampak pada hasil pengelompokan yang lebih halus. Semakin dikit jumlah cluster maka hasil pengelompokan semakin kasar [1]. Dalam teori fuzzy, keanggotaan data tidak diberi dengan nilai tagas dengan nilai 1 (menjadi anggota) dan 0 (tidak menjadi anggota), melainkan setiap data memiliki derajat keanggotaan yang memiliki jangkauan nilai 0 sampai 1.

Dalam clustering dibutuhkan kriteria-kriteria yang optimal sehingga clustering yang dihasilkan optimal. Pada tugas akhir ini, untuk menentukan kriteria digunakan feature selection. Menurut Rehat pada Zurnila Feature Selection adalah suatu kegiatan pemodelan atau penganalisaan data yang umumnya dapat dilakukan secara preprocessing dan bertujuan untuk memilih kriteria yang berpengaruh (fitur optimal) dan mengesampingkan kriteria yang tidak berpengaruh [2]. Untuk menemukan kriteria yang optimal, pada penelitian ini digunakan algoritma ReliefF. ReliefF adalah perbaikan metode dari relief, algoritma relief sendiri adalah metode estimasi pembobotan sebuah fitur. Semakin besar bobot sebuah fitur, maka dianggap semakin relevan fitur tersebut dengan output. Namun, relief sudah lama tidak digunakan lagi karena ketidakstabilan akurasi yang dihasilkan, dikarenakan ketidakmampuannya untuk mengambil dan mengevaluasi sampel berulang kali

dengan bobot fitur yang sama [3]. Sedangkan ReliefF dapat mengevaluasi nilai fitur dengan berulang kali mengambil sampel instance dan mempertimbangkan nilai fitur yang diberikan untuk instance terdekat dari kelas yang sama dan yang berbeda.. Berdasarkan latar belakang permasalahan ini diharapkan metode clustering dengan menggunakan metode Fuzzy C-Means dan Feature Selection ReliefF ini mampu menentukan kelompok UKT yang optimal.

1.2 Rumusan Masalah

Berdasarkan uraian pada latar belakang maka permasalahan yang akan diangkat dalam tugas akhir ini sebagai berikut :

1. Bagaimana menentukan kriteria dalam penentuan UKT?
2. Bagaimana membuat sistem pendukung keputusan untuk dapat mengelompokkan UKT mahasiswa baru dengan metode Fuzzy C-Means?

1.3 Tujuan

Tujuan dari penelitian ini adalah sebagai berikut :

1. Melakukan eksperimen untuk mencari kriteria-kriteria optimal yang berpengaruh pada pengelompokan UKT sehingga menghasilkan nilai validasi cluster tertinggi.
2. Mengimplementasikan metode Fuzzy C-Means untuk mengelompokkan UKT di Institut Teknologi Sumatera.

2. ANALISIS PERANCANGAN

2.1 Analisis Persoalan

Sistem UKT (Uang Kuliah Tunggal) yang ditetapkan pemerintah pada tahun 2013 merupakan sebagian Biaya Kuliah Tunggal (BKT) yang ditanggukan kepada setiap mahasiswa berdasarkan kemampuan ekonominya. Biaya kuliah Tunggal merupakan seluruh biaya operasional per mahasiswa per semester pada program studi di perguruan tinggi negeri dan UKT itu ditetapkan berdasarkan BKT dikurangi dengan biaya yang ditanggung oleh pemerintah. yang akan dilaksanakan oleh perguruan tinggi negeri, kebijakan ini bertujuan untuk meringankan beban uang kuliah yang harus dibayarkan oleh mahasiswa dari awal masuk perkuliahan hingga lulus nanti.

Di Institut Teknologi Sumatera sendiri telah menerapkan sistem UKT sejak tahun ajar 2015/2016. Penetapan nominal UKT itu sendiri telah diatur oleh

undang- undang sesuai dengan kebutuhan operasional universitas negeri yang bersangkutan dalam menjalankan kegiatan belajar mengajar, hal ini telah diatur pada Permendikbud No. 55 Tahun 2013 tertanggal 23 Mei 2013. Penentuan UKT di ITERA dilakukan oleh panitia Penerimaan Mahasiswa Baru (PMB). Dalam menentukan golongan UKT dilakukan dengan menghitung bobot dari keterkaitan antara penghasilan orang tua, tanggungan anak , listrik, dan Pajak Bumi dan Bangunan (PBB). Penentuan bobot kriteria yang dilakukan subjektif karena dilakukan dengan cara skema percobaan pada bobot kriteria. Terdapat kesulitan dalam mengolah data mahasiswa baru selama pengelompokan, salah satunya yaitu dalam menyesuaikan bobot kriteria untuk mencapai hasil yang optimal. Kondisi seperti ini menjadikan dilema dalam pengelompokan, karena kondisi ekonomi mahasiswa baru setiap tahunnya akan berebeda. Hal ini berdampak pada setiap tahunnya mahasiswa baru banyak melakukan banding karena pengelompokan UKT yang dihasilkan masih dianggap tidak sesuai.

2.2 Analisis Solusi

Berdasarkan analisis persoalan yang telah dijabarkan di atas, maka pada penelitian ini diusulkan suatu gagasan untuk membuat sebuah aplikasi untuk mengatasi persoalan tersebut. Sistem yang akan dikembangkan ini merupakan sebuah sistem yang menjadi pendukung keputusan pengelompokan UKT di Institut Teknologi Sumatera. Sistem pendukung keputusan ini diharapkan dapat mengelompokkan UKT sesuai dengan tingkat ekonomi mahasiswa.

