



CLUSTER :
PENELITIAN PENGEMBANGAN PENDIDIKAN TINGGI

**KLUSTERISASI BIDANG KEAHLIAN
MAHASISWA MENGGUNAKAN TEKNIK**

DATA MINING DAN FUZZY C-MEANS

**UNTUK Mendukung Rencana Strategis
PROGRAM STUDI TEKNIK INFORMATIKA**



Oleh :

Dr. Elin Haerani, ST, M.Kom (19810522007102003/2023058101)

Novriyanto, ST, M.Sc (197711282007101003/2028117702)

**LEMBAGA PENELITIAN DAN PENGABDIAN KEPADA MASYARAKAT
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU
PEKANBARU**

2019

ABSTRAK

Penelitian ini merupakan salah satu upaya untuk membantu memberikan rekomendasi kepada mahasiswa dalam memilih bidang keahlian yang sesuai dengan kemampuan mereka. Penelitian ini juga memberikan informasi penting terkait rencana strategis bagi Jurusan TIF berdasarkan interpretasi dari hasil penelitian dalam pengelolaan mahasiswa dengan memfasilitasi memfokuskan mahasiswa terhadap ilmu yang didapat dari matakuliah sebelumnya, agar keahlian mahasiswa dapat lebih terarah dan juga bertujuan mengarahkan para mahasiswa dalam pemilihan topik penelitian tugas akhir. Analisa dilakukan terhadap data-data di peroleh dari data akademis mahasiswa Prodi Teknik Informatika tahun 2014/2015 dan di proses menggunakan data mining dan Algoritma *Fuzzy C-Means* Desain terhadap kebutuhan analisa dilakukan dengan menggunakan menggunakan konsep UML meliputi *usecase*, *sequence diagram*, *class diagram* dan *activity diagram*, mulai perancangan interface sampai kepada perancangan database dan pengkodean system. Produk yang dihasilkan adalah Perangkat Lunak untuk klusterisasi bidang keahlian mahasiswa Menggunakan teknik data mining dan *fuzzy c-means* dengan hasil pengujian untuk pengujian validitas perangkat lunak dengan nilai rata-rata 0,80 dan praktikalitas perangkat lunak dengan nilai rata-rata 85,33.

Kata kunci : *Data mining, Fuzzy C-Means, Bidang Keahlian*

DAFTAR ISI

PENGESAHAN	i
ABSTRAK	ii
KATA PEGANTAR	iii
DAFTAR ISI	iv
DAFTAR GAMBAR	vi
DAFTAR TABEL.....	vii
BAB I: PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	5
1.3. Batasan Masalah	5
1.4. Tujuan Penelitian	5
1.5. Manfaat Penelitian	5
1.6. Sistematika Penulisan	6
BAB II: TINJAUAN PUSTAKA	7
2.1. Data Mining	7
2.1.1 Teknik Data Mining	9
2.2. Metode Clustering Algoritma Fuzzy C-Means	10
2.2.1 Clustering	11
2.2.1 Fuzzy C-Means	11
2.2.1 Algoritma Fuzzy Clustering Means (FCM)	12
2.3. Penelitian yang Relevan	14
BAB III: METODOLOGI PENELITIAN	17
3.1. Kerangka Pikir Penelitian	17
BAB IV : ANALISA DAN PEMBAHASAN.....	22
4.1. Analisa Sistem Lama	22
4.2. Analisa Sistem Baru	22

4.2.1. Analisis Kebutuhan Data	25
4.2.1.1. Tahapan Proses Algoritma Fuzzy C-Means.....	25
4.2.1.1. Uraian Algoritma Fuzzy C-Means	26
4.2.2. Analisis Fungsional Sistem	32
4.3. Perancangan Sistem	37
4.3.1. Perancangan Struktur Menu	37
4.3.2. Perancangan Basis Data.....	37
4.3.3. Perancangan Antarmuka (Interface).....	38
 BAB V :PENGUJIAN DAN HASIL	 42
5.1. Implementasi	42
5.1.1. Batasan Implementasi	42
5.1.2. Lingkungan Implementasi	43
5.1.3. Implementasi Antar Muka Sistem	43
5.2. Pengujian	48
5.2.1. Pengujian Sistem dengan <i>Blackbox</i>	49
5.2.2. Uji Validasi Perangkat Lunak	51
5.2.3. Uji Praktikalitas Kepastian Perangkat Lunak Oleh Jurusan	53
 BAB VI : KESIMPULAN DAN SARAN	 55
6.1. Kesimpulan.....	55
6.2. Saran	55

DAFTAR PUSTAKA

KATA PENGANTAR

Assalamu'alaikum Wr. Wb.

