

PEMETAAN BIDANG KEILMUAN MAHASISWA MENGUNAKAN METODE FUZZY C-MEANS (STUDI KASUS : PROGRAM STUDI TEKNIK INFORMATIKA UNIVERSITAS MUHAMMADIYAH GRESIK)

Rachmad Deo Abdika¹⁾ Indra Gita Anugrah²⁾

^{1,2)}Program Studi Teknik Informatika, Universitas Muhammadiyah Gresik
Jl. Sumatera No. 101 Gresik Kota Baru, Kec. Kebomas, Kab. Gresik.
e-mail:deoabdika01@gmail.com¹⁾, indragitaanugrah@umg.ac.id²⁾

(Naskah masuk : 25 Januari 2022 Diterima untuk diterbitkan : 30 November 2022)

ABSTRAK

Program Studi Teknik Informatika merupakan salah satu program studi yang ada di fakultas teknik Universitas Muhammadiyah Gresik. Program Studi Teknik Informatika sudah memiliki bidang keilmuan, bidang keilmuan yang telah ditentukan oleh prodi antara lain adalah Bidang Keilmuan *Artificial Intelegent*, Bidang Keilmuan *Rekayasa Perangkat Lunak* dan Bidang Keilmuan *Jaringan*. Clustering merupakan suatu teknik data mining dengan cara mengelompokkan data kedalam beberapa cluster sehingga data-data di suatu cluster memiliki kemiripan maksimum. Metode *Fuzzy C- Means*(FCM) yaitu salah satu metode pengclusteran data yang mana keberadaan masing- masing titik data dalam sesuatu cluster didetetapkan oleh derajat keanggotaan. Penelitian ini menghasilkan bobot optimal untuk Fuzzy C-Means, menghasilkan nilai *Silhouette Coefficient* 0,532 berkategori sedang dan menghasilkan nilai purity 0,806.

Kata Kunci: *Clustering, Fuzzy C-Means, Informatika*

ABSTRACT

The Informatics Engineering Study Program is one of the study programs in the architecture of the University of Muhammadiyah Gresik. The Informatics Engineering Study Program already has scientific fields, scientific fields that have been determined by the study program include the Artificial Intelligence Scientific Field, Software Engineering Scientific Field and Network Science Field. Clustering is a data mining technique by grouping data into several clusters so that the data in a cluster has a maximum value. The Fuzzy C-Means (FCM) method is a data clustering method in which the existence of each data point in a cluster is determined by the degree of membership. This study resulted in the optimal weight for Fuzzy C-Means, the Silhouette Coefficient value was 0.532 in the medium category and the purity value was 0.806.

Keywords: *Clustering, Fuzzy C-Means, Informatics*

I. PENDAHULUAN

Program Studi Teknik Informatika merupakan salah satu program studi yang ada di fakultas teknik Universitas Muhammadiyah Gresik. Program Studi Teknik Informatika sudah memiliki bidang keilmuan yang telah ditentukan sesuai dengan mata kuliah pilihan yang sudah ada pada prodi teknik informatika fakultas teknik Universitas Muhammadiyah Gresik. Bidang Keilmuan yang telah ditentukan oleh prodi antara lain adalah Bidang Keilmuan *Artificial Intelegent*, Bidang Keilmuan *Rekayasa Perangkat Lunak* dan Bidang Keilmuan *Jaringan*. Di era industri 4.0 selama ini banyak sekali perusahaan yang membutuhkan tenaga IT agar usahanya semakin berkembang di era keterbukaan teknologi saat ini. Perusahaan yang membuka lowongan pekerjaan saat ini membutuhkan calon karyawan yang fokus dalam salah satu bidang ilmu di dunia IT, seperti contoh nya bidang jaringan , bidang rekayasa perangkat lunak. Mahasiswa Teknik Informatika Universitas Muhammadiyah Gresik selama ini kurang memahami bidang keilmuan yang sesuai dengan dirinya, sehingga selama ini ketika mahasiswa ingin mengambil mata kuliah mereka hanya mengikuti saran dari teman-temannya. Hal tersebut membuat mahasiswa tidak memiliki fokus yang jelas mengenai bidang keilmuan yang akan mereka pelajari di Univesitas Muhammadiyah Gresik. Pemetaan Bidang Keilmuan diperlukan dalam Program Studi Teknik Informatika untuk mengetahui Bidang Keilmuan yang dimiliki oleh mahasiswa, sehingga mahasiswa dapat fokus mempelajari bidang keilmuan yang dimilikinya. Pemetaan bidang keilmuan dapat dilakukan dengan metode clustering.

Clustering merupakan suatu teknik data mining dengan cara mengelompokan data kedalam beberapa cluster sehingga data-data di suatu cluster memiliki kemiripan maksimum. Metode yang bisa digunakan untuk melakukan clustering diantaranya : metode K-Means , metode LVQ (Learning Vector Quantization), FCM (Fuzzy C-Means) dan sebagainya. Teknik clustering sangat berguna dalam prediksi dan analisa terhadap masalah tertentu, misalnya masalah pemetaan bidang keilmuan. Teknik clustering ini bisa memprediksi data mana saja yang mempunyai kemiripan yang sama.

Metode *Fuzzy C- Means*(FCM) yaitu salah satu metode pengclusteran data yang mana keberadaan masing- masing titik data dalam sesuatu cluster ditetapkan oleh derajat keanggotaan. Metode ini awal kali diperkenalkan oleh Jim Bezdek pada tahun 1981(Kusumadewi,2010). Konsep bawah FCM, awal kali merupakan memastikan pusat cluster, yang hendak menandai posisi rata-rata buat masing- masing cluster. Pada keadaan dini, pusat cluster ini masih belum akurat. Masing- masing titik informasi mempunyai derajat keanggotaan buat masing- masing cluster. Cara membetulkan pusat cluster serta derajat keanggotaan masing- masing titik informasi secara kesekian, maka akan bisa dilihat kalau pusat cluster akan bergerak mengarah posisi yang pas. Perulangan ini didasarkan pada minimisasi guna objektif yang menggambarkan jarak dari titik data yang diberikan ke pusat cluster yang terbobot oleh derajat keanggotaan titik informasi tersebut (Kusumadewi,2010).

Maka dari itu penulis melakukan penelitian yang berjudul "*Pemetaan Bidang Keilmuan menggunakan Fuzzy C-means (Studi Kasus : Teknik informatika Universitas Muhammadiyah Gresik)*". Penelitian ini akan menghasilkan Bidang keilmuan mahasiswa yang sesuai dengan kriteria yang ditentukan. Kriteria yang digunakan adalah MKP Kecerdasan Komputasional, MKP Analisis Citra & Visi Komputer, MKP Simulasi dan Optimasi, MKP Sistem Temu Kembali Informasi, MKP Sistem Pendukung Keputusan, MKP Rekayasa Kebutuhan, MKP Management Proyek Perangkat Lunak, MKP Management Kualitas Perangkat Lunak, MKP Sistem Terdistribusi, MKP Jaringan Multimedia, MKP Jaringan Komputasi Bergerak dan MKP Desain

Management Jaringan. Diharapkan pada penelitian ini dapat menentukan bidang keilmuan mahasiswa sehingga mahasiswa dapat fokus mempelajari bidang keilmuan yang dimilikinya.

II. LANDASAN TEORI

2.1 Clustering Data Mining

Istilah lain yang memiliki makna yang sama dengan data mining adalah knowledge discovery in database (KDD). Definisi data mining menurut Erma Delima Sikumbang (2018) adalah proses interaktif untuk menemukan pola atau model baru yang sempurna, bermanfaat dan dapat dimengerti dalam suatu database yang sangat besar (massive database). Sementara Eko Prasetyo (2014) menjelaskan pengertian data mining secara naratif yang mempunyai beberapa maksud sebagai berikut :

1. Pencarian otomatis pola dalam basis data besar, menggunakan teknik komputasional campuran dari statistik, pembelajaran mesin, dan pengenalan pola.
2. Pengekstrakan implisit non-trivial, yang sebelumnya belum diketahui secara potensial adalah informasi berguna dari data.
3. Ilmu pengekstrakan informasi yang berguna dari set data atau basis data besar.
4. Eksplorasi otomatis atau semi otomatis dan analisis data dalam jumlah besar dengan tujuan untuk menemukan pola yang bermakna.
5. Proses penemuan informasi otomatis dengan mengidentifikasi pola dan hubungan ‘tersembunyi’ dalam data.

2.2 Clustering Fuzzy C-Means

Pengelompokan menggunakan metode fuzzy C-Means (FCM) didasarkan pada teori logika fuzzy. Teori ini pertama kali diperkenalkan oleh Lotfi Zadeh (1965) dengan nama Fuzzy Set. Dalam logika fuzzy, keanggotaan data tidak secara eksplisit diberi nilai 1 (menjadi anggota) dan 0 (bukan anggota), tetapi nilai keanggotaan antara 0 dan 1. Akan ditetapkan. Nilai keanggotaan data dalam himpunan adalah 0 jika bukan anggota dan 1 jika merupakan anggota penuh himpunan. Nilai keanggotaan biasanya antara 0 dan 1. Semakin tinggi nilai keanggotaan, semakin tinggi tingkat keanggotaan, dan semakin rendah tingkat keanggotaan, semakin rendah tingkat keanggotaan. Adapun K-Means, FCM sebenarnya adalah versi fuzzy K-Means, dengan beberapa perubahan yang membedakannya dari K-Means.

2.3 Normalisasi Data

Beberapa kasus data mining memiliki fitur dengan nilai yang terletak dalam jangkauan nilai yang berbeda. Akibatnya, fitur dengan nilai atau jangkauan yang besar akan mendominasi dan memiliki pengaruh yang lebih besar dalam fungsi biaya dibandingkan dengan fitur yang memiliki nilai atau jangkauan yang kecil. Normalisasi data dilakukan agar semua fitur akan berada dalam jangkauan yang sama sehingga tidak ada fitur yang mendominasi fungsi biaya pada klasifikator. Cara sederhana dan banyak digunakan adalah normalisasi linier. Berikut adalah rumus perhitungan normalisasi linier :

- a. Hitung nilai mean masing-masing fitur menggunakan persamaan (1)

$$\bar{x}_k = \sum_{i=1}^N x_{ik}, k = 1, 2, \dots, r \quad (1)$$

- b. Hitung nilai varian masing-masing fitur menggunakan persamaan (2)

$$\sigma_k^2 = \frac{1}{N-1} \sum_{i=1}^N (x_{ik} - \bar{x}_k)^2 \quad (2)$$

c. Data hasil normalisasi dihitung menggunakan persamaan (3)

$$\hat{x}_{ik} = \frac{x_{ik} - \bar{x}_k}{\sigma_k} \quad (3)$$

Teknik linier yang lain adalah dengan menskalakan jangkauan setiap fitur dalam jangkauan [0,1] atau [-1,1]. Berikut adalah persamaan yang digunakan.

a. Jangkauan [0,1] :

$$\hat{x}_{ik} = \frac{x_{ik} - \min(x_k)}{\max(x_k) - \min(x_k)}$$

b. Jangkauan [-1,1] :

$$\hat{x}_{ik} = \frac{2x_{ik} - (\max(x_k) + \min(x_k))}{\max(x_k) - \min(x_k)}$$

2.4 Jarak Manhattan

Jarak manhattan merupakan jarak yang diukur mengikuti jalur yang tegak lurus. Disebut dengan jarak manhattan, mengingatkan jalan-jalan manhattan yang membentuk garis paralel dan saling tegak lurus satu jalan dengan jalan lainnya. Pengukuran dengan jarak manhattan sering digunakan karena mudah perhitungannya, mudah dimengerti dan untuk beberapa masalah lebih sesuai, misalnya untuk menentukan jarak antar kota, jarak antar fasilitas dimana peralatan pemindahan bahan hanya dapat bergerak searah tegak lurus (Setya Yaniar Nimas 2012). Pengukuran jarak manhattan digunakan notasi pada persamaan (4) berikut :

$$d_1(x, y) = \|x_1 - y_1\| = \sum_{i=1}^r |x_i - y_i| \quad (4)$$

Keterangan :

d (x,y) = Jarak antara data pada titik x dan titik y

x = Data

y = Pusat *cluster*

2.5 Algoritma Clustering Dengan Fuzzy C-Means (FCM)

Algoritma Fuzzy C-Means adalah sebagai berikut.

a. Input data

Langkah ini dapat dilakukan dengan mengasumsikan sejumlah data dalam set X, berupa matriks berukuran n x m (n = jumlah data, m = jumlah atribut). x_{ij} adalah data sampel ke-i (i = 1,2,3,...,n) dan atribut ke-j (j = 1,2,3,...,m).

b. Tentukan jumlah *cluster*

Tentukan jumlah cluster (K), pangkat untuk matriks partisi (w), maksimum iterasi (MaxIter),

error terkecil yang diharapkan (ξ), fungsi objektif awal ($P_0=0$), dan iterasi awal ($t=1$).