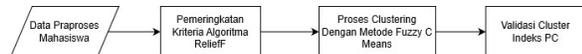
Pada penelitian kali ini, metode yang akan digunakan adalah Fuzzy C-Means. Fuzzy C-Means merupakan algoritma yang mampu mengelompokkan objek-objek yang memiliki tingkat kemiripan yang sama menjadi satu kelompok dengan pembagian kelompok UKT menjadi 8 kelompok. Dalam clustering dibutuhkan kriteria-kriteria yang optimal sehingga clustering yang dihasilkan optimal. Pada penelitian ini, untuk menentukan kriteria digunakan feature selection. feature selection merupakan suatu kegiatan pemodelan atau penganalisaan data yang umumnya dapat dilakukan secara praproses dan bertujuan untuk memilih kriteria yang berpengaruh (fitur optimal) dan mengesampingkan kriteria yang tidak berpengaruh [2]. Ada beberapa algoritma feature selection yang dapat digunakan. untuk menemukan kriteria yang optimal dari sebuah himpunan fitur. Salah satu algoritma Feature Selection adalah algoritma ReliefF. ReliefF merupakan pengembangan dari algoritma

relief. ReliefF dapat mengevaluasi nilai fitur dengan berulang kali mengambil sampel instance dan mempertimbangkan nilai fitur yang diberikan untuk instance terdekat dari kelas yang sama dan yang berbeda. Evaluasi atribut ini memberikan bobot untuk masing-masing fitur berdasarkan kemampuan fitur untuk membedakan antar kelas, dan kemudian memilih fitur-fitur yang nilainya melebihi threshold yang ditetapkan sebelumnya sebagai fitur yang relevan. Hasil yang di dapatkan pada algoritma ReliefF ini akan berupa estimasi bobot dari setiap kriteria, kemudian setiap kriteria akan dilakukan pemeringkatan dari bobot tertinggi ke terendah. Setelah itu akan dilakukan uji validasi indeks partition coefficient setiap.

kombinasi kriteria pada proses clustering Fuzzy C-Means. Maka hasil validasi indeks partition coefficient tertinggi merupakan kombinasi kriteria terbaik dalam proses clustering UKT.

2.3 Perancangan Sistem

Penjelasan Proses dari penerapan metode Fuzzy C-Means Clustering untuk menentukan golongan uang kuliah tunggal pada mahasiswa baru secara umum digambarkan dalam blok diagram pada Gambar 3.1.



Gambar 3. 1 Diagram Blok Sistem

2.3.1 Data Mahasiswa

Data yang dipakai merupakan data induk mahasiswa (DIM) tahun ajaran 2017. Terdapat 37 kriteria data training sebelum dilakukan praproses data. Data akan dilakukan inisialisasi untuk setiap data kategori menjadi numerik karena data akan dilakukan clustering dengan metode Fuzzy C-Means. Inisialisasi dilakukan dengan asumsi nilai yang lebih tinggi dianggap mahasiswa memiliki finansial yang lebih baik.

Praproses selanjutnya yaitu setiap data dilakukan normalisasi pada rentang [0.0 , 1.0] dengan metode normalisasi Min-Max untuk mengurangi waktu komputasi pada proses clustering. Data dilakukan transformasi linier terhadap data asli.

$$X' = \frac{(X - \text{Min}) * (\text{Newmax} - \text{Newmin})}{(\text{Max} - \text{Min}) + \text{Newmin}}$$

X' = Data hasil normalisasi

X = Data awal

Min = Nilai minimum data per kolom

Max = Nilai maksimum data per kolom

Newmin = Batas minimum yang akan diberikan

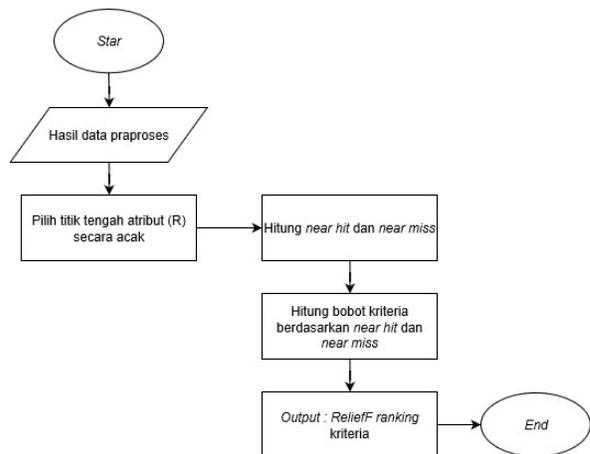
Newmax = Batas maksimum yang akan diberikan

2.3.2 Proses Pemeringkatan Fitur ReliefF

Setelah data kriteria dilakukan praproses, setiap kriteria akan dihitung estimasi bobot dengan algoritma feature selection ReliefF. Pemeringkatan kriteria dengan ReliefF digunakan untuk mengetahui kriteria-kriteria yang paling berpengaruh terhadap UKT yang telah ditetapkan sebelumnya oleh pihak PPMB ITERA.

Gambar 3. 2 Urutan Proses Pemeringkatan Fitur dengan ReliefF

Gambar 3.2 merupakan urutan proses ReliefF. Untuk mendapatkan score dari setiap kriteria, ReliefF memanfaatkan nilai near miss dan near hit. Pada proses ReliefF, sebelum mencari nilai near miss dan



near hit, perlu ditentukan terlebih dahulu nilai tengahnya (R). Penentuan nilai tengah dilakukan secara acak. Setelah nilai R ditentukan, nilai near miss dan nilai near hit diketahui, maka score fitur dapat dihitung. Output pada proses ini adalah urutan ranking dengan melihat setiap estimasi bobot kriteria.

2.3.3 Perhitungan Bobot ReliefF

Berikut merupakan contoh tahap perhitungan ReliefF. Diketahui sebuah training set D, jumlah iterasi m, jumlah nearest neighbors k, jumlah fitur n dilihat pada Tabel 3.2 berikut:

Tabel III. 2 Tabel Set Data D

Nama	Kriteria		Kelas
	Penghasilan	Kendaraan	
Irvan	2,5	3	gol 2
Annisa	2	2	gol 1
Prasetyo	0,5	1	gol 1
Falah	4	5	gol 2
Afif	3,5	2	gol 2
Najib	1	1,5	gol 1

1. Diasumsikan $S=\emptyset$, set semua bobot fitur $W(Ft)=0$, dan $k=1$
2. Pilih sample R dari D secara acak.
3. Cari k nearest neighbors H_i dari kelas yang sama dan k nearest neighbors $M_i(C)$ dari setiap kelas C yang berbeda.
4. Hitung bobot untuk setiap fitur dengan persamaan (2.1) : $\text{diff}(Ft,R1,R2)$ untuk setiap data kontinu maka akan dihitung dengan persamaan (2.3). Setelah itu untuk memudahkan perhitungan maka hitung $|R1 [Ft] - R2[Ft]|$ untuk H_i dan $(\frac{P(C)}{1-P(Class(R))}) * |R1[Ft] - R2[Ft]|$ untuk $M_i(C)$ terlebih dahulu. Berikut perhitungan $|R1 [Ft] - R2[Ft]|$ tiap kriteria dilihat pada Tabel 3.3 dan Tabel 3.4