Alhamdulillah, puji dan syukur senantiasa diucapkan ke hadirat Allah SWT, atas segala limpahan anugerah dan petunjuk-Nya, penelitian dengan judul “**Klusterisasi Bidang Keahlian Mahasiswa Menggunakan Teknik Data Mining Dan *Fuzzy C-Means* Untuk Mendukung Rencana Strategis Program Studi Teknik Informatika**” ini dapat diselesaikan. Shalawat serta salam semoga senantiasa tercurah atas Nabi Muhammad SAW, para kerabat, serta pengikutnya hingga hari akhir nanti.

Penulis menyadari bahwa laporan ini tidaklah lepas dari kekurangan dan kesalahan. Untuk itu, kritik dan saran yang sifatnya membangun sangat penulis harapkan. Penulis berharap semoga laporan ini dapat bermanfaat bagi semua pihak, khususnya dalam bidang pendidikan Teknik Informatika.

Wassalamu'alaikum Wr.Wb.

Pekanbaru, 28 November 2019

Peneliti

BAB I

PENDAHULUAN

1.1. Latar Belakang

Sumber utama yang dapat digunakan untuk menghasilkan informasi adalah data. Pengolahan data secara maksimal akan menghasilkan informasi yang berkualitas dan bermanfaat bagi Perguruan Tinggi. Perguruan Tinggi memiliki banyak data yang terkait dengan masalah akademik dan kemahasiswaan. Dari waktu ke waktu, data ini akan terus meningkat dan menjadi tumpukan data dalam bentuk *softfiles* atau *hardfiles*. Jika data tidak dikelola dengan baik, tidak ada manfaatnya. Dengan memanfaatkan teknik data mining dalam pengolahan data, data dapat diolah menjadi informasi yang berguna dan bermanfaat bagi Perguruan Tinggi.

Data mining adalah penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari data yang sangat besar (Davies, 2004). Menurut Pramudiono (2007) data mining juga dapat disebut sebagai rangkaian proses untuk mengeksplorasi nilai tambah dalam bentuk pengetahuan yang belum diketahui secara manual dari kumpulan data. Data mining juga disebut *Knowledge Discovery in Database* (KDD), yang merupakan kegiatan yang mencakup pengumpulan, menggunakan data untuk menemukan pesan, pola dan hubungan dalam kumpulan data besar. Hasil data mining dapat digunakan untuk membuat keputusan. Diharapkan bahwa menggunakan data mining dapat memberikan pengetahuan dan informasi yang sebelumnya tersembunyi di gudang data sehingga dapat menjadi informasi dan pengetahuan yang menarik dan berguna.

Program Studi Teknik Informatika Fakultas Sains dan Teknologi UIN SUSKA memiliki visi untuk menjadi pusat pendidikan, penelitian, dan pengembangan studi teknik informatika terkemuka dengan mensinergikan antara Sains dan Teknologi (IPTEK) dan Iman dan Takwa (IMTAK) dengan kebutuhan

tenaga kerja lokal, nasional atau internasional. Untuk mewujudkan visi ini Program Studi Teknik Informatika memiliki misi yaitu menyiapkan lulusan Islam yang profesional di bidang teknik informatika, memiliki kemampuan menganalisis masalah secara sistematis, tepat dan logis, serta mampu memberikan solusi yang tepat dalam kerangka kerja layanan dan layanan terbaik bagi masyarakat dan industri.

Sejak berdirinya program studi Teknik Informatika (TIF) dari tahun 1999 hingga 2018, memiliki mahasiswa mencapai 5000 orang, baik yang lulus maupun masih aktif sebagai mahasiswa. Dengan sejumlah besar mahasiswa, data akademik yang dimiliki juga meningkat. Data tersebut dapat digunakan sebagai sumber data strategis yang sangat berguna bagi program studi Teknik Informatika Fakultas Sains dan Teknologi. Jurusan Teknik Informatika mempunyai dua kelompok besar bidang keahlian mahasiswa yaitu, Ilmu Komputer dan Teknologi Informasi. Dari kelompok besar tersebut terdapat turunan bidang keahlian mahasiswa seperti *Software Development, Intelligent System, Mobile Computing, Data Mining, dan Networking*. Bidang keahlian adalah fokus mahasiswa pada bidang studi tertentu sesuai dengan minat mereka. Tujuannya adalah untuk memfokuskan mahasiswa lebih banyak pada pengetahuan yang diperoleh dari mata kuliah sebelumnya, sehingga dapat lebih terarah. Selain itu, bidang keahlian bertujuan untuk mengarahkan mahasiswa, dalam memberikan beberapa pilihan mengenai topik penelitian tugas akhir.