c. Bangkitkan nilai random

Bangkitkan matriks fuzzy pseudo-partition yang diinisialisasikan dengan memberikan nilai sembarang dalam jangkauan $[0,1]$ dan jumlah untuk setiap data (baris) adalah 1. u_{ik} adalah matriks data ke- i ($i = 1,2,3,\dots,n$) dan cluster ke- k ($k=1,2,\dots,k$).

d. Hitung pusat *cluster* ke- k

Hitung pusat cluster ke- k menggunakan persamaan berikut :

$$c_{ij} = \frac{\sum_{i=1}^n (u_{ik})^w \cdot x_{ij}}{\sum_{i=1}^n (u_{ik})^w}$$

Dengan :

c_{ij} = pusat cluster ke- k untuk atribut ke- j

u_{ik} = derajat keanggotaan untuk data ke- i pada cluster ke- k

x_{ij} = data ke- i , atribut ke- j

e. Hitung nilai derajat keanggotaan data pada *cluster*

Hitung nilai derajat keanggotaan data pada cluster menggunakan persamaan berikut:

$$u_{ik} = \frac{D(x_i, c_k)^{\frac{-2}{w-1}}}{\sum_{k=1}^c D(x_i, c_k)^{\frac{-2}{w-1}}}$$

f. Hitung fungsi objektif

Hitung fungsi objektif pada iterasi ke- t menggunakan persamaan berikut :

$$P_t = \sum_{i=1}^N \sum_{k=1}^k (u_{ik})^w D(x_i, c_k)^2$$

Dengan :

J_t = fungsi objektif pada iterasi ke- t

u_{ik} = derajat keanggotaan data ke- i pada cluster ke- k

$D(x_i, c_k)$ = jarak data ke- i dengan pusat cluster ke- k

g. Cek kondisi berhenti

Iterasi akan berhenti perubahan nilai fungsi objektif kurang dari ambang batas yang telah ditentukan atau iterasi telah melebihi maksimal iterasi.

$$(|P_t - P_{t-1}| < \epsilon) \text{ atau } (t > MaxIter)$$

2.6 Silhoutte Coefficient

Pengujian model dilakukan untuk mengetahui seberapa dekat relasi antara objek dalam sebuah cluster dan seberapa jauh sebuah cluster terpisah dengan cluster lain. Metode pengujian yang akan digunakan adalah Silhouette Coefficient. Metode silhouette coefficient merupakan gabungan dari dua metode yaitu metode cohesion yang berfungsi untuk mengukur seberapa dekat relasi antara objek dalam sebuah cluster, dan metode separation yang

berfungsi untuk mengukur seberapa jauh sebuah cluster terpisah dengan cluster lain. Tahapan perhitungan silhouette coefficient :

1. Hitung rata-rata jarak objek dengan semua objek lain yang berada di dalam satu cluster dengan persamaan :

$$a(i) = \frac{1}{[A] - 1} \sum_{j \in A, j \neq i} d(i, j)$$

2. Hitung rata-rata jarak objek dengan semua objek lain yang berada pada cluster lain, kemudian ambil nilai paling minimum dengan persamaan :

$$d(i, C) = \frac{1}{[A]} \sum_{j \in C} d(i, j)$$

3. Hitung nilai silhouette coefficient dengan persamaan :

$$S(i) = \frac{b(i) - a(i)}{\text{Max}\{a(i), b(i)\}}$$

III. ANALISIS DAN PERANCANGAN SISTEM

3.1 Analisis Sistem

Analisis mengenai sistem pemetaan bidang keilmuan mahasiswa Prodi Teknik Informatika Universitas Muhammadiyah Gresik menggunakan metode *fuzzy c-means*. Pengelompokan bidang keilmuan diperlukan dalam Program Studi Teknik Informatika untuk mengetahui Bidang Keilmuan yang dimiliki oleh mahasiswa, sehingga mahasiswa dapat fokus mempelajari bidang keilmuan yang dimilikinya.

Program studi teknik informatika memiliki bidang keilmuan yang telah di tentukan dan diasumsikan pada mata kuliah pilihan yang ada di dalam prodi teknik informatika Universitas Muhammadiyah Gresik. Bidang keilmuan yang telah ditentukan antara lain Bidang Keilmuan *Artificial Intelegent*, Bidang Keilmuan *Rekayasa Perangkat Lunak* dan Bidang Keilmuan *Jaringan*. Selama ini mahasiswa teknik informatika Universitas Muhammadiyah Gresik ketika ingin mengambil mata kuliah mereka hanya mengikuti saran dari teman-temannya sehingga mereka tidak mempunyai fokus yang jelas dalam mendalami bidang keilmuan.

Dalam membuat penelitian ini peneliti mengambil data transkrip nilai mahasiswa dari prodi teknik informatika Universitas Muhammadiyah Gresik untuk diolah dan dikelompokkan sesuai bidang keilmuannya, diharapkan dari hasil penelitan ini program studi teknik informatika Universitas Muhammadiyah Gresik dapat memberikan pembelajaran yang tepat sesuai bidang keilmuan mahasiswa tersebut dan untuk mahasiswa diharapkan bisa lebih fokus dalam mempelajari bidang keilmuan yang diminati. Pemetaan bidang keilmuan dalam penelitian ini menggunakan metode *fuzzy c-means* untuk mendapatkan daftar mahasiswa yang sesuai dengan bidang keilmuan mereka.

3.2 Representasi Data

Dalam pengujian dalam sistem pemetaan bidang keilmuan mahasiswa teknik informatika Universitas Muhammadiyah Gresik menggunakan data 2 periode angkatan yaitu transkrip nilai mahasiswa angkatan 2017 dan mahasiswa angkatan 2018 .

Pada transkrip nilai mahasiswa penelitian ini menggunakan nilai Mata Kuliah Pilihan sebagai atribut, mata kuliah yang dipilih ketika mahasiswa memasuki semester 5 dan semester 6. Mahasiswa angkatan tahun 2017 dan 2018 merupakan angkatan yang sudah mengambil Mata Kuliah Pilihan.

Tabel 3.1 Atribut Data

NO	MATA KULIAH PILIHAN
1	MKP. Kecerdasan Komputasional
2	MKP. Analisis Citra & Visi Komputer
3	MKP. Simulasi dan Optimasi
4	MKP. Sistem Temu Kembali Informasi
5	MKP. Sistem Pendukung Keputusan
6	MKP. Rekayasa Kebutuhan
7	MKP. Management Proyek Perangkat Lunak
8	MKP. Management Kualitas Perangkat Lunak
9	MKP. Sistem Terdistribusi
10	MKP. Jaringan Multimedia
11	MKP. Jaringan Komputasi Bergerak
12	MKP. Desain Manajemen Jaringan

3.3 Perancangan Sistem

Setelah analisis sistem dilakukan maka akan dilanjutkan dengan proses perancangan sistem pemetaan bidang keilmuan. Perancangan ini akan tersusun dalam bentuk *flowchart*, diagram konteks, diagram berjenjang dan *data flow diagram (DFD)*.

3.4 Flowchart

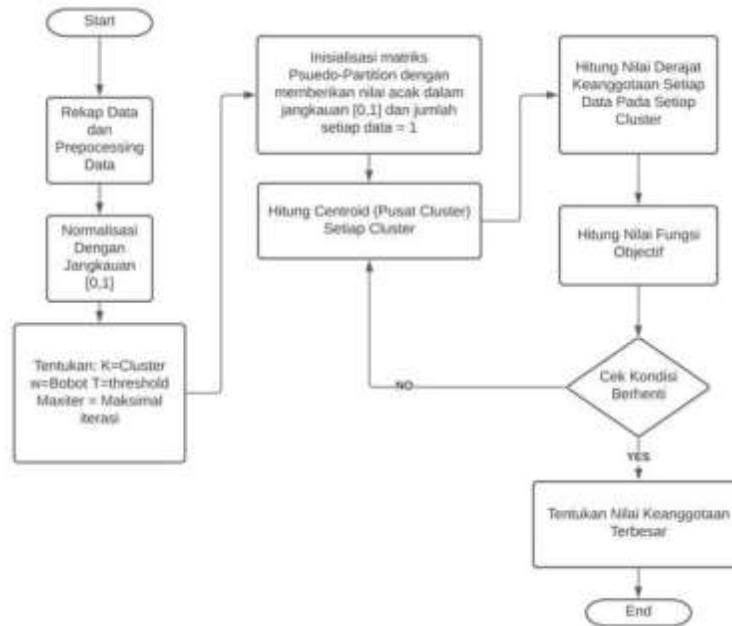


Gambar 3.1 *Flowchart* Sistem Pemetaan Bidang Keilmuan

Berikut ini adalah penjelasan *flowchart* atau diagram yang menggambarkan langkah-langkah dari sistem pemetaan bidang keilmuan mahasiswa:

1. Program Studi Teknik informatika memasukkan Nilai Mata Kuliah Pilihan (MKP) dari mahasiswa yang terdiri dari 12 Mata Kuliah Pilihan.
2. Proses rekap data nilai dan preprocessing data dari nilai Mata Kuliah Pilihan (MKP) dari mahasiswa dari yang semula menggunakan nilai kategorial menjadi nilai numerik.
3. Proses normalisasi data mengubah data yang semula berada di range yang mempunyai jangkauan jarak yang jauh menjadi jangkauan antara 0 sampai 1.
4. Proses perhitungan *clustering* dengan menggunakan metode *Fuzzy C-Means* untuk pemetaan bidang keilmuan mahasiswa. Sebelum melakukan perhitungan, perlu menentukan jumlah *cluster* yang diinginkan. Proses perhitungan ini juga meliputi proses validitas *clustering* Sihoulet Conde.
5. Melaporkan hasil pemetaan bidang keilmuan mahasiswa dengan metode *Fuzzy C-Means*. Hasil pemetaan merupakan informasi tentang kelompok-kelompok mahasiswa sesuai dengan bidang keilmuan yang cocok dengan mahasiswa tersebut.

3.5 Flowchart Sistem Menggunakan Metode Fuzzy C- Means



Gambar 3.2 *Flowchart* Sistem menggunakan Fuzzy C-means

Pada gambar 3.2 merupakan alur dari algoritma Fuzzy C-means:

1. Masukkan data rekap Nilai Mata Kuliah Pilihan Mahasiswa yang terdiri dari 12 mata kuliah dan melakukan preprocessing data yang semula kategorial menjadi numeric.
2. Lakukan normalisasi data dengan jangkauan [0,1].
3. Tentukan jumlah *cluster* yang diinginkan (K), pembobot (w), ambang batas atau *threshold* (T), dan maksimal iterasi.
4. Inisialisasikan matriks *pseudo-partition* dengan memberikan nilai sembarang dalam jangkauan [0,1] dan untuk jumlah nilai setiap data (baris) = 1.

5. Hitung centroid atau pusat *cluster* setiap *cluster*.
6. Hitung nilai derajat keanggotaan setiap data pada setiap *cluster*.
7. Hitung nilai fungsi objektif.
8. Cek kondisi berhenti. Apabila nilai perubahan fungsi objektif telah mencapai di bawah ambang batas atau jumlah iterasi telah melebihi maksimal iterasi, maka iterasi dihentikan. Iterasi akan kembali dilakukan apabila kondisi berhenti belum tercapai. Proses iterasi akan dimulai kembali dari langkah perhitungan centroid (langkah 5).
9. Apabila kondisi berhenti telah tercapai, tentukan nilai derajat keanggotaan terbesar untuk mengetahui *cluster* mana yang diikuti oleh data.

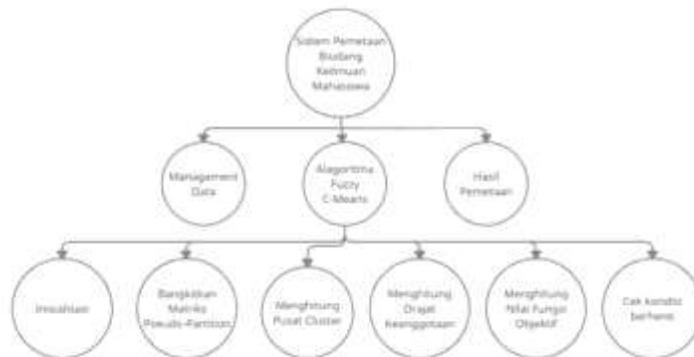
3.6 Diagram Konteks



Gambar 3.3 Diagram Konteks Sistem Pemetaan Bidang Keilmuan

Kepala Program Studi merupakan *entity* dalam sistem pemetaan bidang keilmuan mahasiswa. Kepala Program Studi akan menginputkan data nilai Mata Kuliah Pilihan (MKP) mahasiswa, data nilai Mata Kuliah Pilihan (MKP) digunakan sebagai data yang akan di hitung untuk menentukan bidang keilmuan yang sesuai dengan mahasiswa tersebut. Sebelum dilakukan proses perhitungan menggunakan Metode C-means data nilai Mata Kuliah Pilihan terlebih dahulu dilakukan proses preprocessing data untuk mengubah yang semula berbentuk kategorial menjadi numerik. Nilai Mata Kuliah Pilihan yang berbentuk numerik akan menjadi dasar dalam melakukan proses perhitungan dalam pemetaan bidang keilmuan mahasiswa menggunakan Metode C-means. Setelah proses pemetaan selesai maka kepala program studi bisa melihat kelompok-kelompok mahasiswa yang sesuai di bidang keilmuannya.