Tabel 3. 3 Tabel Hit dan Miss Penghasilan

Hits untuk Penghasilan	Miss untuk Penghasilan
2,5 - 3,5	$\frac{\frac{3}{6}}{1-\frac{3}{6}} * 2,5 - 2 $
2 - 1	$\frac{\frac{3}{6}}{1-\frac{3}{6}} * 2 - 2,5 $
0,5 - 1	$\frac{\frac{3}{6}}{1-\frac{3}{6}} * 0,5 - 2,5 $
4 - 3,5	$\frac{\frac{3}{6}}{1-\frac{3}{6}} * 4 - 2 $
3,5 - 4	$\frac{\frac{3}{6}}{1-\frac{3}{6}} * 3,5 - 2 $
1 - 0,5	$\frac{\frac{3}{6}}{1-\frac{3}{6}} * 1 - 2,5 $

Hits untuk Penghasilan	Miss untuk Penghasilan
1	0,5
1	0,5
0,5	2
0,5	2
0,5	1,5
0,5	1,5

Tabel 3. 4 Tabel Hit dan Miss Kendaraan

Hits untuk Kendaraan	Miss untuk Kendaraan
3 - 2	$\frac{\frac{3}{6}}{1-\frac{3}{6}} * 3 - 2 $
2 - 1,5	$\frac{\frac{3}{6}}{1-\frac{3}{6}} * 2 - 2 $
1 - 1,5	$\frac{\frac{3}{6}}{1-\frac{3}{6}} * 1 - 2 $
5 - 3	$\frac{\frac{3}{6}}{1-\frac{3}{6}} * 5 - 2 $
2 - 3	$\frac{\frac{3}{6}}{1-\frac{3}{6}} * 2 - 2 $
1,5 - 2	$\frac{\frac{3}{6}}{1-\frac{3}{6}} * 1,5 - 2 $

Hits untuk Kendaraan	Miss untuk Kendaraan
1	1
0,5	0
0,5	1
2	3
1	0
0,5	0,5

Kemudian hitung total $|R1[Ft] - R2[Ft]|$ hits dan miss tiap fitur. Tabel 3.5 merupakan total $|R1[Ft] - R2[Ft]|$ hits dan miss.

Tabel 3. 5 Tabel Total Perhitungan Hit dan Miss

Penghasilan	Total	Kendaraan	Total
HIT	4	HIT	5,5
MISS	8	MISS	5,5

Hitung tiap *nearest Hits* penghasilan :

$$\sum_{i=1}^k \frac{\text{diff}(Ft,R,H_i)}{(mk)} = \frac{4}{6*1} = 0.1904$$

Hitung tiap *nearest Miss* penghasilan:

$$\sum_{c \in \text{ClassR}} \frac{(\frac{P(C)}{1-P(Class(R))} \sum_{i=1}^k \text{diff}(Ft,R,M_i(C)))}{(mk)} = \frac{8}{6*1} = 0.3809$$

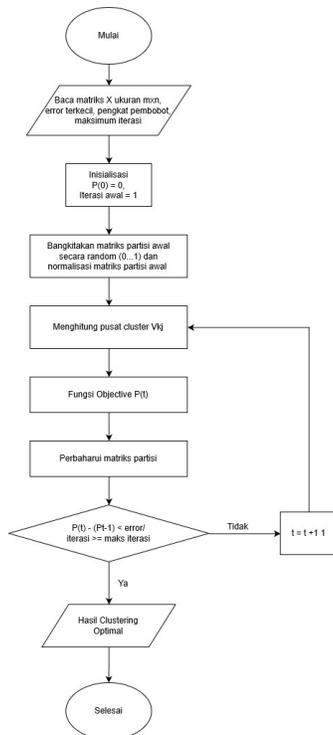
Maka $W[Ft]$ penghasilan = $0,1904 + 0,3809 = 0,5714$

2.3.4 Proses Clustering Fuzzy C-Means

Pada tahapan ini dilakukan proses utama yaitu segmentasi data kriteria mahasiswa dengan menggunakan metode Fuzzy C-Means. Setelah didapatkan hasil estimasi bobot setiap kriteria pada proses algoritma ReliefF akan dilakukan uji clustering untuk setiap kombinasi kriteria. Pengujian dilakukan dengan cara eliminasi setiap threshold bobot. Setiap threshold akan dilakukan clustering untuk mengetahui kombinasi kriteria yang memiliki nilai tertinggi dalam proses validasi cluster.

Diagram flowchart pada Gambar 3.3, menunjukan beberapa langkah yang dilakukan oleh algoritma Fuzzy C-Means dalam melakukan clustering, berikut penjelasannya:

1. Bagian admin melakukan set banyaknya iterasi dan nilai error yang akan dilakukan.
2. Sistem membuka data inputan berupa data eksperimen kombinasi kriteria.
3. Langkah selanjutnya yaitu sistem proses algoritma Fuzzy C-Means.
4. Output dari hasil pengelompokan adalah daftar alternatif mahasiswa dengan status kelompok 1-8.



Gambar 3. 3 Flowchart Algoritma Fuzzy C-Means

2.3.5 Perhitungan Fuzzy C-Means

Berikut merupakan contoh tahap penyelesaian metode Fuzzy C-Means.

1. Sejumlah data set yang digunakan untuk mencari kesejahteraan orang tua dari kriteria penghasilan dan kendaraan dilihat pada Tabel 3.6.

Tabel 3. 6 Tabel Set Data X

Nama	Kriteria	
	Penghasilan	Kendaraan
Irvan	3	3
Annisa	2	2
Prasetyo	1	1
Falah	3	5

2. Melakukan seting kondisi awal variable perhitungan dari Fuzzy C-Means seperti menentukan banyak cluster, kondisi bobot pemangkat, maksimum error, maksimum iterasi, dan iterasi awal. Tabel 3.7 ini merupakan contoh seting variabel awal dari perhitungan Fuzzy C-Means.

Tabel 3. 7 Tabel Setting Variabel

Jumlah Cluster	3
Max Iterasi	100
Pembobot	2
Nilai Error	0.000001

3. Mulai membentuk nilai keanggotaan (μ_{ik}) secara acak dengan asumsi nilai acak dengan total pada setiap baris data selalu bernilai 1. Tabel 3.8 merupakan contoh tabel nilai keanggotaan (μ_{ik}) yang didapat secara random dari contoh perhitungan Fuzzy C-Means.