Melaksanakan pendidikan di bidang yang relevan dengan bidang keahlian dan minat akan dapat membawa efek yang baik dalam proses pembelajaran. Pendidikan berbasis kemampuan dan minat bertujuan untuk memastikan bahwa saat belajar atau nanti dalam pekerjaan seseorang akan berada di bidang yang konsisten dengan minat dan kemampuan yang dimiliki, sehingga dapat mengembangkan kemampuan dalam belajar dan bekerja secara optimal. Pemilihan bidang keahlian yang tidak tepat akan memiliki dampak bahwa nilai yang diperoleh akan kurang memuaskan dalam bidang keahlian telah dipilih, serta mahasiswa akan mengalami kesulitan dalam menyerap materi kuliah.

Dalam pemilihan bidang keahlian banyak mahasiswa mengalami kesulitan menentukan bidang keahlian mana yang akan mereka ambil. Masalah yang dihadapi oleh mahasiswa adalah kebingungan dan kesulitan dalam menentukan pilihan yang sesuai dengan kemampuan akademik dan minat mereka. Pemilihan bidang keahlian terkadang dipengaruhi oleh teman, atau dengan kata lain memilih hanya berdasarkan rekomendasi dari teman dan tidak berdasarkan kemampuan. Pilihan bidang keahlian tanpa pertimbangan yang matang terlebih dahulu dan sesuai dengan kemampuan akademik dan minat dapat membawa dampak tidak baik. Minat adalah sumber motivasi yang mendorong seseorang untuk melakukan apa yang diinginkan jika orang tersebut diberi kebebasan untuk memilih (Hurlock, 1999). Minat adalah kecenderungan seseorang untuk tertarik pada pengalaman dan berkelanjutan. Oleh karena itu perlu adanya suatu bentuk rekomendasi dari program studi sebagai panduan bagi mahasiswa dalam memilih bidang keahlian yang tepat dan sesuai dengan tingkat kemampuan akademik dan minatnya.

Untuk memberikan rekomendasi kepada mahasiswa berdasarkan begitu banyak data mahasiswa tentu tidak mudah karena harus mengelola data sebesar itu yang memiliki banyak *field* dan jumlah *record*. Teknologi data mining adalah salah satu alat untuk mengekstraksi data dalam database besar dan dengan spesifikasi kerumitan yang tinggi. Penelitian yang dilakukan oleh Varghese et al., (2011), Inyang & Joshua, (2013) berjudul "*Clustering Student Data to Characterize Performance Patterns*" dan "*Fuzzy Clustering of Students Data Repository for At-Risks Students Identification and Monitoring*", menggunakan teknik data mining dengan jumlah data 400 mahasiswa yang dilakukan oleh Varghese et al dan 496 data mahasiswa yang dilakukan oleh Inyang & Joshua untuk mengekstrak pola sehingga pengambilan keputusan dapat dibuat dari database besar. Model Fuzzy C-Means dapat membantu dalam mencapai proses pemantauan kemajuan mahasiswa dan mengidentifikasi fitur apa yang mempengaruhi kinerja mahasiswa sehingga dapat membantu pendidik mengetahui tingkat kemampuan setiap mahasiswa. Penelitian yang dilakukan oleh Setiyono & Isnanto (2008) membahas pendekatan fuzzy clustering dengan membagi kelas berdasarkan nilai prestasi mahasiswa dalam mata pelajaran yang merupakan syarat untuk mengambil mata pelajaran baru

dimana jumlah data yang akan dikelompokkan adalah 70 mahasiswa. data. Hasil penelitian menunjukkan bahwa Fuzzy C-Means lebih cocok untuk alokasi mahasiswa ke dalam beberapa kelas berdasarkan masukan dari jumlah cluster yang diinginkan. Munandar et al., (2013) menganalisis 126 data mahasiswa sesuai dengan bobot subjek tertentu dengan menggunakan konsep Fuzzy C-Means sehingga mereka mampu memberikan dukungan keputusan tentang pengelompokan data nilai mahasiswa untuk menentukan klasifikasi konsentrasi di mana seorang mahasiswa harus dimasukkan.