3.7 Diagram Berjenjang



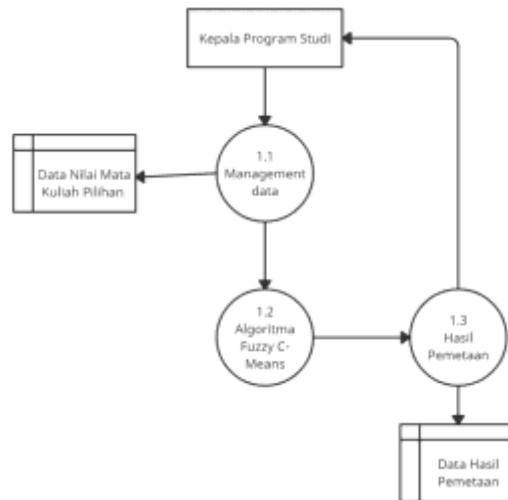
Gambar 3.4 Diagram Berjenjang Sistem Pemetaan Bidang Keilmuan

Dari gambar 3.4 dapat dilihat secara keseluruhan proses yang nantinya dilakukan pada sistem pemetaan bidang keilmuan menggunakan *Fuzzy C-Means* (FCM). Penjelasan dari gambar 3.3:

1. Top level: Aplikasi menggunakan konsep Data Mining yaitu Pemetaan Bidang Keilmuan menggunakan *Fuzzy C-Means* (FCM).
2. Level 1 proses: Berisi proses dalam sistem yang meliputi manajemen data, Algoritma *Fuzzy C-Means*, dan Hasil Pemetaan. Manajemen data merupakan suatu proses memasukkan data nilai Mata Kuliah Pilihan Mahasiswa . Data tersebut yang nantinya akan digunakan untuk melakukan proses pemetaan bidang keilmuan mahasiswa.
3. Level 2: Merupakan proses pemetaan dengan menggunakan FCM yang memuat perhitungan atau tahapan-tahapan dalam menggunakan algoritma *Fuzzy C-Means* (FCM).

3.8 Data Flow Diagram (DFD)

3.8.1 DFD Level 1

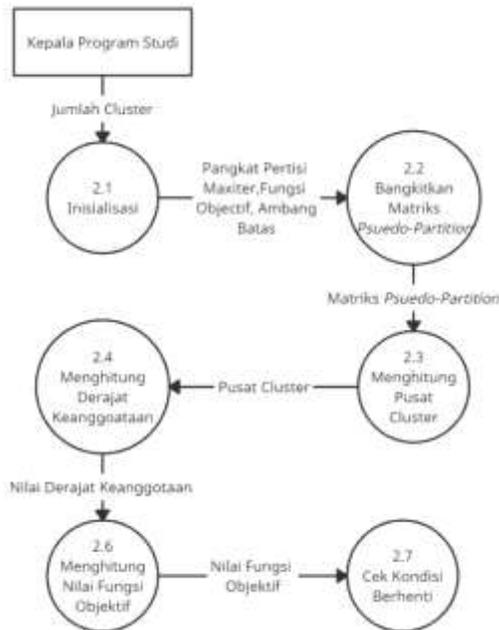


Gambar 3.5 *Data Flow Diagram Level 1* Sistem Pemetaan Bidang Keilmuan

Dapat Gambar 3.5 yang merupakan Data Flow Diagram Level 1 terdapat beberapa proses antara lain:

1. Kepala Program Studi memasukkan data Nilai Mata Kuliah Pilihan Mahasiswa ke dalam sistem database.
2. Data Nilai Mata Kuliah Pilihan di lakukan perhitungan menggunakan Metode C-means.
3. Hasil akan di simpan kedalam database dan akan dilaporkan kepada kepala program studi.

3.8.2 DFD Level 2



Gambar 3.6 Data Flow Diagram Level 2 Sistem Pemetaan Bidang Keilmuan

Dapat Gambar 3.6 yang merupakan Data Flow Diagram Level 2 terdapat beberapa proses antara lain:

1. Kepala Program Studi memasukan jumlah *cluster* yang ingin dihitung.
2. Proses 2.1 menentukan inisialisasi pangkat partisi (w), maksimal iterasi (MaxIter), fungsi objektif awal (P_0) dan ambang batas.
3. Proses 2.2 membangkitkan matriks *pseudo-partition*.
4. Proses 2.3 menghitung pusat *cluster* atau centroid setiap *cluster* menggunakan data transaksi pelanggan dan matriks *pseudo-partition*.
5. Proses 2.4 menghitung derajat keanggotaan. Nilai derajat keanggotaan ini juga akan digunakan sebagai matriks untuk iterasi berikutnya.
6. Proses 2.5 menghitung nilai fungsi objektif.
7. Proses 2.6 mengecek kondisi berhenti. Kondisi berhenti dapat dicapai apabila nilai perubahan fungsi objektif ada di bawah ambang batas atau jumlah iterasi telah melebihi maksimal iterasi.

3.9 Perancangan Dan Pengujian

Perancangan pengujian ini akan menggunakan sejumlah data nilai Mata Kuliah Pilihan(MKP) mahasiswa dimana jumlah data yang digunakan sebanyak 60 data. Dimana data tersebut diambil dari KRS mahasiswa angkatan 2017-2018 Berikut ini adalah data transkrip nilai mahasiswa yang akan ditampilkan pada Tabel 3.7.

Tabel 3.7 Data Transkrip Mahasiswa

No	Nama	Nim	A	B	C	D	E	F	G	H	I	J	K	L
1	Denas Subito abron	170902004	BC	NON	NON	B	AB	AB	NON	AB	C	NON	NON	NON
2	Wahid ghuhy	170902005	BC	NON	A	AB	A	NON	NON	A	A	NON	NON	NON
3	Edu satrio	170902007	B	AB	AB	B	AB	NON	NON	NON	E	NON	NON	NON
4	Ryan Haris Banafi	170902008	B	NON	AB	B	AB	NON	NON	A	A	NON	NON	NON
5	Achmad masadi	170902010	B	C	A	AB	AB	NON	NON	NON	A	NON	NON	NON
50	Ahmad Subaini Rahman	190602036	AB	NON	AB	AB	AB	AB	NON	AB	NON	NON	NON	NON
51	Ahmad Fathul Akbar	190602040	D	NON	BC	AB	AB	B	NON	AB	NON	NON	NON	NON
52	Muhammad Adhyan Firdaus	190602042	A	NON	AB	A	AB	NON	NON	A	A	NON	NON	NON
53	Moch Nasrudin	190602044	AB	NON	A	AB	AB	A	NON	AB	NON	NON	NON	NON
54	Moch. Rivaldy Syarifata	190602045	BC	NON	AB	A	AB	AB	NON	AB	NON	NON	NON	NON
55	Muhammad Iqbal Walipadli	190602046	A	NON	AB	A	AB	NON	NON	A	A	NON	NON	NON
56	Diyah Utami	190602048	A	NON	A	AB	AB	A	NON	AB	NON	NON	NON	NON
57	Muhammad Wahyu Syarif Firdaus	190602049	A	NON	AB	AB	AB	AB	NON	AB	NON	NON	NON	NON
58	Rachmad Deo Abdika	190602053	A	NON	A	A	A	A	NON	AB	NON	NON	NON	NON
59	Gilang Asala Parhared	190602056	A	NON	A	A	A	A	NON	AB	NON	NON	NON	NON
60	Mochamad Ihsan	190602057	A	NON	AB	A	A	A	NON	AB	NON	NON	NON	NON

Keterangan :

- A = MKP. Kecerdasan Komputasional
- B = MKP. Analisis Citra & Visi Komputer
- C = MKP. Simulasi dan Optimasi
- D = MKP. Sistem Temu Kembali Informasi
- E = MKP. Sistem Pendukung Keputusan
- F = MKP. Rekayasa Kebutuhan
- G = MKP. Management Proyek Perangkat Lunak
- H = MKP. Management Kualitas Perangkat Lunak
- I = MKP. Sistem Terdistribusi
- J = MKP Jaringan Multimedia
- K = MKP Jaringan Komputasi Bergerak
- L = MKP. Desain Management Jaringan

Data yang digunakan dalam perhitungan proses pemetaan bidang keilmuan dengan menggunakan metode *Fuzzy C-Means* adalah data yang terdapat pada tabel 3.7. Sebelum dilakukan Pehitungan data terlebih dahulu akan di lakukan Preprocessing data :

3.9.1 Proses Preprocessing Data

Preproccasing data di lakukan untuk mengubah Nilai yang ada di dalam data nilai yang semula berupa kategorial lalu akan di rubah menjadi numerik untuk memudahkan perhitungan :

Tabel 3.8 Preprocessing Data

Nilai Kategorial MKP	Nilai Numerik MKP
A	4
AB	3,5
B	3
BC	2,5
C	2
D	1,5
E	1
NON	0

Tabel 3.9 Nilai Numerik Mata Kuliah Pilihan

No.	Nama	Nim.	A	B	C	D	E	F	G	H	I	J	K	L
1	Dimas habib ahro	170602004	2,5	0	0	3	3,5	3,5	0	3,5	2	0	0	0
2	Wahed ghaly	170602005	2,5	0	4	3,5	4	0	0	4	4	0	0	0
3	Edhi safir	170602007	3	3,5	3,5	3	3,5	0	0	0	2	0	0	0
4	Ryan hani bawati	170602008	3	0	3,5	3	3,5	0	0	4	4	0	0	0
5	Achmad nazari	170602010	3	2	4	3,5	3,5	0	0	0	4	0	0	0
6	Moh malik fajar	170602011	2	3	3,5	3,5	4	0	0	0	4	0	0	0
7	Ervin setyanan	170602012	4	0	4	3,5	4	0	0	4	4	0	0	0
8	Abdurrobbil anurillah	170602013	3	3	4	4	4	0	0	0	4	0	0	0
9	Sri mulyani	170602016	3	0	0	3	0	4	0	0	4	3,5	3,5	0
10	M. Fahrudin	170602018	2,5	2,5	3,5	3,5	3	0	0	0	4	0	0	0
50	Akmat wahari rahman	180602036	3,5	0	3,5	3,5	3,5	3,5	0	3,5	0	0	0	0
51	Alfa nuril fatih albar Muhammad adyan	180602040	1,5	0	2,5	3,5	3,5	3	0	3,5	0	0	0	0
52	Fidans	180602042	4	0	3,5	4	3,5	0	0	4	4	0	0	0
53	Nadhif nasrudin	180602044	3,5	0	4	3,5	3,5	4	0	3,5	0	0	0	0
54	Moch. Rizvaldy byantara	180602045	2,5	0	3,5	4	3,5	3,5	0	3,5	0	0	0	0
55	Muhammad iqbal walyyuddin	180602046	4	0	3,5	4	3,5	0	0	4	4	0	0	0
56	Diyah utami	180602048	4	0	4	3,5	3,5	4	0	3,5	0	0	0	0
57	Muhammad wahyu fajar fidans	180602049	4	0	3,5	3,5	3,5	3,5	0	3,5	0	0	0	0
58	Rachmad deo abdika	180602053	4	0	4	4	4	4	0	3,5	0	0	0	0
59	Gleng stala panharso	180602056	4	0	4	4	4	4	0	3,5	0	0	0	0
60	Mochammad ichsan	180602057	4	0	3,5	4	4	4	0	3,5	0	0	0	0

Setelah dilakukan Preprocessing data dilakukan perhitungan *Fuzzy C-Means* menggunakan 60 data. Langkah-langkah perhitungan *clustering* sesuai dengan alur sistem adalah sebagai berikut.

1. Memasukkan data 60 data nilai MKP serta menentukan nilai maksimum dan minimum data set sejumlah. Nilai tersebut akan digunakan untuk proses normalisasi data pada langkah berikutnya.