Tabel 3. 8 Tabel Buat μ_{ik}

	Keanggotaan Cluster (random)			
	C1	C2	C3	Jumlah
Irvan	0.5	0	0.5	1
Annisa	0.5454	0.4545	0	1
Prasetyo	0.5384	0.3846	0.0769	1
Falah	0.3333	0	0.6667	1

4. Mulai Iterasi

5. Hitung nilai pusat vector V_{kj} dengan dimensi k = banyak criteria(c) dan j banyak attribut dengan menggunakan persamaan (2.5)

Pusat Cluster (V)	
2.075	2.309
1.582	1.582
2.983	4.252

Tahap pertama penghitungan pusat cluster kita hitung dengan nilai vektor μ_{ik}^2 agar perhitungan pusat cluster akan menjadi mudah. Tabel 3.9 merupakan contoh tabel perhitungan nilai keanggotaan (μ_{ik}) kwadrat dari contoh perhitungan dari Fuzzy C- Means.

Tabel 3. 9 Tabel Hitung μ_{ik}^2

Miu Kuadrat (μ_{ik}) ²		
0.25	0	0.25
0.2975	0.2066	0
0.2899	0.1479	0.00591
0.1111	0	0.44444
Total 0.9845	Total 0.35454	Total 0.70036

Setelah nilai (μ_{ik})² didapatkan, maka dihitung dulu nilai (μ_{ik})² x X_{ij} untuk dicari total dari setiap kolom (μ_{ik}) agar $\Sigma(\mu_{ik})^2$ x X_{ij} . Tabel 3.13 merupakan contoh tabel perhitungan nilai vector pusat (V) dengan persamaan (2.5) dari contoh perhitungan Fuzzy C- Means. Berikut Tabel 3.10 merupakan tabel hitung V.

Tabel 3. 10 Tabel Hitung V

	(μ_{i1}) ² x X_{i1}	(μ_{i1}) ² x X_{i2}	(μ_{i2}) ² x X_{i1}	(μ_{i2}) ² x X_{i2}	(μ_{i3}) ² x X_{i1}	(μ_{i3}) ² x X_{i2}
1	0.75	0.75	0	0	0.75	0.75
2	0.595	0.595	0.413	0.413	0	0
3	0.289	0.289	0.147	0.147	0.005	0.005
4	0.333	0.555	0	0	1.333	2.222
Total	1.968	2.190	0.561	0.561	2.089	2.978

Pusat Cluster (V)	
1.96/0.98	2.19 /0.98
0.56/0.35	0.56/0.35
2.08/0.70	2.97/0.70

Hitung fungsi objektif P dari perhitungan metode Fuzzy C-Means yang telah dilakukan menggunakan persamaan (2.6). Tahap pertama penghitungan Nilai Objektif kita hitung dulu nilai vektor $(X_{ij}-V_{kj})^2$ (X_V) agar perhitungan nilai objektif iterasi akan menjadi mudah. Tabel 3.11 merupakan contoh tabel perhitungan nilai X_V dari contoh perhitungan Fuzzy C-Means.

Tabel 3. 11 Tabel Hitung X_V

($X_{ij}-V_{kj}$) ² +($X_{ij}-V_{kj}$) ²	($X_{ij}-V_{kj}$) ² +($X_{ij}-V_{kj}$) ²	($X_{ij}-V_{kj}$) ² +($X_{ij}-V_{kj}$) ²
(3-2.075) ² +(3-2.309) ²	(3-1.582) ² +(3-1.582) ²	(3-2.833) ² +(3-2.157) ²
(2-2.075) ² +(2-2.309) ²	(2-1.582) ² +(2-1.582) ²	(2-2.833) ² +(2-2.157) ²
(1-2.075) ² +(1-2.309) ²	(1-1.582) ² +(1-1.582) ²	(1-2.833) ² +(1-2.157) ²
(3-2.075) ² +(5-2.309) ²	(3-1.582) ² +(5-1.582) ²	(3-2.833) ² +(5-2.157) ²

($X_{ij}-V_{kj}$) ² +($X_{ij}-V_{kj}$) ²	($X_{ij}-V_{kj}$) ² +($X_{ij}-V_{kj}$) ²	($X_{ij}-V_{kj}$) ² +($X_{ij}-V_{kj}$) ²
1.33263	4.01714	1.56855
0.10129	0.34818	6.03929
2.86995	0.67921	14.5100
8.09544	13.6861	0.55935

Hasil perhitungan $\Sigma(X_{ij}-V_{kj})^2$ akan dikalikan (μ_{ik})² pada yang telah dihitung pada Table 3.9. Berikut adalah contoh perhitungan $\Sigma(X_{ij}-V_{kj})^2 * (\mu_{ik})^2$ (L) untuk mendapatkan $\Sigma[\Sigma(X_{ij}-V_{kj})^2 * (\mu_{ik})^2]$ (Total L) dapat dilihat pada Tabel 3.12 dan Tabel 3.13

Tabel 3. 12 Tabel Hitung L

L1	L2	L3
1.33263*0.25	4.01714*0	1.56855*0.25
0.10129*0.297	0.34818*0.206	6.03929*0
2.86995*0.289	0.67921*0.147	14.5100*0.005
8.09544*0.111	13.6861*0	0.55935*0.44

Tabel 3. 13 Tabel Total L

L1	L2	L3	Total L
0.333159	0	0.392127	0.7252
0.030137	0.071938	0	0.1020
0.832115	0.100475	0.085858	1.0184
0.899493	0	0.248603	1.1480
Total P			2.9939

Pada iterasi pertama didapat hasil nilai fungsi objektif (P) sebesar: 2.9939

Kemudian dilakukan contoh penghitungan nilai fungsi objektif iterasi ke-1 dari data dan nilai yang dibangkitkan serta nilai pusat cluster yang telah dijelaskan pada poin sebelumnya.

7. Kemudian dilakukan pengecekan pemberhentian iterasi dengan syarat kondisi untuk berhenti, yaitu:

Jika $(|Pt-Pt-1| < \xi)$ atau $(t > \text{MaxIter})$ maka berhenti

- Jika tidak: $t=t+1$, mengulang langkah ke-5 dengan melakukan langkah 9 dulu.
- Disini karena iterasi masih $<$ maksimal iterasi,
- maka dilakukan pengecekan P.
- karena P sekarang $>$ error epsilon yaitu sebesar 2.9939 maka dilakukan langkah selanjutnya.