Analisis data mining yang digunakan dalam penelitian ini adalah teknik *clustering* dengan algoritma Fuzzy C-means yang bertujuan untuk *clustering* (mengelompokkan) mahasiswa berdasarkan bidang keahlian dan minat yang tepat berdasarkan data-data akademik mahasiswa salah satunya seperti IPK (indeks prestasi kumulatif) yang diperoleh mahasiswa. Menurut (Bora, 2014) menyatakan bahwa "*clustering* membagi dataset ke dalam kelompok yang berbeda, proses pengelompokan objek ke dalam kelompok yang objek dalam kelompok yang sama memiliki kesamaan dan objek dalam kelompok yang berbeda tidak memiliki kemiripan". Data yang digunakan untuk proses analisis data mining adalah data *warehouse* yang diperoleh dari proses awal dalam tahap data mining. Data *warehouse* berisi data IPK, nilai mata kuliah yang terkait dengan bidang keahlian dan konsentrasi pilihan subjek yang diambil oleh mahasiswa. Oleh karena itu, teknologi data mining dapat digunakan untuk mengelompokkan bidang keahlian mahasiswa berdasarkan IPK, nilai mata kuliah yang terkait dengan bidang keahlian dan minat mahasiswa sehingga dapat memberikan gambaran kepada mahasiswa untuk memilih bidang keahlian yang sesuai dengan kemampuan mereka.

Dengan adanya penelitian ini bagi Program Studi Teknik Informatika merupakan salah satu upaya untuk membantu memberikan rekomendasi kepada mahasiswa dalam memilih bidang keahlian yang sesuai dengan kemampuan mereka. Penelitian ini juga memberikan informasi penting terkait rencana strategis bagi Jurusan TIF berdasarkan interpretasi dari hasil penelitian dalam pengelolaan mahasiswa dengan memfasilitasi memfokuskan mahasiswa terhadap ilmu yang didapat dari matakuliah sebelumnya, agar keahlian mahasiswa dapat lebih terarah

dan juga bertujuan mengarahkan para mahasiswa dalam pemilihan topik penelitian tugas akhir. Dengan adanya tindakan ini maka akan dapat membantu tujuan dari Jurusan yaitu melahirkan sumber daya manusia Islami yang mampu menerapkan dan mengembangkan pengetahuan dan teknologi di bidang informatika, khususnya dalam mempersiapkan tenaga ahli yang terampil dan profesional di bidang teknologi informasi.

1.2. Rumusan Masalah

Berdasarkan uraian pada latar belakang, penelitian ini perumusan masalah yang diambil sebagai berikut: "Bagaimana merancang, membangun dan menerapkan data mining dan metode *Clustering* dengan algoritma *Fuzzy C-means* dalam pengelompokan mahasiswa berdasarkan data akademis mahasiswa sebagai rekomendasi untuk mengambil bidang keahlian yang sesuai dengan kemampuan dan minat mahasiswa dalam upaya mendukung rencana strategis untuk mengelola program studi Teknik Informatika”.

1.3. Batasan masalah

Keterbatasan masalah untuk penelitian ini meliputi:

- a. Penelitian ini diterapkan pada program studi S1 Teknik Informatika Fakultas Sains dan Teknologi UIN Suska Riau.
- b. Data Akademik yang digunakan adalah, Nama, Nim, IPK, Nilai Mata Kuliah.

1.4. Tujuan penelitian

Berdasarkan rumusan masalah, tujuan dari penelitian ini adalah untuk merancang, membangun dan menerapkan data mining dan metode *Clustering* dengan *algoritma Fuzzy C-Means* dalam pengelompokan mahasiswa berdasarkan data akademis mahasiswa untuk merekomendasikan pemilihan keahlian yang sesuai dengan kemampuan dan minat mahasiswa dalam upaya mendukung rencana strategis pengelolaan program studi Teknik Informatika.

1.5. Manfaat Penelitian

Manfaat dari penelitian ini adalah:

- a. Bagi penulis, penelitian ini berguna untuk menambah wawasan ke dalam data mining Clustering method dengan Fuzzy C-means algorithm dan aplikasinya.
- b. Untuk program studi Teknik Informatika, penelitian ini adalah salah satu upaya untuk membantu mahasiswa memilih bidang keahlian mereka dalam upaya mendukung rencana strategis untuk pengelolaan program studi.
- c. Hasil penelitian ini juga dapat digunakan sebagai referensi untuk penelitian lebih lanjut.

1.6. Sistematika Penulisan

Sistematika penulisan yang digunakan dalam laporan penelitian ini dibagi menjadi lima bab, yaitu:

- a. **BAB I PENDAHULUAN**

Bab ini berisi latar belakang, perumusan masalah, batasan masalah, manfaat penelitian, tujuan penelitian dan sistematika penulisan.

- b. **BAB II TINJAUAN LITERATUR**

Bab ini menjelaskan studi sebelumnya (tinjauan pustaka) dan teori yang mendukung penelitian (landasan teori).

- c. **BAB III METODOLOGI PENELITIAN**

Bab ini berisi metode dan langkah yang akan digunakan dalam penelitian.