Tabel 3.10 Data Perhitungan FCM

No.	Nama	Nim.	A	B	C	D	E	F	G	H	I	J	K	L
1	Dimas habib ahro	170602004	2,5	0	0	3	3,5	3,5	0	3,5	2	0	0	0
2	Wahed ghaly	170602005	2,5	0	4	3,5	4	0	0	4	4	0	0	0
3	Edhi safir	170602007	3	3,5	3,5	3	3,5	0	0	0	2	0	0	0
4	Ryan hani bawati	170602008	3	0	3,5	3	3,5	0	0	4	4	0	0	0
5	Achmad nazari	170602010	3	2	4	3,5	3,5	0	0	0	4	0	0	0
6	Moh malik fajar	170602011	2	3	3,5	3,5	4	0	0	0	4	0	0	0
7	Ervin setyanan	170602012	4	0	4	3,5	4	0	0	4	4	0	0	0
8	Abdurrobbil anurillah	170602013	3	3	4	4	4	0	0	0	4	0	0	0
9	Sri mulyani	170602016	3	0	0	3	0	4	0	0	4	3,5	3,5	0
10	M. Fahrudin	170602018	2,5	2,5	3,5	3,5	3	0	0	0	4	0	0	0
50	Akmat wahari rahman	180602036	3,5	0	3,5	3,5	3,5	3,5	0	3,5	0	0	0	0
51	Alfa nuril fatih albar Muhammad adyan	180602040	1,5	0	2,5	3,5	3,5	3	0	3,5	0	0	0	0
52	Fidans	180602042	4	0	3,5	4	3,5	0	0	4	4	0	0	0
53	Nadhif nasrudin	180602044	3,5	0	4	3,5	3,5	4	0	3,5	0	0	0	0
54	Moch. Rizvaldy byantara	180602045	2,5	0	3,5	4	3,5	3,5	0	3,5	0	0	0	0
55	Muhammad iqbal walyyuddin	180602046	4	0	3,5	4	3,5	0	0	4	4	0	0	0
56	Diyah utami	180602048	4	0	4	3,5	3,5	4	0	3,5	0	0	0	0
57	Muhammad wahyu fajar fidans	180602049	4	0	3,5	3,5	3,5	3,5	0	3,5	0	0	0	0
58	Rachmad deo abdika	180602053	4	0	4	4	4	4	0	3,5	0	0	0	0
59	Gleng stala panharso	180602056	4	0	4	4	4	4	0	3,5	0	0	0	0
60	Mochammad ichsan	180602057	4	0	3,5	4	4	4	0	3,5	0	0	0	0

pembaca:														
60	Mochammadichsan	180602057	4	0	3,5	4	4	4	0	3,5	0	0	0	0
	MIN		0	0	0	0	0	0	0	0	0	0	0	
	MAX		4	3,5	4	4	4	4	3,5	4	4	3,5	3,5	

2. Proses normalisasi. Normalisasi yang digunakan adalah normalisasi data dengan jangkauan [0,1]

$$x_i = \frac{Xi - MIN}{MAX - MN}$$

$$x_{11} = \frac{2,5 - 0}{4 - 0} = \frac{2,5}{4} = 0,625$$

$$x_{12} = \frac{0 - 0}{3,5 - 0} = \frac{0}{3,5} = 0$$

$$x_{13} = \frac{0 - 0}{4 - 0} = \frac{0}{4} = 0$$

$$x_{14} = \frac{3 - 0}{4 - 0} = \frac{3}{4} = 0,75$$

$$x_{15} = \frac{3,5 - 0}{4 - 0} = \frac{3,5}{4} = 0,875$$

$$x_{16} = \frac{3,5 - 0}{4 - 0} = \frac{3,5}{4} = 0,875$$

$$x_{17} = \frac{0 - 0}{3,5 - 0} = \frac{0}{3,5} = 0$$

$$x_{18} = \frac{3,5 - 0}{4 - 0} = \frac{3,5}{4} = 0,875$$

$$x_{19} = \frac{2 - 0}{4 - 0} = \frac{2}{4} = 0,5$$

$$x_{110} = \frac{0 - 0}{3,5 - 0} = \frac{0}{3,5} = 0$$

$$x_{111} = \frac{0 - 0}{3,5 - 0} = \frac{0}{3,5} = 0$$

$$x_{112} = \frac{0 - 0}{4 - 0} = \frac{0}{4} = 0$$

Hasil proses normalisasi adalah sebagai berikut.

Tabel 3.11 Hasil Normalisasi Data

No	Nama	Nilai	A	B	C	D	E	F	G	H	I	J	K	L
1	Dimas kabdo abour	170602004	0,625	0	0	0,75	0,875	0,875	0	0,875	0,5	0	0	0
2	Wahid ghufr	170602005	0,625	0	1	0,875	1	0	0	1	1	0	0	0
3	Edu safir	170602007	0,75	1	0,875	0,75	0,875	0	0	0	0,5	0	0	0
4	Ryan hana harafi	170602008	0,75	0	0,875	0,75	0,875	0	0	1	1	0	0	0
5	Achmad amali	170602010	0,75	0,571429	1	0,875	0,875	0	0	0	1	0	0	0
6	Moh malik fajar	170602011	0,5	0,857143	0,875	0,875	1	0	0	0	1	0	0	0
7	Ervin setyan	170602012	1	0	1	0,875	1	0	0	1	1	0	0	0
8	Abdunoyid amriyah	170602013	0,75	0,857143	1	1	1	0	0	1	0	0	0	0
9	Semmlyan	170602015	0,75	0	0	0,75	0	1	0	0	1	1	1	0
10	M Fahrudin	170602018	0,625	0,714286	0,875	0,875	0,75	0	0	1	0	0	0	0
...
50	Ahmad subhan rahman	180602036	0,875	0	0,875	0,875	0,875	0,875	0	0,875	0	0	0	0
51	Alfie umil farid alfar	180602040	0,375	0	0,875	0,875	0,875	0,75	0	0,875	0	0	0	0
52	Muhammad adnan Sadana	180602042	1	0	0,875	1	0,875	0	0	1	1	0	0	0
53	Nadhif nasudin	180602044	0,875	0	1	0,875	0,875	1	0	0,875	0	0	0	0
54	Moch. Rhuaidy	180602045	0,625	0	0,875	1	0,875	0,875	0	0,875	0	0	0	0

No	Proses	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
55	Mahamanal ipah	1	0	0,875	0	0,875	0	0	0	0	0	0	0	0	0	0
56	Mahamanal Diyah utami	1	0	1	0,875	0,875	0	0	0,875	0	0	0	0	0	0	0
57	Mahamanal Makrya Agw Fitria	1	0	0,875	0,875	0,875	0,875	0	0,875	0	0	0	0	0	0	0
58	Karomel dan abdiha	1	0	1	1	1	1	0	0,875	0	0	0	0	0	0	0
59	Changana pachani	1	0	1	1	1	1	0	0,875	0	0	0	0	0	0	0
60	Sorkhamat ichan	1	0	0,875	0	1	1	0	0,875	0	0	0	0	0	0	0

Langkah-langkah berikutnya adalah langkah perhitungan *Fuzzy C-Means* untuk pemetaan bidang keilmuan.

- Inisialisasi. Jumlah *cluster* (*c*) yang dicari adalah 3, pangkat untuk matriks partisi (*w*) adalah 2, nilai fungsi objektif awal (*Po*) yang digunakan adalah 10000, maksimum iterasi (*maxIter*) untuk mengecek kondisi berhenti adalah 100 iterasi dan ambang batas yang digunakan adalah 0,1.
- Membangkitkan matriks *pseudo-partition* dengan memberikan nilai sembarang dalam jangkauan [0,1] untuk data ke-*i* (*i* = 1,2,3,...,n) dan *cluster* ke-*k* (*k*=1,2,...,k). Jumlah untuk setiap data (baris) adalah 1. Matriks *pseudo-partition* yang dibangkitkan ditampilkan pada tabel dibawah ini.

Tabel 3.12 Matriks *pseudo-partition*

No	u1	u2	u3
1	0,237	0,2689	0,4941
2	0,2748	0,2854	0,4398
3	0,3298	0,3015	0,3687
4	0,1923	0,288	0,5197
5	0,2151	0,1683	0,6166
6	0,1882	0,2249	0,5869
7	0,2763	0,2416	0,4821
8	0,2414	0,3075	0,4511
9	0,3034	0,2277	0,4689
10	0,2526	0,207	0,5404
...
50	0,3265	0,3044	0,3691
51	0,2442	0,2694	0,4864
52	0,2107	0,3258	0,4635
53	0,2846	0,2615	0,4539
54	0,2953	0,2607	0,444
55	0,3277	0,1797	0,4926
56	0,2102	0,1787	0,6111
57	0,2835	0,3328	0,3837
58	0,2553	0,2198	0,5249
59	0,2863	0,3014	0,4123
60	0,1902	0,2406	0,5692

- Hitung centroid setiap *cluster* menggunakan persamaan 3.7. Berikut ini adalah contoh perhitungan centroid atau pusat *cluster*.

52	6	0,106	0,000	0,093	0,106	0,093	0,000	0,000	0,106	0,106	0,000	0,000	0,000
53	8	0,060	0,000	0,068	0,060	0,060	0,068	0,000	0,060	0,000	0,000	0,000	0,000
54	8	0,042	0,000	0,058	0,048	0,058	0,058	0,000	0,042	0,000	0,000	0,000	0,000
55	2	0,032	0,000	0,028	0,032	0,028	0,000	0,000	0,032	0,032	0,000	0,000	0,000
56	2	0,032	0,000	0,032	0,028	0,028	0,032	0,000	0,028	0,000	0,000	0,000	0,000
57	1	0,111	0,000	0,097	0,097	0,097	0,097	0,000	0,097	0,000	0,000	0,000	0,000
58	8	0,048	0,000	0,048	0,048	0,048	0,048	0,000	0,042	0,000	0,000	0,000	0,000
59	1	0,091	0,000	0,091	0,091	0,091	0,091	0,000	0,079	0,000	0,000	0,000	0,000
60	8	0,058	0,000	0,051	0,058	0,058	0,058	0,000	0,051	0,000	0,000	0,000	0,000
61	5	3,072	0,604	3,068	3,411	3,408	3,888	0,172	2,766	2,338	0,188	0,224	0,091

Tabel 3.15 Perhitungan centroid untuk cluster 3 iterasi ke-1 (lanjutan)

N	U _{ik} ^w	U _{ik} ^w x _{i1}	U _{ik} ^w x _{i2}	U _{ik} ^w x _{i3}	U _{ik} ^w x _{i4}	U _{ik} ^w x _{i5}	U _{ik} ^w x _{i6}	U _{ik} ^w x _{i7}	U _{ik} ^w x _{i8}	U _{ik} ^w x _{i9}	U _{ik} ^w x _{i10}	U _{ik} ^w x _{i11}	U _{ik} ^w x _{i12}
1	0,344	0,153	0,000	0,000	0,183	0,214	0,214	0,000	0,214	0,122	0,000	0,000	0,000
2	0,193	0,121	0,000	0,193	0,169	0,193	0,000	0,000	0,193	0,193	0,000	0,000	0,000
3	0,158	0,102	0,158	0,119	0,122	0,119	0,000	0,000	0,000	0,000	0,000	0,000	0,000
4	0,210	0,203	0,000	0,216	0,203	0,206	0,000	0,000	0,210	0,210	0,000	0,000	0,000
5	0,380	0,285	0,217	0,380	0,333	0,333	0,000	0,000	0,000	0,380	0,000	0,000	0,000
6	0,344	0,172	0,285	0,301	0,301	0,344	0,000	0,000	0,000	0,344	0,000	0,000	0,000
7	0,233	0,233	0,000	0,233	0,233	0,233	0,000	0,000	0,233	0,233	0,000	0,000	0,000
8	0,203	0,153	0,174	0,203	0,203	0,203	0,000	0,000	0,000	0,203	0,000	0,000	0,000
9	0,220	0,183	0,000	0,000	0,183	0,000	0,220	0,000	0,220	0,220	0,000	0,000	0,000
10	0,292	0,183	0,289	0,256	0,256	0,218	0,000	0,000	0,000	0,292	0,000	0,000	0,000
11	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
12	0,138	0,119	0,000	0,119	0,119	0,119	0,000	0,119	0,000	0,000	0,000	0,000	0,000
13	0,257	0,088	0,000	0,148	0,207	0,207	0,177	0,000	0,207	0,000	0,000	0,000	0,000
14	0,215	0,215	0,000	0,188	0,215	0,188	0,000	0,000	0,215	0,215	0,000	0,000	0,000
15	0,204	0,180	0,000	0,206	0,180	0,180	0,206	0,000	0,180	0,000	0,000	0,000	0,000
16	0,197	0,123	0,000	0,172	0,197	0,172	0,172	0,000	0,172	0,000	0,000	0,000	0,000
17	0,243	0,243	0,000	0,212	0,243	0,212	0,000	0,000	0,243	0,243	0,000	0,000	0,000
18	0,373	0,373	0,000	0,373	0,327	0,327	0,373	0,000	0,327	0,000	0,000	0,000	0,000
19	0,147	0,147	0,000	0,129	0,129	0,129	0,129	0,000	0,129	0,000	0,000	0,000	0,000
20	0,278	0,278	0,000	0,278	0,278	0,278	0,278	0,000	0,241	0,000	0,000	0,000	0,000
21	0,170	0,170	0,000	0,170	0,170	0,170	0,170	0,000	0,149	0,000	0,000	0,000	0,000
22	0,324	0,324	0,000	0,283	0,324	0,324	0,324	0,000	0,283	0,000	0,000	0,000	0,000
23	14,651	11,151	2,800	11,475	12,417	12,782	6,532	0,333	9,528	8,864	0,537	0,388	0,219

Pusat cluster didapatkan dengan menggunakan persamaan 3.7 yaitu dari pembagian penjumlahan $u_{ik}^w \cdot x_{ij}$ dengan penjumlahan u_{ik}^w sebagai berikut.

$$c_i = \frac{\sum u_{ik}^w \cdot x_{ij}}{u_{ik}^w} \quad (3.2)$$

$$c_{11} = \frac{3,149}{4,036} = 0,780$$

$$c_{12} = \frac{0,613}{4,036} = 0,152$$

$$c_{13} = \frac{3,185}{4,036} = 0,789$$

$$c_{14} = \frac{3,490}{4,036} = 0,865$$

$$c_{15} = \frac{3,469}{4,036} = 0,860$$

$$c_{16} = \frac{1,788}{4,036} = 0,443$$

$$c_{17} = \frac{0,093}{4,036} = 0,023$$

$$c_{18} = \frac{2,835}{4,036} = 0,702$$

$$c_{19} = \frac{2,306}{4,036} = 0,571$$

$$c_{110} = \frac{0,188}{4,036} = 0,047$$

$$c_{11} = \frac{0,109}{4,036} = 0,052$$

$$c_{12} = \frac{0,051}{4,036} = 0,013$$

Hasil centroid masing-masing atribut untuk setiap *cluster* yang didapatkan adalah sebagai berikut.