8. Hitung nilai perubahan dari nilai yang dibangkitkan menggunakan persamaan (2.7). Untuk memperoleh nilai (μ_{ik}) baru, maka dilakukan perhitungan terlebih dahulu nilai $\sum(X_{ij}-V_{kj})^2$ (X_V) pada Tabel 3.14 berikut :

Tabel 3. 14 Tabel Hitung X_V (2)

$(X_{ij}-V_{k1})^2+(X_{ij}-V_{k2})^2$	$(X_{ij}-V_{k2})^2+(X_{ij}-V_{k3})^2$	$(X_{ij}-V_{k3})^2+(X_{ij}-V_{k1})^2$
1.33263	4.01714	1.56855
0.10129	0.34818	6.03929
2.86995	0.67921	14.5100
8.09544	13.6861	0.55935

Kemudian hitung nilai $[\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{2-1}$. Berikut adalah contoh perhitungan dari nilai $[\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{2-1}$ (LT) dapat dilihat pada Tabel 3.15.

Tabel 3.5 Tabel Hitung LT

$[\sum_{j=1}^m (X_{ij} - V_{k1})^2]^{2-1}$	$[\sum_{j=1}^m (X_{ij} - V_{k2})^2]^{2-1}$	$[\sum_{j=1}^m (X_{ij} - V_{k3})^2]^{2-1}$
$[1.3326]^{2-1}$	$[4.0171]^{2-1}$	$[1.5685]^{2-1}$
$[0.1012]^{2-1}$	$[0.3481]^{2-1}$	$[6.0392]^{2-1}$
$[2.8699]^{2-1}$	$[0.6792]^{2-1}$	$[14.510]^{2-1}$
$[8.095]^{2-1}$	$[13.686]^{2-1}$	$[0.5593]^{2-1}$

9. Lakukan langkah 5 sampai nilai $P <$ epsilon atau iterasi \geq maksimal iterasi

2.3.6 Perhitungan Indeks Partition Coefficient

Setiap hasil kombinasi Kriteria yang diuji pada proses clustering Fuzzy C-Means selesai maka di dapat nilai keanggotaan baru (μ_{ij}) . Kemudian digunakan Indeks PC untuk mengukur validitas cluster dengan

persamaan (2.8). Kombinasi dalam penentuan kriteria terbaik merupakan kombinasi kriteria yang memiliki nilai validitas tertinggi.

Berikut merupakan contoh tahap perhitungan indeks PC.

1. Hasil dari keanggotaan proses *clustering* didapatkan dari sampel Tabel 3.18.

Tabel 3. 18 Tabel Keanggotaan (PC)

μ_{i1}	μ_{i2}	μ_{i3}
0.6221	0.1099	0.2678
0.9268	0.0620	0.0110
0.0141	0.9834	0.0024
0.0121	0.0055	0.9823

2. Selanjutnya akan ditentukan (μ_{ij}^2) dari keanggotaan. Berikut (μ_{ij}^2) dilihat pada Tabel 3.19.

Tabel 3. 19 Tabel Hitung (μ_{ij}^2) (PC)

μ_{i1}	μ_{i2}	μ_{i3}
$(0.6221)^2$	$(0.1099)^2$	$(0.2678)^2$
$(0.9268)^2$	$(0.0620)^2$	$(0.0110)^2$
$(0.0141)^2$	$(0.9834)^2$	$(0.0024)^2$
$(0.0121)^2$	$(0.0055)^2$	$(0.9823)^2$

μ_{i1}	μ_{i2}	μ_{i3}
0.3870	0.0120	0.0717
0.8589	0.0038	0.0001
0.0002	0.9670	0.000005
0.0001	0.00003	0.9649

3. Kemudian penjumlahan setiap data (μ_{ij}^2) akan dibagi dengan jumlah data.

$$PC(c) = \frac{1}{N} \left(\sum_{i=1}^c \sum_{j=1}^N (\mu_{ij}^2) \right)$$

$$PC = \frac{3.265735}{5} = 0,65$$

Maka indeks PC yang didapat adalah 0,65

3. IMPLEMENTASI DAN PENGUJIAN

3.1 Perangkat Pembuatan Sistem

Clustering Fuzzy C-Means dilakukan untuk menentukan kelompok UKT pada mahasiswa baru dengan perangkat yang memiliki spesifikasi sebagai berikut :

1. Laptop yang digunakan dengan spesifikasi - Intel Core i7
2. Sistem Operasi Windows 10
3. Database Mysql XAMPP Bundled
4. Microsoft Office 2016
5. Weka 3.8

3.2 Skema Penentuan Uang Kuliah Tunggal

3.2.1 Praproses Data

Data yang digunakan dalam penelitian ini merupakan Data Induk Mahasiswa (DIM) tahun ajaran 2017 yang terdiri dari 544 field mahasiswa. Terdapat praproses yang dilakukan sebelum proses seleksi kriteria dilakukan. Beberapa tahapan praproses yaitu menghilangkan setiap kriteria yang tidak dapat dibandingkan seperti *kd_jurusan*, *jenis_kelamin*, *gol_darah*, *anak_ke*, *jns_bayar_listrik*, *id_agama*, *penguasaan_komputer*, *hubungan_tg_jwb* dan *jurusan_sma*. Kemudian menghilangkan kriteria yang 80% memiliki nilai nol nilai seperti *pengeluaran_kesehatan*, dan *penghasilan_tambahan_ayah* dan *ibu*, serta dilakukan eliminasi untuk kriteria yang memiliki nilai yang sama seperti kriteria *kd_penghasilan_ayah* dan *kd_penghasilan_ibu* yang sudah terdapat kriteria *penghasilan_ayah* dan *penghasilan_ibu* sehingga didapatkan 22 kriteria yang akan dilakukan proses clustering. Tabel 4.1 merupakan 22 kriteria yang akan dilakukan proses cluster.

Tabel 4.1 Data Kriteria

Data Kriteria			
1	Pilihan	12	<i>kd_pekerjaan_ibu</i>
2	<i>jns_tinggal</i>	13	<i>penghasilan_ibu</i>
3	<i>jml_saudara</i>	14	<i>Bidikmisi</i>
4	<i>kd_pendidikan_tg_jwb</i>	15	<i>pengeluaran_makan_transport</i>
5	<i>penghasilan_tg_jwb</i>	16	<i>status_rumah</i>
6	<i>kd_pekerjaan_tg_jwb</i>	17	<i>Biling</i>
7	<i>jml_tanggungan_ortu</i>	18	<i>ikut_asrama</i>
8	<i>kd_pendidikan_ayah</i>	19	<i>Njopbangunanpbb</i>
9	<i>kd_pekerjaan_ayah</i>	20	<i>Luasbangunanpbb</i>
10	<i>penghasilan_ayah</i>	21	<i>Njopbumipbb</i>
11	<i>kd_pendidikan_ibu</i>	22	<i>Luasbumipbb</i>

Data dilakukan inisialisasi untuk setiap data kategori menjadi numerik dengan asumsi nilai yang lebih tinggi dianggap mahasiswa memiliki finansial yang lebih baik. Berikut merupakan inisialisasi data categorical pada Tabel 4.2.