- d. **BAB IV PEMBAHASAN**

Bab ini berisi hasil dan diskusi dari penelitian yang dilakukan. Diskusi ini berisi analisis data, langkah penyelesaian, desain dan analisis proses dan hasil.

- e. **BAB V KESIMPULAN DAN SARAN**

Bab ini berisi kesimpulan dari hasil penelitian yang telah dilakukan dan saran untuk pengembangan penelitian lebih lanjut.

BAB II

TINJAUAN PUSTAKA

2.1. Data Mining

Data mining adalah serangkaian proses untuk mengeksplorasi nilai tambah dalam bentuk informasi yang belum diketahui secara manual dari database. Informasi yang dihasilkan diperoleh dengan mengekstraksi dan mengenali pola penting atau menarik dari data yang terdapat dalam database (Hastuti, 2012). Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin untuk mengekstrak dan mengidentifikasi informasi yang berguna dan pengetahuan terkait dari berbagai database besar (Larose, 2006).

Data mining adalah penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari data yang sangat besar (Davies, 2004). Menurut Pramudiono (2007) data mining juga dapat disebut sebagai rangkaian proses untuk mengeksplorasi nilai tambah dalam bentuk pengetahuan yang belum diketahui secara manual dari kumpulan data. Menurut Davies (2004), data mining memiliki beberapa karakteristik, yaitu:

- a. Data mining terkait dengan penemuan sesuatu yang tersembunyi dan pola data tertentu yang tidak diketahui sebelumnya.
- b. Data mining biasanya menggunakan data yang sangat besar yang digunakan untuk membuat hasil lebih dapat diandalkan.
- c. Data mining berguna untuk membuat keputusan penting, terutama dalam strategi

Data mining juga disebut *Knowledge Discovery in Database (KDD)*, yang merupakan kegiatan yang mencakup pengumpulan, menggunakan data untuk menemukan pesan, pola dan hubungan dalam kumpulan data besar. Hasil data mining dapat digunakan untuk membuat keputusan. Diharapkan bahwa menggunakan data mining dapat memberikan pengetahuan dan informasi yang

sebelumnya tersembunyi di gudang data sehingga dapat menjadi informasi dan pengetahuan yang menarik dan berguna.

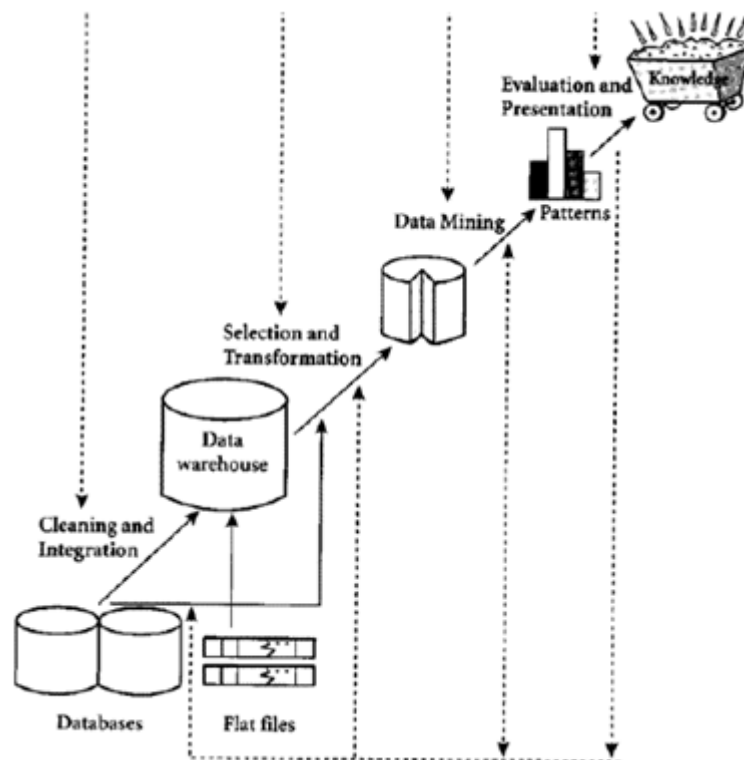
Menurut Gartner Group, data mining adalah proses menemukan hubungan, pola, dan kecenderungan yang bermakna dengan memeriksa dalam kumpulan besar data yang disimpan dalam penyimpanan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika. Data mining bukanlah bidang yang sepenuhnya baru. Salah satu kesulitan dalam mendefinisikan data mining adalah fakta bahwa data mining mewarisi banyak aspek dan teknik dari bidang sains yang mapan terlebih dahulu. Mulai dari beberapa disiplin ilmu, data mining bertujuan untuk meningkatkan teknik tradisional sehingga mereka dapat menangani:

- a. Jumlah data yang sangat besar
- b. Dimensi data yang tinggi
- c. Data yang heterogen dan berbeda bersifat

Menurut para ahli, data mining adalah analisis pengamatan sejumlah besar data untuk menemukan hubungan yang tidak diketahui sebelumnya dan metode baru untuk meringkas data agar mudah dipahami dan kegunaannya bagi pemilik data. Data yang ada tidak dapat diproses secara langsung menggunakan sistem data mining. Data-data ini harus dipersiapkan sebelumnya sehingga hasilnya dapat dimaksimalkan, dan waktu perhitungannya minimal. Proses persiapan data ini saja dapat mencapai 60% dari keseluruhan proses dalam data mining. Tahapan yang harus dilalui dalam proses data mining meliputi:

- a. *Data cleaning*, untuk membersihkan data dari noise data dan data yang tidak konsisten.
- b. *Data integration*, mengkombinasikan atau mengintegrasikan beberapa sumber data.
- c. *Data selection*, mengambil data-data yang relevan dari database untuk dianalisis.
- d. *Data transformation*, mentransformasikan data *summary* ataupun operasi agregasi.

- e. Data mining, merupakan proses yang esensial dimana metode digunakan untuk mengekstrak pola data yang tersembunyi.
- f. *Pattern evaluation*, untuk mengidentifikasi pola sehingga merepresentasikan pengetahuan berdasarkan nilai-nilai yang menarik
- g. *Knowledge presentation*, dimana teknik representasi dan visualisasi data digunakan untuk mempresentasikan pengetahuan yang didapat kepada user.



Gambar 2.1 Tahap-tahap Data Mining

2.1.1 Teknik data mining

Ada banyak jenis teknik analisis yang dapat diklasifikasikan dalam data mining. Namun ada tiga teknik data mining yang populer, yaitu:

1. *Association Rule Mining*

Association Rule Mining adalah teknik penambangan untuk menemukan aturan asosiatif antara kombinasi atribut. Contoh aturan asosiatif dari analisis pembelian di supermarket diketahui seberapa besar kemungkinan bagi

pelanggan untuk membeli roti bersama dengan susu. Dengan pengetahuan ini, pemilik supermarket dapat mengatur penempatan barang-barang mereka atau merancang strategi pemasaran menggunakan kupon diskon untuk kombinasi barang tertentu.

2. Klasifikasi.

Klasifikasi adalah proses menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui. Model itu sendiri dapat menjadi aturan "if-then", dalam bentuk pohon keputusan, rumus matematika atau jaringan saraf. Proses klasifikasi biasanya dibagi menjadi dua fase: pembelajaran dan pengujian. Dalam fase pembelajaran, beberapa data yang telah dikenal untuk kelas data diumpangkan untuk membentuk model perkiraan. Kemudian pada tahap uji coba model yang telah terbentuk diuji dengan beberapa data lain untuk menentukan keakuratan model. Jika akurasi cukup, model ini dapat digunakan untuk prediksi kelas data yang tidak diketahui.

3. *Clustering*.

Tidak seperti association rule mining dan klasifikasi di mana kelas data telah ditentukan sebelumnya, pengelompokan melakukan pengelompokan data tanpa berdasarkan pada kelas data spesifik. Bahkan pengelompokan dapat digunakan untuk memberi label pada kelas data tidak dikenal. Oleh karena itu pengelompokan sering diklasifikasikan sebagai metode pembelajaran yang tidak diawasi. Prinsip pengelompokan adalah untuk memaksimalkan kesamaan antara anggota satu kelas dan meminimalkan kesamaan antar kelompok. Clustering dapat dilakukan pada data yang memiliki beberapa atribut yang dipetakan sebagai ruang multidimensional.

2.2. Metode *Clustering Algoritma Fuzzy C-Means*

Logika fuzzy pertama kali dikembangkan oleh Lotfi A. Zadeh, seorang ilmuwan Amerika Iran dari *California University* di Berkeley, melalui tulisannya pada tahun 1965. (Munir, R. 2005). Fuzzy diartikan sebagai bahasa yang samar atau

tidak jelas. Nilai bisa benar atau salah pada saat yang sama. Dalam fuzzy diketahui bahwa derajat keanggotaan memiliki rentang 0 (nol) hingga 1 (satu). Berbeda dengan set ketat yang memiliki nilai 1 atau 0 (ya atau tidak). Logika fuzzy adalah cara yang tepat untuk memetakan ruang input ke dalam ruang output, memiliki nilai yang berkesinambungan. Fuzzy diungkapkan dalam tingkat keanggotaan dan tingkat kebenaran. Oleh karena itu sesuatu dapat dikatakan sebagian benar dan sebagian salah pada waktu yang sama (Kusumadewi, 2003).