Tabel 3.16 Centroid iterasi ke-1

No	A	B	C	D	E	F	G	H	I	J	K	L
1	0,78 0	0,15 2	0,7 89	0,8 65	0,86 0	0,4 43	0,0 23	0,7 2	0,57 1	0,047	0,052	0,0 13
2	0,75 8	0,14 9	0,7 57	0,8 41	0,84 0	0,4 66	0,0 42	0,6 2	0,57 6	0,049	0,055	0,0 23
3	0,76 1	0,19 1	0,7 83	0,8 48	0,87 2	0,4 46	0,0 24	0,6 0	0,60 5	0,037	0,040	0,0 15

6. Lakukan perhitungan nilai derajat keanggotaan data pada *cluster* menggunakan persamaan Berikut adalah contoh perhitungan jarak data dengan pusat *cluster* menggunakan perhitungan jarak Manhattan dan perhitungan nilai derajat keanggotaan.

$$D(x_1, c_1) = |0,625 - 0,780| + |0 - 0,015| + |0 - 0,789| + |0,75 - 0,865| + |0,875 - 0,860| + |0,875 - 0,443| + |0 - 0,023| + |0,875 - 0,702| + |0,5 - 0,571| + |0 - 0,047| + |0 - 0,052| + |0 - 0,013| = \mathbf{0,241}$$

$$D(x_1, c_2) = |0,625 - 0,758| + |0 - 0,149| + |0 - 0,757| + |0,75 - 0,841| + |0,875 - 0,040| + |0,875 - 0,466| + |0 - 0,042| + |0,875 - 0,682| + |0,5 - 0,576| + |0 - 0,049| + |0 - 0,055| + |0 - 0,023| = \mathbf{0,247}$$

$$D(x_1, c_3) = |0,625 - 0,761| + |0 - 0,191| + |0 - 0,783| + |0,75 - 0,484| + |0,875 - 0,872| + |0,875 - 0,446| + |0 - 0,024| + |0,875 - 0,650| + |0,5 - 0,605| + |0 - 0,037| + |0 - 0,040| + |0 - 0,015| = \mathbf{0,230}$$

$$u_{11} = \frac{D(x_1, c_1)^{-2}}{\sum_{k=1}^c D(x_i, c_k)^{-2}} = \frac{0,241^{-2}}{0,241^{-2} + 0,247^{-2} + 0,230^{-2}} = 0,336$$

$$u_{12} = \frac{D(x_1, c_2)^{-2}}{\sum_{k=1}^c D(x_i, c_k)^{-2}} = \frac{0,247^{-2}}{0,241^{-2} + 0,247^{-2} + 0,230^{-2}} = 0,344$$

$$u_{13} = \frac{D(x_1, c_3)^{-2}}{\sum_{k=1}^c D(x_i, c_k)^{-2}} = \frac{0,230^{-2}}{0,241^{-2} + 0,247^{-2} + 0,230^{-2}} = 0,320$$

Hasil perhitungan nilai derajat keanggotaan ditampilkan sebagai berikut.

Tabel 3.17 Nilai derajat keanggotaan iterasi ke-1

No	Jarak ke centroid			Derajat keanggotaan		
	D(xi,c1)	D(xi,c2)	D(xi,c3)	ui1	ui2	ui3
1	0,241	0,247	0,230	0,336	0,344	0,320
2	0,257	0,228	0,249	0,350	0,311	0,339
3	0,167	0,160	0,184	0,327	0,313	0,360
4	0,345	0,317	0,346	0,343	0,314	0,343
5	0,174	0,162	0,198	0,326	0,303	0,371
6	0,116	0,110	0,129	0,328	0,309	0,364
7	0,241	0,206	0,225	0,359	0,306	0,335
8	0,116	0,110	0,129	0,328	0,309	0,364
9	0,032	0,034	0,033	0,326	0,341	0,333
10	0,144	0,139	0,159	0,326	0,314	0,359
...
50	0,359	0,308	0,308	0,368	0,316	0,316
51	0,268	0,262	0,246	0,345	0,337	0,317
52	0,274	0,231	0,254	0,360	0,305	0,335
53	0,272	0,238	0,238	0,364	0,318	0,318
54	0,291	0,265	0,263	0,355	0,324	0,321
55	0,274	0,231	0,254	0,360	0,305	0,335
56	0,239	0,211	0,211	0,362	0,319	0,319
57	0,311	0,269	0,269	0,366	0,317	0,317
58	0,190	0,170	0,170	0,359	0,321	0,321
59	0,190	0,170	0,170	0,359	0,321	0,321
60	0,213	0,189	0,189	0,360	0,320	0,320

7. Hitung nilai fungsi objektif menggunakan persamaan 3.7. Berikut ini adalah contoh perhitungan nilai fungsi objektif sesuai dengan persamaan.

$$\begin{aligned}
 (u_{11})^w \cdot D(x_1, c_1)^2 &= (0,336)^2 \cdot (0,241)^2 = 0,468 \\
 (u_{12})^w \cdot D(x_1, c_2)^2 &= (0,344)^2 \cdot (0,247)^2 = 0,479 \\
 (u_{13})^w \cdot D(x_1, c_3)^2 &= (0,320)^2 \cdot (0,230)^2 = 0,446
 \end{aligned}$$

Hasil perhitungan adalah sebagai berikut.

Tabel 3.18 Nilai fungsi objektif iterasi ke-1

No	$(u_{11})^w.D(x_1,c_1)^2$	$(u_{12})^w.D(x_1,c_2)^2$	$(u_{13})^w.D(x_1,c_3)^2$
1	0,468	0,479	0,446
2	0,478	0,423	0,462
3	0,638	0,612	0,703
4	0,340	0,312	0,340
5	0,609	0,567	0,693
6	0,921	0,868	1,023
7	0,534	0,456	0,499
8	0,921	0,868	1,023
9	3,301	3,455	3,366
10	0,738	0,710	0,812
...
50	0,377	0,324	0,324
51	0,445	0,435	0,409
52	0,475	0,401	0,441
53	0,487	0,426	0,426
54	0,433	0,395	0,392
55	0,475	0,401	0,441
56	0,547	0,482	0,482
57	0,430	0,373	0,373
58	0,677	0,605	0,605
59	0,677	0,605	0,605
60	0,610	0,542	0,542

8. Lakukan pengecekan kondisi berhenti dengan membandingkan nilai perubahan fungsi objektif ($|P_t - P_{t-1}|$). Perubahan nilai fungsi objektif dilakukan dengan melakukan pengurangan nilai fungsi objektif awal atau iterasi sebelumnya dengan nilai fungsi objektif yang baru. Pada iterasi 1 ini, dikarenakan data belum masuk dalam *cluster*, maka nilai fungsi objektif awal adalah nilai yang telah ditentukan sebelumnya, yaitu 10000. Nilai fungsi objektif yang didapat pada iterasi ke-1 adalah jumlah seluruh data pada tabel 3.9 diatas sesuai dengan persamaan 2.9, yaitu 135,609
 Perubahan nilai fungsi objektif = $10000 - 135,609 = 9864,391$
9. Karena perubahan nilai fungsi objektif masih ambang batas, maka proses dilanjutkan ke iterasi berikutnya, yaitu dimulai dengan menghitung kembali pusat *cluster*. Hasil perhitungan fungsi objektif dan perubahan nilai fungsi objektif setiap iterasi yang telah dilakukan ditampilkan sebagai berikut.

Tabel 3.19 Perubahan nilai fungsi objektif

Iterasi ke-	Jumlah Fungsi Objektif	Perubahan Fungsi Objektif
1	135,609	9864,391
2	135,371	0,238
3	134,896	0,475
4	133,644	1,252
5	130,743	2,900
6	124,582	6,161
7	115,799	8,783
8	109,793	6,006
9	107,111	2,682
10	104,375	2,736
11	98,932	5,443
12	91,791	7,141
13	87,558	4,234
14	86,300	1,258
15	86,015	0,284
16	85,952	0,063
17	85,938	0,014
18	85,935	0,003

Sesuai dengan nilai perubahan fungsi objektif pada tabel 3.19 diatas, maka perhitungan berhenti pada iterasi ke-18. Hal tersebut dikarenakan nilai perubahan fungsi objektif pada iterasi ke-18 yang bernilai 0,0158 telah mencapai batas ambang.

10. Langkah selanjutnya adalah menentukan *cluster* yang diikuti oleh setiap data. *Cluster* yang diikuti adalah *cluster* yang memiliki nilai derajat keanggotaan terbesar. Berikut ini adalah hasil perhitungan nilai derajat keanggotaan pada iterasi ke-18.

Tabel 3.20 Nilai akhir derajat keanggotaan dan *cluster* yang diikuti

No	Derajat Keanggotaan Matriks			Terbesar	<i>Cluster</i> yang diikuti
	ui1	ui2	ui3		
1	0,579	0,286	0,135	0,579	1
2	0,060	0,873	0,067	0,873	2
3	0,074	0,120	0,806	0,806	3
4	0,038	0,923	0,039	0,923	2
5	0,048	0,174	0,777	0,777	3

6	0,033	0,097	0,870	0,870	3
7	0,069	0,878	0,053	0,878	2
8	0,053	0,155	0,792	0,792	3
9	0,322	0,332	0,346	0,346	3
10	0,023	0,075	0,902	0,902	3
...
50	0,983	0,013	0,004	0,983	1
51	0,791	0,145	0,063	0,791	1
52	0,062	0,892	0,046	0,892	2
53	0,958	0,031	0,011	0,958	1
54	0,923	0,054	0,023	0,923	1
55	0,062	0,892	0,046	0,892	2
56	0,938	0,045	0,017	0,938	1
57	0,966	0,025	0,009	0,966	1
58	0,913	0,063	0,024	0,913	1
59	0,913	0,063	0,024	0,913	1
60	0,935	0,048	0,017	0,935	1

Tabel 3.21 Centroid akhir

No	A	B	C	D	E	F	G	H	I	J	K	L
1	0,878	0,004	0,844	0,901	0,896	0,909	0,008	0,854	0,056	0,014	0,016	0,007
2	0,805	0,010	0,813	0,880	0,904	0,058	0,013	0,943	0,977	0,019	0,021	0,008
3	0,623	0,765	0,827	0,828	0,869	0,103	0,025	0,022	0,863	0,032	0,036	0,009

Program Studi teknik informatika Univeristas Muhammadiyah Gresik sudah menetapkan terdapat 3 Bidang Keilmuan yang ada didalamnya yaitu Artificial Intelegent, Rekayasa Perangkat Lunak , dan Jaringan Komputer . Dari ke 3 bidang ini di inisialisasikan ke dalam 3 cluster.

Tabel 3.22 Cluster Bidang Keilmuan

No	Mata Kuliah Pilihan	Inisialaisasi
1	Artificial Intelegent	Cluster 1
2	Rekayasa Perangkat Lunak	Cluster 2
3	Jaringan Komputer	Cluster 3

Tabel 3.23 Hasil Data Mahasiswa sesuai dengan Bidang Keilmuan

No	Nama	Nim	Bidang Keilmuan
1	DIMAS HABIB ABROR	170602004	Artificial Intelegent
2	WAJED GHALY	170602005	Rekayasa Perangkat Lunak
3	EDHI SAFITR	170602007	Jaringan Komputer
4	RYAN HARIS BAWAFI	170602008	Rekayasa Perangkat Lunak
5	ACHMAD MUZADI	170602010	Jaringan Komputer
6	MOH MALIK FAJAR	170602011	Jaringan Komputer
7	ERVIN SETYAWAN	170602012	Rekayasa Perangkat Lunak
8	ABDURROSYIID AMRULLAH	170602013	Jaringan Komputer
9	SRI MULYANI	170602016	Jaringan Komputer
10	M. FAHRUDIN	170602018	Jaringan Komputer
...