Tabel 4.2 Inisialisasi data *categorical*

<i>penguasaan_komputer</i>	inisialisasi
Tidak Terampil	1
Sedang	2
Terampil	3

<i>status_rumah</i>	inisialisasi
Kontrak	1
Menumpang	2
Rumah-Dinas	3
Milik-Sendiri	4

<i>jns_byr_listrik</i>	inisialisasi
Pra Bayar	1
Pasca Bayar	2

<i>ikut_asrama</i>	inisialisasi
Tidak	1
Ya	2

<i>biling</i>	inisialisasi
Ya	1
Tidak	2

Data yang sebelumnya telah dilakukan inisialisasi oleh pihak UPT TIK dilakukan penyesuaian kembali dengan asumsi nilai yang lebih tinggi dianggap mahasiswa memiliki finansial yang lebih baik. Kemudian data kriteria yang inisialisasi memiliki rentang nilai yang jauh (98,99) seperti data 'lainnya' diubah menjadi nilai rata rata dari di kolom kriteria tersebut .Berikut merupakan penyesuaian inisialisasi data categorical pada Tabel 4.3.

Tabel 4. 3 Penyesuaian Data yang Telah di Inisialisasi

kd_pekerjaan	Nilai
Sudah Meninggal	0
Tidak Bekerja	1
Buruh	2
Nelayan	3
Petani	4
Peternak	5
Pedagang Kecil	6
Pensiunan	7
PNS/TNI/Polri	8
Karyawan Swasta	9
Pedagang Besar	10
Wiraswasta	11
Wirasaha	12
Lainnya	6

kd_pendidikan	Nilai
Tidak Sekolah	0
PAUD	1
TK/ sederajat	2
Putus SD	3
Paket A	4
SD/ sederajat	5
Paket B	6
SMP/ sederajat	7
Paket C	8
SMA/ sederajat	9
D1	10
D2	11
D3	12
D4	13
Profesi	14

kd_pendidikan	Nilai
S1	15
Sp-1	16
S2	17
Sp-2	18
S3	19
Non Formal	9,5
Informal	9,5
(tidak diisi)	9,5
Lainnya	9,5

Praproses selanjutnya yaitu setiap data dilakukan normalisasi pada rentang [0.0 , 1.0] dengan metode normalisasi Min-Max untuk mengurangi waktu komputasi pada proses clustering. Sampel hasil data sebelum dan sesudah praproses dapat dilihat pada lembar Lampiran Sampel Data Kriteria.

Setelah dilakukan praproses, data praproses disimpan dalam file teks dengan attribute-relation file format (arff). Tahapan selanjutnya yaitu seleksi kriteria dengan metode ReliefF untuk menentukan ranking dan bobot dari setiap kriteria.

Hasil ranking kriteria akan dilatih untuk mendapatkan nilai validasi partition coefficient terbaik. Eksperimen akan dilakukan dengan menggunakan metode clustering Fuzzy C-Means. Pengujian ini dilakukan dengan mencoba beberapa threshold dengan harapan mencapai nilai optimum pada validasi partition coefficient pada proses clustering Fuzzy C-Means. Beberapa kriteria terbaik akan dilakukan kombinasi dengan kriteria yang digunakan oleh pihak itera dalam menentukan kelompok UKT. Hasil terbaik dari eksperimen uji kriteria merupakan kombinasi kriteria terbaik dalam penentuan UKT.

3.2.2 Estimasi Bobot Kriteria Dengan ReliefF

Untuk menentukan seleksi kriteria yang terpilih dilakukan penentuan kriteria dengan menguji seluruh data induk mahasiswa yang dapat dibandingkan dengan menggunakan algoritma feature selection ReliefF. Pada penelitian ini penentuan estimasi bobot ReliefF menggunakan library Weka 3.8. Berikut data seleksi dan ranking kriteria dengan ReliefF. dapat dilihat pada Tabel 4.4 dan Tabel 4.5

Tabel 4. 5 Data Ranking Kriteria Dengan ReliefF

No	Kriteria	Bobot
1	Bidiknisi	0.256405
2	Biling	0.199308
3	kd pekerjaan tg_jwb	0.171078
4	penghasilan tg_jwb	0.147673
5	kd pendidikan tg_jwb	0.14454
6	kd pekerjaan ayah	0.139858
7	penghasilan ayah	0.13379
8	kd pendidikan ayah	0.126106
9	kd pendidikan ibu	0.104202
10	pengeluaran makan transport	0.099488
11	kd pekerjaan ibu	0.086858
12	penghasilan ibu	0.086149
13	Njobumipbb	0.030862
14	Luasbangunanpbb	0.024853
15	Njobangunanpbb	0.024853
16	ikut asrama	0.018306
17	status_rumah	0.013522
18	jml saudara	0.003019
19	Pilihan	0.00211
20	Luasbumipbb	0.002042
21	ins tinggal	0.001928
22	jml tanggungan_ortu	-0.000748

Hasil Tabel 4.5 menunjukkan ranking dari bobot dari data seleksi kriteria ReliefF. Nilai bobot kriteria yang dihasilkan menunjukkan kuat atau tidaknya hubungan linear pada kriteria UKT (kelas yang dituju). Nilai bobot yang mendekati -1 atau +1 menunjukkan hubungan yang kuat antara dua kriteria tersebut dan nilai bobot yang mendekati 0 mengindikasikan lemahnya hubungan antara dua variabel tersebut. Sedangkan tanda + (positif) dan - (negatif) memberikan informasi mengenai arah hubungan antara kriteria tersebut. Jika bernilai + (positif) maka kedua variabel tersebut memiliki hubungan yang searah. Dalam arti lain peningkatan X akan bersamaan dengan peningkatan Y dan begitu juga sebaliknya. Jika bernilai - (negatif) artinya korelasi antara kedua kriteria tersebut bersifat berlawanan. Peningkatan nilai X akan dibarengi dengan penurunan Y.

Hasil dari ranking bobot ReliefF ini akan dilanjutkan pada tahap eksperimen dengan kombinasi kriteria dengan uji eliminasi threshold yang di implementasikan pada pengelompokan Fuzzy C-Means. Hasil dari ranking bobot ReliefF ini akan dilanjutkan pada tahap eksperimen dengan kombinasi kriteria dengan uji eliminasi threshold yang di implementasikan pada pengelompokan Fuzzy C-Means.