2.2.1. Clustering

Clustering adalah metode pengelompokan berdasarkan ukuran kedekatan (*similarity*). Clustering berbeda dari grup, jika grup berarti grup yang sama, kondisi atau bukan pasti bukan grup. Tetapi jika cluster tidak harus sama, pengelompokan didasarkan pada kedekatan karakteristik sampel yang ada. Berbagai metode pengelompokan:

1. Berdasarkan metode statistik.

Sebuah. Metode pengelompokan hirarkis: dalam hal jumlah kelompok belum ditentukan sebelumnya, misalnya data dari survei kuesioner. Jenis metode ini:

1. Berbasis Metode Statistikk

a. Hierarchical clustering method : pada kasus untuk jumlah kelompok belum ditentukan terlebih dulu, contoh data-data hasil survey kuisisioner. Macam-metode jenis ini: Single Linkage, Complete Linkage, Average Linkage dll.

b. Non Hierarchical clustering method: Jumlah kelompok telah ditentukan terlebih dulu. Metode yang digunakan : K-Means.

2. Berbasis Fuzzy : Fuzzy C-Means

3. Berbasis Neural Network : Kohonen SOM, LVQ

4. Metode lain untuk optimasi centroid atau lebar cluster : Genetik Algoritma (GA)

2.2.2. Fuzzy C-Means

Fuzzy clustering adalah proses penentuan tingkat keanggotaan, dan kemudian menggunakannya dengan memasukkannya ke dalam elemen data ke

dalam satu *cluster* atau lebih. Ini akan memberikan informasi kesamaan untuk setiap objek. Salah satu dari banyak algoritma pengelompokan fuzzy yang digunakan adalah algoritma *fuzzy clustering c means*. Vektor pengelompokan fuzzy, $V = \{v_1, v_2, v_3, \dots, v_c\}$, adalah fungsi obyektif yang ditentukan oleh tingkat keanggotaan data X_j dan pusat kluster V_j . Algoritma fuzzy clustering c berarti membagi data yang tersedia dari masing-masing elemen data hingga kemudian memasukkannya ke dalam bagian-bagian kumpulan cluster yang dipengaruhi oleh beberapa kriteria yang diberikan. Berikan satu set data terbatas. $X = \{x_1, \dots, x_n\}$ dan pusat data.

$$J_m(X, U, V) = \sum_{j=1}^n \sum_{i=1}^c (\mu_{ij})^m d^2(X_j, V_i) \dots \dots \dots (1)$$

Dimana μ_{ij} adalah tingkat keanggotaan X_j dan pusat dari cluster adalah bagian dari keanggotaan matriks $[\mu_{ij}]$. d^2 adalah akar dari jarak Euclidean dan m adalah parameter fuzzy yang rata-rata tingkat keburaman dari setiap data derajat keanggotaan tidak lebih besar dari 1.0 Ravichandran (2009). Output Fuzzy C-Means adalah deretan pusat cluster dan beberapa derajat keanggotaan untuk setiap titik data. Informasi ini dapat digunakan untuk membangun sistem inferensi fuzzy.

2.2.3. Algoritma Fuzzy Clustering Means (FCM)

Algoritma Fuzzy C-Means adalah sebagai berikut:

1. Input data yang akan dicluster X , berupa matriks berukuran $n \times m$ (n =jumlah sample data, m =atribut setiap data). X_{ij} =data sample ke- i ($i=1,2,\dots,n$), atribut ke- j ($j=1,2,\dots,m$).
2. Tentukan :
 - a. Jumlah cluster = c
 - b. Pangkat = w
 - c. Maksimum iterasi = $MaxIter$
 - d. Error terkecil yang diharapkan = ζ
 - e. Fungsi obyektif awal = $P_0 = 0$
 - f. Iterasi awal = $t =$

3. Bangkitkan nilai acak μ_{ik} , $i=1,2,\dots,n$; $k=1,2,\dots,c$ sebagai elemen-elemen matriks partisi awal μ_{ik} . μ_{ik} adalah derajat keanggotaan yang merujuk pada seberapa besar kemungkinan suatu data bisa menjadi anggota ke dalam suatu cluster. Posisi dan nilai matriks dibangun secara random. Dimana nilai keanggotaan terletak pada interval 0 sampai dengan 1. Pada posisi awal matriks partisi U masih belum akurat begitu juga pusat clusternya. Sehingga kecendrungan data untuk masuk suatu cluster juga belum akurat.

$$Q_i = \sum_{k=1}^c \mu_{ik} \dots\dots\dots (2)$$

4. Hitung pusat *Cluster* ke- k : V_{kj} , dengan $k=1,2,\dots,c$ dan $j=1,2,\dots,m$. dimana X_{ij} adalah variabel fuzzy yang digunakan dan w adalah bobot.