50	AHMAT SUBAIRI RAHMAN	180602036	Artifical Intelegant
51	ALFIN NIURIL FATHIL AKBAR	180602040	Artifical Intelegant
52	MUHAMMAD ADRYAN FIRDAUS	180602042	Rekayasa Perangkat Lunak
53	NADHIF NASRUDIN	180602044	Artifical Intelegant
54	MOCH. RIVNALDY BYANTARA	180602045	Artifical Intelegant
55	MUHAMMAD IQBAL WALIYUDDIIN	180602046	Rekayasa Perangkat Lunak
56	DIYAH UTAMI	180602048	Artifical Intelegant
57	MUHAMMAD WAHYU FAJAR FIRDAUS	180602049	Artifical Intelegant
58	RACHMAD DEO ABDIKA GILANG ATALA	180602053	Artifical Intelegant
59	PANHARESI	180602056	Artifical Intelegant
60	MOHAMMAD ICHSAN	180602057	Artifical Intelegant

3.10 Pengujian dan Evaluasi Sistem

Pada pengujian sistem pemetaan bidang keilmuan kali ini dilakukan pengujian dengan cara:

1. Menggunakan data mahasiswa teknik informatika Universitas Muhammadiyah Gresik angkatan 2017 -2018
2. Pengujian sistem yang akan dilakukan dengan cara memasukkan 60 data nilai mahasiswa yang berasal dari angkatan 2017-2018 yang selanjutnya akan di kelompokkan menggunakan metode *fuzzy C-means*
3. Data nilai mahasiswa teknik informatika Universitas Muhammadiyah Gresik yang di uji pada metode *fuzzy C-means* merupakan nilai Mata Kuliah Pilihan mahasiswa teknik informatika Universitas Muhammadiyah Gresik

3.10.1 Validitas Sihoulet Coefficient

Indeks validitas silhouette merupakan suatu ukuran statistik yang digunakan untuk menyeleksi permasalahan penentuan jumlah cluster optimal yang dapat merepresentasikan grafis singkat seberapa baik setiap objek terletak dalam cluster. Asumsikan data sudah dikelompokkan ke dalam cluster. Untuk setiap objek, misalkan $a(i)$ adalah rata-rata jarak objek ke semua objek dalam cluster yang sama dan $b(i)$ adalah rata-rata jarak minimum objek ke semua objek pada suatu cluster dimana bukan anggota cluster. Dari penjelasan yang telah dipaparkan indeks validitas silhouette dapat ditulis dengan perhitungan cara sebagai berikut:

1. Hitung rata-rata jarak objek dengan semua objek lain yang berada di dalam satu cluster dengan persamaan :

$$a_j^i = \frac{1}{m_j - 1} \sum_{\substack{r=1 \\ r \neq i}}^{m_j} d(x_j^i, x_r^i) \quad (3)$$

Ket :

- i : Indeks data
- j : Cluster
- $d(x_j^i, x_r^i)$: Jarak data ke-i dengan data ke-r dalam satu cluster j
- m_j : Jumlah data dalam cluster ke-j
- x : Data

2. Hitung rata-rata jarak objek dengan semua objek lain yang berada pada *cluster* lain, kemudian ambil nilai paling minimum dengan persamaan :

$$b_j^i = \min \left\{ \frac{1}{m_n} \sum_{\substack{r=1 \\ r \neq i}}^{m_n} d(x_j^i, x_r^n) \right\} \quad (4)$$

Ket:

b : Nilai minimum dari rata-rata jarak data ke-i terhadap semua data dari *cluster* yang lain

j : *Cluster*

$d(x_j^i, x_r^i)$: Jarak data ke-i dengan data ke- r dalam satu *cluster* j

m_n : Banyak data dalam 1 *cluster*

x : Data

3. Hitung nilai silhouette coefficient dengan persamaan :

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (5)$$

Ket :

i : Index

j : *Cluster*

m_j : Banyaknya data dalam *cluster* j

SI_j : *Silhouette Index cluster*

Hasil perhitungan menggunakan Algoritma *fuzzy c-means* dari ke 60 data sebagai berikut:

No	Nama	Nim	A	B	C	D	E	F	G	H	I	J	K	L	Cluster
1	Dimas	170602004	2,5	0	0	3	3,5	3,5	0	3,5	2	0	0	0	1
2	Wahid	170602005	2,5	0	4	3,5	4	0	0	4	4	0	0	0	2
3	Ezki	170602007	3	3,5	3,5	3	3,5	0	0	0	2	0	0	0	3
4	Ryan	170602008	3	0	3,5	3	3,5	0	0	4	4	0	0	0	2
5	Achmad	170602010	3	2	4	3,5	3,5	0	0	0	4	0	0	0	3
6	Mikhael	170602011	2	3	3,5	3,5	4	0	0	0	4	0	0	0	3
7	Erwin	170602012	4	0	4	3,5	4	0	0	4	4	0	0	0	2
8	Adhany	170602013	3	3	4	4	4	0	0	4	0	0	0	0	3
9	Sri	170602016	3	0	0	3	0	4	0	0	4	3,5	3,5	0	3
10	SL	170602018	2,5	2,5	3,5	3,5	3	0	0	0	4	0	0	0	3
...
50	Alman	180602036	3,5	0	3,5	3,5	3,5	3,5	0	3,5	0	0	0	0	1
51	Ahli	180602040	3,5	0	2,5	3,5	3,5	3	0	3,5	0	0	0	0	1
52	Mikhael	180602042	4	0	3,5	4	3,5	0	0	4	4	0	0	0	2
53	Nafal	180602044	3,5	0	4	3,5	3,5	4	0	3,5	0	0	0	0	1
54	Moh.	180602045	2,5	0	3,5	4	3,5	3,5	0	3,5	0	0	0	0	1

- Menghitung rata-rata jarak objek dengan semua objek lain yang berada di dalam satu cluster :

$$a_j^i = \frac{1}{24 - 1} \sum_{\substack{r=1 \\ r \neq 1}}^{m_j} d(x_1, x_1) + d(x_1, x_{27}) + \dots$$

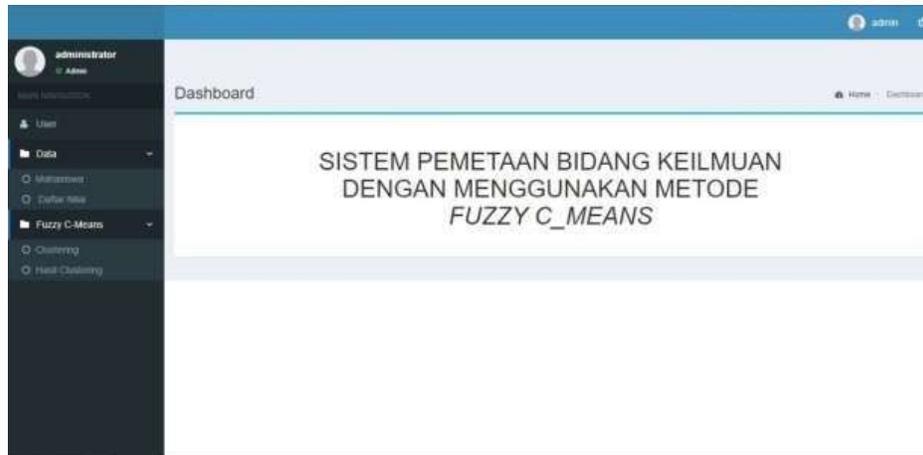
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27			
1	-	0,200	0,400	0,600	0,800	1,000	1,200	1,400	1,600	1,800	2,000	2,200	2,400	2,600	2,800	3,000	3,200	3,400	3,600	3,800	4,000	4,200	4,400	4,600	4,800	5,000	5,200	5,400	5,600	
2	0,200	-	0,400	0,600	0,800	1,000	1,200	1,400	1,600	1,800	2,000	2,200	2,400	2,600	2,800	3,000	3,200	3,400	3,600	3,800	4,000	4,200	4,400	4,600	4,800	5,000	5,200	5,400	5,600	
3	0,400	0,400	-	0,600	0,800	1,000	1,200	1,400	1,600	1,800	2,000	2,200	2,400	2,600	2,800	3,000	3,200	3,400	3,600	3,800	4,000	4,200	4,400	4,600	4,800	5,000	5,200	5,400	5,600	
4	0,600	0,600	0,600	-	0,800	1,000	1,200	1,400	1,600	1,800	2,000	2,200	2,400	2,600	2,800	3,000	3,200	3,400	3,600	3,800	4,000	4,200	4,400	4,600	4,800	5,000	5,200	5,400	5,600	
5	0,800	0,800	0,800	0,800	-	1,000	1,200	1,400	1,600	1,800	2,000	2,200	2,400	2,600	2,800	3,000	3,200	3,400	3,600	3,800	4,000	4,200	4,400	4,600	4,800	5,000	5,200	5,400	5,600	
6	1,000	1,000	1,000	1,000	1,000	-	1,200	1,400	1,600	1,800	2,000	2,200	2,400	2,600	2,800	3,000	3,200	3,400	3,600	3,800	4,000	4,200	4,400	4,600	4,800	5,000	5,200	5,400	5,600	
7	1,200	1,200	1,200	1,200	1,200	1,200	-	1,400	1,600	1,800	2,000	2,200	2,400	2,600	2,800	3,000	3,200	3,400	3,600	3,800	4,000	4,200	4,400	4,600	4,800	5,000	5,200	5,400	5,600	
8	1,400	1,400	1,400	1,400	1,400	1,400	1,400	-	1,600	1,800	2,000	2,200	2,400	2,600	2,800	3,000	3,200	3,400	3,600	3,800	4,000	4,200	4,400	4,600	4,800	5,000	5,200	5,400	5,600	
9	1,600	1,600	1,600	1,600	1,600	1,600	1,600	1,600	-	1,800	2,000	2,200	2,400	2,600	2,800	3,000	3,200	3,400	3,600	3,800	4,000	4,200	4,400	4,600	4,800	5,000	5,200	5,400	5,600	
10	1,800	1,800	1,800	1,800	1,800	1,800	1,800	1,800	1,800	-	2,000	2,200	2,400	2,600	2,800	3,000	3,200	3,400	3,600	3,800	4,000	4,200	4,400	4,600	4,800	5,000	5,200	5,400	5,600	
11	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	-	2,200	2,400	2,600	2,800	3,000	3,200	3,400	3,600	3,800	4,000	4,200	4,400	4,600	4,800	5,000	5,200	5,400	5,600	
12	2,200	2,200	2,200	2,200	2,200	2,200	2,200	2,200	2,200	2,200	2,200	-	2,400	2,600	2,800	3,000	3,200	3,400	3,600	3,800	4,000	4,200	4,400	4,600	4,800	5,000	5,200	5,400	5,600	
13	2,400	2,400	2,400	2,400	2,400	2,400	2,400	2,400	2,400	2,400	2,400	2,400	-	2,600	2,800	3,000	3,200	3,400	3,600	3,800	4,000	4,200	4,400	4,600	4,800	5,000	5,200	5,400	5,600	
14	2,600	2,600	2,600	2,600	2,600	2,600	2,600	2,600	2,600	2,600	2,600	2,600	2,600	-	2,800	3,000	3,200	3,400	3,600	3,800	4,000	4,200	4,400	4,600	4,800	5,000	5,200	5,400	5,600	
15	2,800	2,800	2,800	2,800	2,800	2,800	2,800	2,800	2,800	2,800	2,800	2,800	2,800	2,800	-	3,000	3,200	3,400	3,600	3,800	4,000	4,200	4,400	4,600	4,800	5,000	5,200	5,400	5,600	
16	3,000	3,000	3,000	3,000	3,000	3,000	3,000	3,000	3,000	3,000	3,000	3,000	3,000	3,000	3,000	-	3,200	3,400	3,600	3,800	4,000	4,200	4,400	4,600	4,800	5,000	5,200	5,400	5,600	
17	3,200	3,200	3,200	3,200	3,200	3,200	3,200	3,200	3,200	3,200	3,200	3,200	3,200	3,200	3,200	3,200	-	3,400	3,600	3,800	4,000	4,200	4,400	4,600	4,800	5,000	5,200	5,400	5,600	
18	3,400	3,400	3,400	3,400	3,400	3,400	3,400	3,400	3,400	3,400	3,400	3,400	3,400	3,400	3,400	3,400	3,400	-	3,600	3,800	4,000	4,200	4,400	4,600	4,800	5,000	5,200	5,400	5,600	
19	3,600	3,600	3,600	3,600	3,600	3,600	3,600	3,600	3,600	3,600	3,600	3,600	3,600	3,600	3,600	3,600	3,600	3,600	-	3,800	4,000	4,200	4,400	4,600	4,800	5,000	5,200	5,400	5,600	
20	3,800	3,800	3,800	3,800	3,800	3,800	3,800	3,800	3,800	3,800	3,800	3,800	3,800	3,800	3,800	3,800	3,800	3,800	3,800	-	4,000	4,200	4,400	4,600	4,800	5,000	5,200	5,400	5,600	
21	4,000	4,000	4,000	4,000	4,000	4,000	4,000	4,000	4,000	4,000	4,000	4,000	4,000	4,000	4,000	4,000	4,000	4,000	4,000	4,000	-	4,200	4,400	4,600	4,800	5,000	5,200	5,400	5,600	
22	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,200	-	4,400	4,600	4,800	5,000	5,200	5,400	5,600	
23	4,400	4,400	4,400	4,400	4,400	4,400	4,400	4,400	4,400	4,400	4,400	4,400	4,400	4,400	4,400	4,400	4,400	4,400	4,400	4,400	4,400	4,400	-	4,600	4,800	5,000	5,200	5,400	5,600	
24	4,600	4,600	4,600	4,600	4,600	4,600	4,600	4,600	4,600	4,600	4,600	4,600	4,600	4,600	4,600	4,600	4,600	4,600	4,600	4,600	4,600	4,600	4,600	-	4,800	5,000	5,200	5,400	5,600	
25	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800	-	5,000	5,200	5,400	5,600	
26	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	-	5,200	5,400	5,600	
27	5,200	5,200	5,200	5,200	5,200	5,200	5,200	5,200	5,200	5,200	5,200	5,200	5,200	5,200	5,200	5,200	5,200	5,200	5,200	5,200	5,200	5,200	5,200	5,200	5,200	5,200	-	5,400	5,600	
28	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	-	5,600	
29	5,600	5,600	5,600	5,600	5,600	5,600	5,600	5,600	5,600	5,600	5,600	5,600	5,600	5,600	5,600	5,600	5,600	5,600	5,600	5,600	5,600	5,600	5,600	5,600	5,600	5,600	5,600	5,600	-	
30	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800	5,800