3.2.3 Proses Fuzzy C-Means Clustering

Seperti yang telah dijelaskan di akhir bab estimasi bobot ReliefF, setiap kombinasi kriteria akan dilakukan proses clustering untuk mendapatkan tujuan dalam penelitian ini yaitu kombinasi kriteria yang terbaik serta hasil cluster yang terbentuk. Untuk mencari kombinasi kriteria yang terbaik pada penelitian ini akan diuji tiap kombinasi kriteria dengan cara menaikkan nilai threshold dari threshold yang terendah hingga mendapatkan nilai validasi cluster tertinggi.

Tahapan pertama dalam Fuzzy C-Means ini yaitu dilakukan Inisialisasi terlebih dahulu. Jumlah Cluster ($C = 8$) karena pengelompokan UKT di ITERA dibagi menjadi 8 cluster, Pangkat ($w = 2$), Maksimum iterasi ($MaxIter = 100$), Error terkecil yang diharapkan ($\xi = 0,0001$), Fungsi Obyektif awal ($P_0 = 0$). Kemudian bentuk matriks partisi awal secara acak, berdasarkan persamaan (2.4) dengan syarat jumlah derajat keanggotaan setiap data pada setiap kolom harus bernilai 1. Setelah mendapatkan matriks awal μ_0 , kemudian hitung pusat cluster V_0 dengan menggunakan persamaan (2.5). Setelah didapatkan pusat cluster V_0 , hitung nilai fungsi obyektif menggunakan persamaan (2.6). Kemudian dilakukan perbaikan derajat keanggotaan / matriks μ (matriks partisi) berdasarkan persamaan (2.7). Setelah mendapatkan matriks baru (μ_1), selanjutnya dilakukan pengecekan kondisi berhenti dengan melihat selisih fungsi obyektif terhadap epsilon $|P_t - P_{t-1}| < \zeta$. Apabila nilai fungsi obyektif sudah lebih kecil dari pada nilai maksimum error yang diberikan maka hasil cluster sudah dianggap konvergen. Kemudian nilai keanggotaan μ akhir yang terbentuk akan dilakukan validasi cluster dengan indeks PC.

3.2.4 Eksperimen Hasil Data Ranking ReliefF Pada Fuzzy C-Means

Eksperimen dilakukan untuk mencari nilai indeks partition coefficient terbaik. Hasil eksperimen uji kriteria dapat dilihat pada Tabel 4.17

Tabel 4. 17 Hasil Eksperimen Kombinasi Kriteria

No	Kriteria	Indeks PC
1	D1	0,1251
2	D2	0,1253
3	D3	0,1255
4	D4	0,4645
5	D5	0,5048
6	D6	0,41
7	D7	0,6394
8	D8	0,4972
9	D9	0,5537
10	D10	0,625
11	D11	0,3011
12	D12	0,3051

Eksperimen diawali dengan threshold rendah yaitu melibatkan seluruh kriteria (D1), threshold 0.01 (D2), dan threshold 0.02 (D3) menunjukkan bahwa kriteria dengan threshold yang rendah tidak memiliki banyak pengaruh terhadap perubahan indeks PC. Perubahan yang terjadi pada kriteria D1,D2,D3 kecil berbanding lurus dengan perubahan bobot ReliefF. Untuk mencari nilai indeks PC yang signifikan maka pada percobaan selanjutnya dilakukan dengan menggunakan threshold 0.08 (D4) dan 0.1 (D5). Penentuan threshold 0.08 didasari rentang yang cukup jauh dari bobot sebelumnya yaitu 0.03. Hasil eksperimen uji kriteria D4 dan D5.

terjadi peningkatan indeks PC yang signifikan jika dibandingkan percobaan sebelumnya. Hasil terbaik didapatkan pada D5 yang merupakan indeks dengan threshold 0.1. Pada D5 hanya tersisa kombinasi dari kriteria biling-bidikmisi, penghasilan, kd_pekerjaan, kd_pendidikan, dan pengeluaran_makan_transport. Sehingga pada percobaan selanjutnya dilakukan percobaan D5 tanpa mengikut sertakan salah satu biling-bidikmisi, kd_pendidikan, kd_pekerjaan, pengeluaran_makan_transport dan penghasilan.

D6 = D5 tanpa biling dan bidikmisi

D7 = D5 tanpa kd_pendidikan_ayah, kd_pendidikan_ibu, dan kd_pendidikan_tg_jwb

D8 = D5 tanpa kd_pekerjaan_tg_jwb dan kd_pekerjaan_ayah

D9 = D5 tanpa pengeluaran_makan_transport

D10 = D5 tanpa penghasilan_ayah dan penghasilan_tg_jwb

Dapat dilihat pada Tabel 4.17 terjadi penurunan indeks PC pada kombinasi kriteria D5 tanpa biling dan bidikmisi, namun data uji yang lain menunjukkan peningkatan indeks PC. Peningkatan indeks PC tertinggi dari kelanjutan percobaan data D5 merupakan data D7 yaitu kombinasi D5 tanpa kd_pendidikan. Percobaan selanjutnya dilakukan perbandingan dengan nilai indeks PC menggunakan kombinasi kriteria yang telah diterapkan di itera (D11) dengan hasil eksperimen D7 dan menggabungkan D7 dan D11 (D12).

D7 = D5 tanpa kd_pendidikan

D11 = Kombinasi kriteria di ITERA (penghasilan_tg_jwb, penghasilan_ayah, penghasilan_ibu, njobbangunanpbb, njobbumipbb, dan jml_tanggung_oru.

D12 = kombinasi kriteria D7+D11

Hasil percobaan D7 menghasilkan indeks PC yang lebih baik jika dibandingkan dengan kombinasi kriteria D11 dan D12 dengan perbedaan yang cukup signifikan. Oleh karena itu, kombinasi kriteria yang akan dipilih sebagai kombinasi kriteria terbaik adalah D7. Kombinasi tersebut adalah dengan menyeleksi kriteria penghasilan_tg_jwb, kd_pekerjaan_tg_jwb, kd_pekerjaan_ayah, penghasilan_ayah, bidikmisi, pengeluaran_makan_transport, dan bidikmisi dengan indeks PC 0.6394. Indeks PC dengan 8 cluster terbilang tidak terlalu baik untuk suatu cluster, hal ini disebabkan karena penentuan cluster yang mengharuskan cluster untuk terbentuk menjadi 8 cluster. Hal ini tercantum pada lampiran IV Peraturan Menteri Riset, Teknologi, dan Pendidikan Tinggi Nomor 39 Tahun 2016 tentang biaya uang kuliah tunggal dan uang kuliah tunggal pada perguruan tinggi negeri di lingkungan kementerian riset, teknologi, dan pendidikan tinggi. Tabel 4. 18 merupakan hasil cluster dari percobaan D7.

Tabel 4. 18 Hasil Cluster D7

C1	C2	C3	C4	C5	C6	C7	C8
64	78	23	150	64	76	56	33

Hasil cluster pada Tabel 4.18 merupakan hasil pengelompokan yang belum dilabeli kedalam kelompok UKT. Dapat dilihat bahwa cluster 4 (C4) merupakan cluster dengan jumlah terbanyak. Untuk melihat hasil perbandingan persebaran hasil cluster jika dibandingkan dengan hasil cluster UKT yang telah ditetapkan pihak PPMB ITERA dapat dilihat pada Tabel 4.19

Tabel 4. 19 Perbandingan Hasil *Cluster*

UKT 1	C1	C2	C3	C4	C5	C6	C7	C8
47	0	6	0	34	0	7	0	0
UKT 2	C1	C2	C3	C4	C5	C6	C7	C8
115	0	22	3	70	1	19	0	0
UKT 3	C1	C2	C3	C4	C5	C6	C7	C8
66	1	23	0	34	1	7	0	0
UKT 4	C1	C2	C3	C4	C5	C6	C7	C8
44	3	11	7	7	5	9	0	2
UKT 5	C1	C2	C3	C4	C5	C6	C7	C8
115	16	15	8	5	28	25	6	12
UKT 6	C1	C2	C3	C4	C5	C6	C7	C8
81	26	0	3	0	22	7	10	13
UKT 7	C1	C2	C3	C4	C5	C6	C7	C8
41	13	1	2	0	6	1	14	4
UKT 8	C1	C2	C3	C4	C5	C6	C7	C8
35	5	0	0	0	1	1	26	2

Hasil persebaran cluster terhadap hasil UKT yang telah ditetapkan pihak PPMB itera dapat dilihat bahwa persebaran tegas terlihat pada UKT 1 dan UKT 8 serta cluster yang saling beririsan merata terlihat pada UKT 4 dan UKT 5.

4. KESIMPULAN

Berdasarkan hasil penelitian terhadap sistem untuk menentukan kelompok UKT dengan Fuzzy C-Means, maka dapat disimpulkan :

1. Berdasarkan hasil eksperimen pengujian kombinasi indeks partition coefficient tertinggi diperoleh sebesar 0.6394 yang merupakan kombinasi dari kriteria penghasilan_tg_jwb, kd_pekerjaan_tg_jwb, kd_pekerjaan_ayah, penghasilan_ayah, bidikmisi, pengeluaran_makan_transport, dan biling.
2. Fuzzy C-Means dapat di implementasikan untuk menentukan kelompok UKT. Dari hasil pengelompokan dengan Fuzzy C-Means terbentuk C1 sebanyak 64, C2 sebanyak 78, C3 sebanyak 23, C4 sebanyak 150, C5 sebanyak 64, C6 sebanyak 76, C7 sebanyak 56, C8 sebanyak 33.

DAFTAR PUSTAKA

[1] Hariyanti Ni Kadek D., “Analisis Segemntasi Citra Satelit Menggunakan Metode Clustering

Fuzzy C-Means” Politeknik Negeri Bali, Bali, 2013.

[2] Kesuma Zurnila M., “Feature Selection Data Indeks Kesehatan Masyarakat Menggunakan Algoritma Relief” Biastatistics, vol. 6, 1-8, ISSN 1907-6274. 2011.

[3] Durgabai R.P.L., “Feature Selection using ReliefF Algorithm” International Journal of Advanced Research in Computer and Communication Engineering, 2014.

[4] Dash M., Liu H., “Feature Selection for Classification” Department of Information System & Computer Science, National University of Singapore, Singapore, 1997.

[5] Turban E., “Decision Support System and Expert System” Prentice Hall International, United States, 1995.

[6] Kusriani Mukhsin A., “Konsep dan Aplikasi Sistem Pendukung Keputusan,” Andi, Yogyakarta, 2015.

[7] Mulyanto A., “Sistem Informasi dan Konsep Aplikasi,” Pustaka Pelajar, Yogyakarta, 2009.

[8] A. O’Brien James, “Pengantar Sistem Informasi Perspektif Bisnis dan Manajerial” Jakarta: Salemba Empat, 2005.

[9] Novakovic J., “Toward Optimal Feature Selection Using Ranking Method and Classification Algorithm,” Yugoslav Journal of Operations Research, Serbia, 2011.

[10] Turban E.,”Decision Support Systems and Intelligent Systems” Andi, Yogyakarta, 2005.

[11] D.A. Irawan, Z.A. Baizal, E.G. Perdana, ”Analisis dan Implementasi Algoritma ReliefF untuk Feature Selection pada Klasifikasi Data Multiclass.” M.S. Thesis, Universitas Telkom, Jakarta, 2011.

[12] Muhamad F. P. B., “Perbaikan Prediksi Kesalahan Perangkat Lunak Menggunakan Seleksi Fitur dan Cluster-Based Classification” Institut Teknologi Sepuluh November, Surabaya, 2017.

[13] K. Kira and L. Rendell, "The Feature Selection Problem : Traditional Method and a New Algorithm," in AAAI-92 Proceedings, 1992.

[14] Klawonn F. dan Keller A, "Fuzzy Clustering and Fuzzy Rules". Science Journal, 1997. [15] Li Chun., "The Improved Partition Coefficient. Procedia Engineering" 24. 534-538. 10.1016/j.proeng.2011.11.2691, 2011.

[16] Yohannes., "Analisis Perbandingan Algoritma Fuzzy C-Means dan K- Means " Universitas Sriwijaya, Palembang, 2016.

[17] Gorunescu F., "Data Mining : Concepts, Models and Techniques". New York: Springer-Verlag, 2011. [

18] L.A. Zadeh, "Fuzzy Sets*". Department of Electrical Engineering and Electronics Research Laboratory, University of California, Berkeley, California, 1965.

[19] L. Wahyu, "Penerapan Sistem Fuzzy Untuk Diagnosis Campuran Bahan Bakar dan Udara Pada Mobil". F15 Gurt. Thesis, Universitas Negeri Yogyakarta, 2016.