- Menghitung rata-rata jarak objek dengan semua objek lain yang berada pada cluster lain, kemudian ambil nilai paling minimum

$$b_j^i = \min\{(5,852,6,832)\}$$

		Data ke-1																			
Jarak		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
		3	5.74	5.77	5.83	5.89	5.91	5.98	6.04	6.11	6.18	6.25	6.31	6.38	6.45	6.51	6.58	6.65	6.71	6.78	6.85
4	5.85	5.88	5.94	6.00	6.02	6.09	6.15	6.22	6.29	6.36	6.42	6.49	6.56	6.62	6.69	6.76	6.82	6.89	6.96	7.02	7.09
7	6.03	6.06	6.12	6.18	6.20	6.27	6.33	6.40	6.47	6.54	6.60	6.67	6.74	6.80	6.87	6.94	7.00	7.07	7.14	7.20	7.27
10	6.20	6.23	6.29	6.35	6.37	6.44	6.50	6.57	6.64	6.71	6.77	6.84	6.91	6.97	7.04	7.11	7.17	7.24	7.31	7.37	7.44
12	6.30	6.33	6.39	6.45	6.47	6.54	6.60	6.67	6.74	6.81	6.87	6.94	7.01	7.07	7.14	7.21	7.27	7.34	7.41	7.47	7.54
15	6.45	6.48	6.54	6.60	6.62	6.69	6.75	6.82	6.89	6.96	7.02	7.09	7.16	7.22	7.29	7.35	7.42	7.49	7.55	7.62	7.69
17	6.55	6.58	6.64	6.70	6.72	6.79	6.85	6.92	6.99	7.05	7.12	7.19	7.25	7.32	7.38	7.45	7.52	7.58	7.65	7.72	7.79
20	6.70	6.73	6.79	6.85	6.87	6.94	7.00	7.07	7.14	7.20	7.27	7.34	7.40	7.47	7.53	7.60	7.67	7.73	7.80	7.87	7.94
25	7.00	7.03	7.09	7.15	7.17	7.24	7.30	7.37	7.44	7.50	7.57	7.64	7.70	7.77	7.83	7.90	7.97	8.03	8.10	8.17	8.24
30	7.30	7.33	7.39	7.45	7.47	7.54	7.60	7.67	7.74	7.80	7.87	7.94	8.00	8.07	8.13	8.20	8.27	8.33	8.40	8.47	8.54
35	7.60	7.63	7.69	7.75	7.77	7.84	7.90	7.97	8.04	8.10	8.17	8.23	8.30	8.36	8.43	8.49	8.56	8.62	8.69	8.75	8.82
40	7.90	7.93	7.99	8.05	8.07	8.14	8.20	8.27	8.33	8.40	8.46	8.53	8.59	8.65	8.72	8.78	8.85	8.91	8.98	9.04	9.11
45	8.20	8.23	8.29	8.35	8.37	8.44	8.50	8.57	8.63	8.70	8.76	8.83	8.89	8.95	9.02	9.08	9.15	9.21	9.28	9.34	9.41
50	8.50	8.53	8.59	8.65	8.67	8.74	8.80	8.87	8.93	9.00	9.06	9.13	9.19	9.25	9.32	9.38	9.45	9.51	9.58	9.64	9.71
55	8.80	8.83	8.89	8.95	8.97	9.04	9.10	9.17	9.23	9.30	9.36	9.43	9.49	9.55	9.62	9.68	9.75	9.81	9.88	9.94	10.01
60	9.10	9.13	9.19	9.25	9.27	9.34	9.40	9.47	9.53	9.60	9.66	9.73	9.79	9.85	9.92	9.98	10.05	10.11	10.18	10.24	10.31
Rate-rata	5.74	5.77	5.83	5.89	5.91	5.98	6.04	6.11	6.18	6.25	6.31	6.38	6.45	6.51	6.58	6.65	6.71	6.78	6.85	6.91	6.98

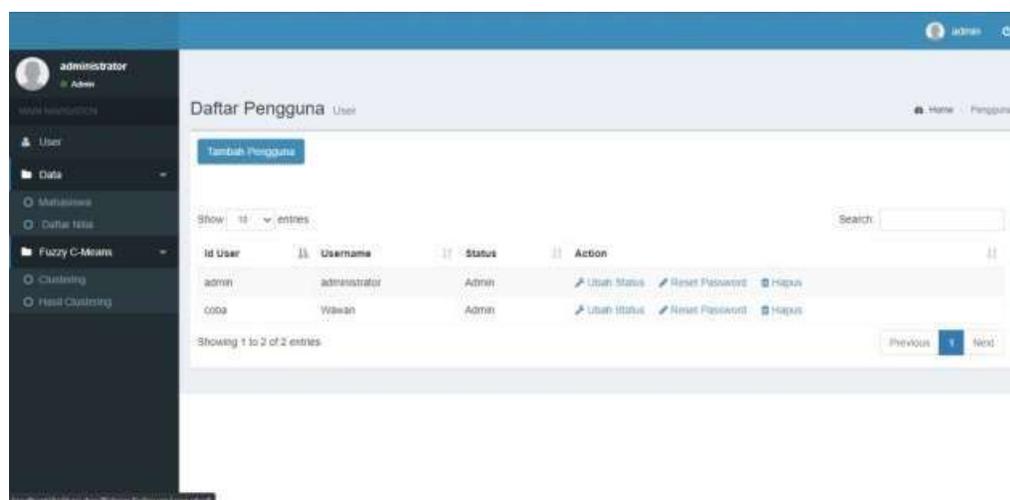
		Data ke-1																			
Jarak		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
		5	7.01	7.04	7.10	7.16	7.18	7.25	7.31	7.38	7.45	7.51	7.58	7.64	7.71	7.77	7.84	7.90	7.97	8.03	8.10
7	7.20	7.23	7.29	7.35	7.37	7.44	7.50	7.57	7.64	7.71	7.77	7.84	7.91	7.97	8.04	8.10	8.17	8.23	8.30	8.36	8.43
9	7.40	7.43	7.49	7.55	7.57	7.64	7.70	7.77	7.84	7.90	7.97	8.03	8.10	8.16	8.23	8.29	8.35	8.42	8.48	8.55	8.61
11	7.60	7.63	7.69	7.75	7.77	7.84	7.90	7.97	8.03	8.10	8.16	8.23	8.29	8.35	8.42	8.48	8.55	8.61	8.68	8.74	8.81
13	7.80	7.83	7.89	7.95	7.97	8.04	8.10	8.17	8.23	8.30	8.36	8.42	8.49	8.55	8.61	8.68	8.74	8.81	8.87	8.94	9.00
15	8.00	8.03	8.09	8.15	8.17	8.24	8.30	8.37	8.43	8.50	8.56	8.62	8.69	8.75	8.81	8.88	8.94	9.01	9.07	9.14	9.20
17	8.20	8.23	8.29	8.35	8.37	8.44	8.50	8.57	8.63	8.70	8.76	8.83	8.89	8.95	9.02	9.08	9.15	9.21	9.28	9.34	9.41
19	8.40	8.43	8.49	8.55	8.57	8.64	8.70	8.77	8.83	8.90	8.96	9.03	9.09	9.15	9.22	9.28	9.34	9.41	9.47	9.54	9.60
21	8.60	8.63	8.69	8.75	8.77	8.84	8.90	8.97	9.03	9.10	9.16	9.23	9.29	9.35	9.42	9.48	9.54	9.61	9.67	9.74	9.80
23	8.80	8.83	8.89	8.95	8.97	9.04	9.10	9.17	9.23	9.30	9.36	9.42	9.49	9.55	9.61	9.68	9.74	9.81	9.87	9.94	10.00
25	9.00	9.03	9.09	9.15	9.17	9.24	9.30	9.37	9.43	9.50	9.56	9.62	9.69	9.75	9.81	9.88	9.94	10.01	10.07	10.14	10.20
27	9.20	9.23	9.29	9.35	9.37	9.44	9.50	9.57	9.63	9.70	9.76	9.83	9.89	9.95	10.02	10.08	10.15	10.21	10.28	10.34	10.41
30	9.50	9.53	9.59	9.65	9.67	9.74	9.80	9.87	9.93	10.00	10.06	10.13	10.19	10.25	10.32	10.38	10.45	10.51	10.58	10.64	10.71
Rate-rata	7.01	7.04	7.10	7.16	7.18	7.25	7.31	7.38	7.45	7.51	7.58	7.64	7.71	7.77	7.84	7.90	7.97	8.03	8.10	8.16	8.23



Gambar 4.2 Halaman Beranda

4.13 Halaman User

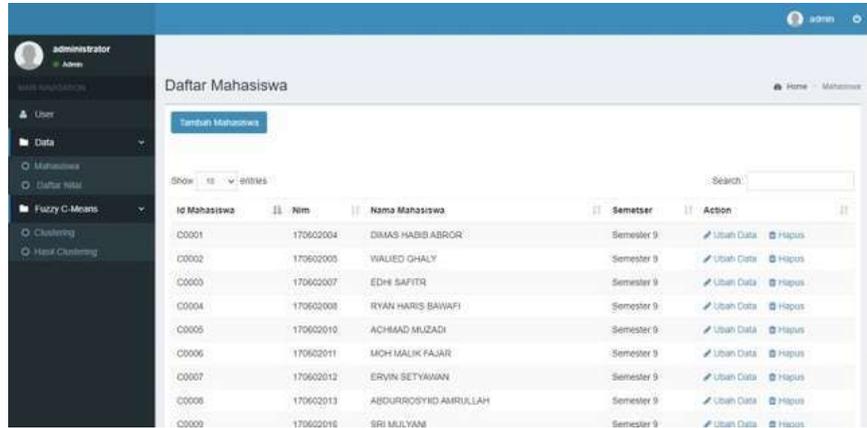
Pada halaman user merupakan sebuah halaman untuk menambahkan user, mengubah dan menghapus user yang ada pada sistem.



Gambar 4.3 Halaman Daftar Pengguna

4.14 Halaman Mahasiswa

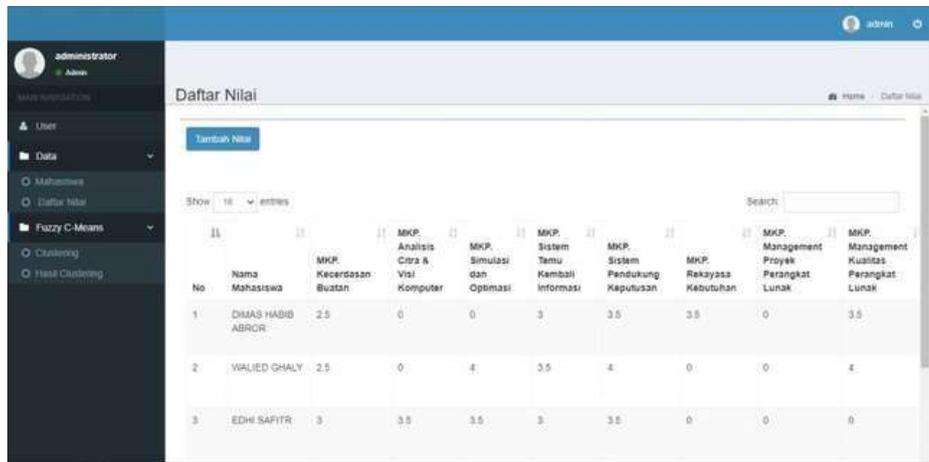
Pada halaman mahasiswa merupakan sebuah halaman yang berisi daftar mahasiswa teknik informatika Universitas Muhammadiyah Gresik. Disini juga bisa menambahkan mengubah dan menghapus data mahasiswa.



Gambar 4.4 Halaman Daftar Mahasiswa

4.15 Halaman Daftar Nilai

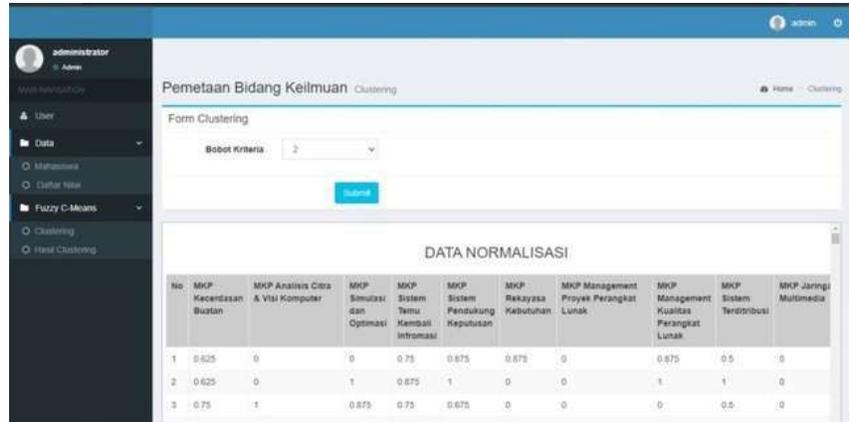
Pada halaman daftar nilai merupakan halaman yang menampung daftar nilai mahasiswa dimana daftar nilai ini terdiri dari 12 mata kuliah pilihan.



Gambar 4.5 Halaman Daftar Nilai

4.16 Halaman Clustering

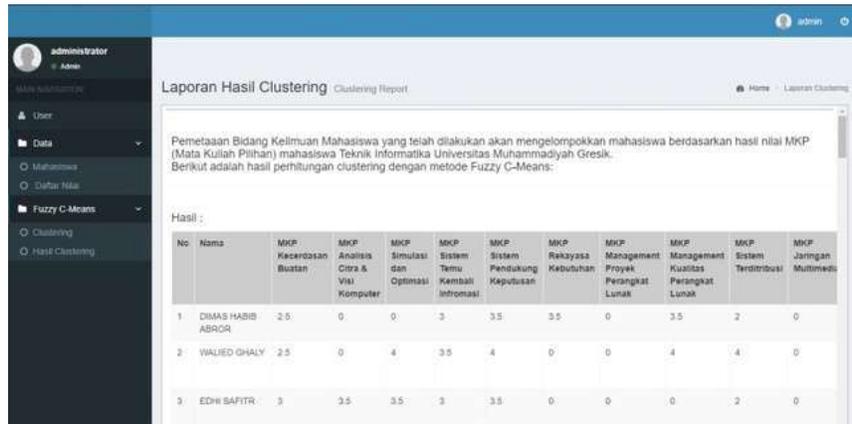
Pada halaman clustering merupakan halaman dimana data nilai mahasiswa yang ada di daftar nilai akan di olah menggunakan metode Fuzzy C-means yang mana data akan di uji sesuai dengan algoritma fuzzy C-means. Disini akan menghasilkan nilai Silhouette Index dan Purty pada setiap perhitungannya.



Gambar 4.6 Halaman Clustering

4.1.7 Halaman Hasil Clustering

Pada halaman hasil Clustering data yang sudah diolah menggunakan Fuzzy C-Means akan di beri label apakah mahasiswa tersebut masuk kebidang keilmuan yang mana.



Gambar 4.7 Halaman Hasil Clustering

4.2 Pengujian Sistem

Pengujian sistem pada pemetaan bidang keilmuan kali ini menggunakan Silhouette Coefficient (SC) dan Purity.

4.2.1 Silhouette Coefficient (SC)

Sebelum mencari Silhouette Coefficient (SC) dilakukan perhitungan Silhouette Index (SI) digunakan untuk memvalidasi sebuah data dan keseluruhan cluster. Untuk mendapatkan Silhouette Index (SI) dapat menggunakan persamaan berikut :

$$S(i) = \frac{b(i) - a(i)}{\text{Max}\{a(i), b(i)\}}$$

Keterangan:

a: Rata-rata jarak data ke-i terhadap semua data lainnya dalam satu cluster

b: Nilai minimum dari rata-rata jarak data ke-i terhadap semua data dari cluster yang lain $\max(a_j^i, b_j^i)$

Nilai maksimum dari nilai a dan b dari satu data

SI: Silhouette Index.

Nilai SC didapat dengan mencari nilai maksimum SI global dari jumlah cluster sampai jumlah cluster n-1, seperti pada persamaan berikut :

$$SI = \frac{1}{k} \sum_j^k SI_j$$

$$SC = \text{Max}_j SI(j)$$

Keterangan :

k : Jumlah Cluster

j : Cluster

SC : Silhouette Coefficient

SI_j : Silhouette Index Cluster

SI : Silhouette Index global

4.2.2 Purity

Sebuah cluster dikatakan murni (pure) apabila semua objek data dengan class yang sama berada pada cluster yang sama pula (Nurul, et al 2015). Purity digunakan untuk menghitung kemurnian suatu cluster yang direpresentasikan sebagai anggota cluster yang paling cocok pada suatu class. Untuk menghitung purity setiap cluster, digunakan rumus berikut:

$$Purity(j) = 1/N_j * (n_i)$$

Keterangan:

Purity(j) = nilai *purity* untuk variabel ke-j

N_j = jumlah data yang menjadi anggota *cluster ke-j*

i,j = indeks dari *cluster*

Pengujian kali ini akan dilakukan dengan beberapa kali percobaan pengujian dengan menguji bobot kriteria dimana sebelumnya pemberian nilai bobot dalam clustering Fuzzy C-Means dilakukan secara manual. Dipengujian kali ini akan dilakukan pengujian bobot yang paling optimal digunakan Fuzzy C-Means.

Tabel 4.1 Hasil Pengujian *Silhouette Index* dan *Purity*

Bobot 2

PERCOBA AN	DATA CLUSTER			SI	Purity
	AI	RPL	JARINGAN		
1	24	16	20	0.53247	0.8096
2	24	32	4	0.44935	1.6345
3	24	16	20	0.53247	0.8096
4	4	24	32	0.44935	1.6345
5	20	24	16	0.53247	0.8096
6	16	24	20	0.53247	0.8096
7	20	16	24	0.53247	0.8096
8	20	24	16	0.53247	0.8096
9	24	4	32	0.44935	1.6345
10	24	20	16	0.53247	0.8096
11	24	16	20	0.53247	0.8096

Bobot 3

PERCOBA AN	DATA CLUSTER			SI	Purity
	AI	RPL	JARINGAN		
1	16	24	20	0.53247	0.8096
2	20	16	24	0.53247	0.8096
3	24	20	16	0.53247	0.8096
4	20	24	16	0.53247	0.8096
5	20	16	24	0.53247	0.8096
6	16	24	20	0.53247	0.8096

7	20	16	24	0.53247	0.8096
8	16	24	20	0.53247	0.8096
9	20	16	24	0.53247	0.8096
10	24	20	16	0.53247	0.8096
11	20	24	16	0.53247	0.8096

Bobot 4

PERCOBA AN	DATA CLUSTER			SI	Purity
	AI	RPL	JARINGAN		
1	25	19	16	0.49954	0.8172
2	21	15	24	0.52609	0.8196
3	27	5	28	0.26421	1.4398
4	20	24	16	0.53247	0.81102
5	16	20	24	0.53247	0.8110
6	30	5	25	0.35809	1.4365
7	31	5	24	0.43604	1.4454
8	24	20	16	0.53247	0.81102
9	43	12	5	0.42498	1.6966
10	30	24	6	0.39671	1.272
11	25	19	16	0.49954	0.8172

Bobot 5

PERCOBA AN	DATA CLUSTER			SI	Purity
	AI	RPL	JARINGAN		
1	23	7	30	0.39720	1.1544
2	26	3	31	0.41477	2.1233
3	3	19	38	0.41477	2.1708
4	12	26	22	0.42093	0.88127
5	2	24	34	0.42093	3.0039
6	24	8	28	0.32200	0.88155
7	15	19	26	0.53117	0.83057
8	2	24	34	0.50305	3.0005
9	7	15	38	0.28088	1.2410
10	31	24	5	0.43604	1.4385
11	29	8	23	0.26596	1.0687

Bobot 6

PERCOBA AN	DATA CLUSTER			SI	Purity
	AI	RPL	JARINGAN		
1	5	24	31	0.37774	1.4364
2	7	29	24	0.40066	1.2857
3	25	18	17	0.43763	0.81332
4	26	15	19	0.53117	0.83030
5	4	32	24	0.43138	1.6938
6	15	24	21	0.52891	0.82121
7	20	31	9	0.39410	1.0164
8	10	26	24	0.36581	0.95026
9	23	28	9	0.38934	0.99822
10	26	7	27	0.40829	0.83030
11	16	25	19	0.53397	0.813030

Bobot 7

PERCOBA AN	DATA CLUSTER			SI	Purity
	AI	RPL	JARINGAN		
1	19	15	26	0.53117	0.83042
2	24	20	16	0.53247	0.8096
3	9	24	27	0.35297	1.0021
4	12	26	22	0.42093	0.8814
5	25	27	8	0.37399	1.0621
6	13	19	28	0.25237	0.87052
7	3	32	25	0.53117	0.83080
8	26	19	15	0.53117	0.83080
9	17	18	25	0.35911	1.1027
10	27	29	4	0.34216	1.6903
11	19	15	26	0.53117	0.83042

Bobot 8

PERCOBA AN	DATA CLUSTER			SI	Purity
	AI	RPL	JARINGAN		
1	13	22	25	0.44091	0.85407
2	25	15	20	0.34877	1.4308
3	12	28	20	0.52848	0.82451
4	30	26	4	0.44865	1.6981
5	24	31	5	0.43604	1.4361
6	28	26	6	0.36506	1.2626
7	4	25	31	0.43604	1.6893
8	20	26	14	0.53616	1.2657
9	13	22	25	0.44091	0.85407
10	25	15	20	0.34877	1.4308
11	12	28	20	0.52848	0.82451

Kriteria subjektif pengukuran baik atau tidaknya pengelompokan berdasarkan SC disajikan dalam tabel berikut (Kaufman & Peter Rousseeuw, 1990):

Tabel 4.2 Kriteria Pengelompokan berdasarkan Silhouette Coefficient

Nilai SC	Interpretasi SC
0.71-100	Struktur Kuat
0,51 – 0,70	Struktur Sedang
0,26 – 0,50	Struktur Lemah
<0,25	Tidak Terstruktur

V. PENUTUP

5.1 Kesimpulan

Hasil dari pembuatan penelitian ini algoritma fuzzy c-means dapat mengelompokkan bidang keilmuan mahasiswa teknik informatika yaitu bidang keilmuan artificial intelegent. Rekayasa perangkat lunak, dan jaringan. Pengujian validasi cluster dengan metode Silhouette Coefficient dan Purity pengujian menggunakan bobot yang berbeda-beda untuk menentukan bobot paling optimal yang bisa digunakan. Pada pengujian pada data nilai transkrip mahasiswa tahun 2017-2018 menyatakan bahwa bobot paling optimal yang bisa digunakan adalah bobot 3, dari 3 cluster yang terbentuk mendapatkan nilai SC 0,532 merupakan struktur yang sedang, sehingga cluster yang terbentuk masih dapat perpotongan dengan cluster yang lain. Untuk melihat kemurnian sebuah cluster maka di lakukan uji purity, uji ini dilakukan untuk melihat apakah sebuah cluster dapat merepresentasikan data anggota cluster didalamnya, dengan nilai purity yang 0,806 maka bisa dikatakan cluster yang terbentuk bisa mencerminkan anggota clusternya.

5.2 Saran

Dalam metode fuzzy c-means pada penelitian ini data yang digunakan sebagai data uji harus diganti karena data pada nilai transkrip mata kuliah pilihan mahasiswa cenderung naik turun sehingga sangat sulit untuk dilakukan pemetaan. Selain itu untuk menghasilkan bobot nilai yang bagus sebaiknya menggabungkan dengan metode weighted fuzzy c-means dengan fuzzy c-means agar mendapatkan bobot nilai yang maksimal.

DAFTAR PUSTAKA

- Kusumadewi, S., Purnomo, H., 2010, *Aplikasi Logika Fuzzy untuk Pendukung Keputusan edisi 2*, Graha Ilmu, Yogyakarta.
- Luthfi, Emha Taufiq. 2007. *Fuzzy C-Means untuk clustering data (studi kasus: data performance mengajar dosen)*. Seminar Nasional Teknologi 2007 (SNT 2007)
- Sapto, D. Anggoro. 2014. *Aplikasi Pencarian Potensi Akademik Mahasiswa Di Bidang Keahlian Rekayasa Perangkat Lunak Dan Sistem Cerdas Menggunakan Fuzzy C-Means*. Universitas Muhammadiyah Gresik, Gresik
- Mirza, M. Faisal. 2009. *Metode Clustering dengan Algoritma Fuzzy C-Means untuk rekomendasi bidang keahlian pada program studi Teknik Informatika*. Universitas Dian Nurwantoro, Semarang.
- Sikumbang, E. Delima. 2018, *Penerapan Data Mining Penjualan Sepatu Menggunakan Metode Algoritma Apriori*, AMIK BSI Jakarta, Jakarta.
- Varani, B. Yorangga. 2018, *Pengelompokan Data Guru Untuk Pemilihan Calom Pengawas Satuan Pendidikan Menggunakan Algoritma Fuzzy C-Means*, Universitas Sanata Dharma, Yogyakarta