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w * X_{ij})}{\sum_{i=1}^n (\mu_{ik})^w} \dots\dots\dots (4)$$

Fungsi objektif digunakan sebagai syarat perulangan untuk mendapatkan pusat cluster yang tepat. Sehingga diperoleh kecendrungan data untuk masuk ke *cluster* mana pada *step* akhir.

5. Hitung fungsi obyektif pada iterasi ke- t , P_t

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^w \right) \dots\dots\dots (5)$$

Perhitungan fungsi objektif P_t dimana nilai variabel fuzzy X_{ij} di kurang dengan dengan pusat cluster V_{kj} kemudian hasil pengurangannya di

kuadratkan lalu masing-masing hasil kuadrat di jumlahkan untuk dikali dengan kuadrat dari derajat keanggotaan μ_{ik} untuk tiap *cluster*. Setelah itu jumlahkan semua nilai di semua *cluster* untuk mendapatkan fungsi objektif P_t .

6. Hitung perubahan matriks partisi:

$$\mu_{ik} = \frac{\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{-1}{w-1}}}{\sum_{k=1}^c \left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{-1}{w-1}}} \dots\dots\dots(6)$$

dengan: $i=1,2,\dots,n$ dan $k=1,2,\dots,c$.

Untuk menemukan perubahan matriks partisi μ_{ik} , nilai dari variabel fuzzy X_{ij} dikurangi lagi ke pusat cluster V_{kj} , dan juga dikuadratkan. Maka jumlah tersebut kemudian dinaikkan oleh $-1 / (w-1)$ dengan berat, $w = 2$ hasilnya adalah bahwa setiap data dinaikkan sebesar -1 . Setelah proses perhitungan dilakukan, menormalkan semua data gelar keanggotaan baru dengan menambahkan derajat keanggotaan baru $k = 1, \dots, c$, hasilnya kemudian dibagi dengan gelar keanggotaan baru. Proses ini dilakukan sehingga tingkat keanggotaan baru memiliki rentang antara 0 dan tidak lebih dari 1

7. Periksa kondisi berhenti: a) jika $(| P_t - P_{t-1} | < \xi)$ atau $(t > \maxIter)$ kemudian berhenti. B) jika tidak, $t = t + 1$, ulangi langkah 4.

2.3 Penelitian yang Relevan

Berikut ini adalah beberapa studi yang relevan dengan penelitian yang dibangun, yaitu:

1. Penelitian Yadav & Ahmed, (2012) berjudul "Evaluasi Kinerja Akademik Menggunakan *Fuzzy C-Means*", dengan jumlah data adalah 20 data mahasiswa untuk menganalisis dan mengetahui pemodelan kinerja akademik dan

meningkatkan kualitas mahasiswa dan guru di dunia pendidikan. Hasil penelitian menunjukkan bahwa *Fuzzy C-Means* menjadi model terbaik untuk pemodelan kinerja akademik yang berfungsi sebagai patokan dalam memantau perkembangan pemodelan mahasiswa di dunia pendidikan. Diharapkan di masa depan *Fuzzy C-Means* dapat dikombinasikan dengan metode lain seperti *Neuro Dynamic Fuzzy Expert System* untuk mengevaluasi kinerja mahasiswa dan guru dan mengembangkan sistem pembelajaran yang dapat mendukung proses pembelajaran jarak jauh berbasis internet.

2. Penelitian yang dilakukan oleh Varghese et al., (2011), Inyang & Joshua, (2013) berjudul "Data Mahasiswa *Clustering* untuk Karakterisasi Pola Kinerja" dan "*Fuzzy Clustering* Mahasiswa Data Repositori untuk At-Risiko Identifikasi dan Pemantauan Mahasiswa", menggunakan Teknik data mining dengan jumlah data 400 mahasiswa dilakukan oleh Varghese et al dan 496 data mahasiswa yang dilakukan oleh Inyang & Joshua untuk mengekstrak pola sehingga pengambilan keputusan dari database besar dapat dilakukan. Model *Fuzzy C-Means* dapat membantu dalam mencapai proses pemantauan kemajuan mahasiswa dan mengidentifikasi fitur apa yang mempengaruhi kinerja mahasiswa sehingga dapat membantu pendidik mengetahui tingkat kemampuan setiap mahasiswa.

3. Penelitian yang dilakukan oleh Agbonifo & Catherine (2013) berjudul "*Fuzzy C-Means Clustering Model* untuk Identifikasi Preferensi Belajar Mahasiswa Lingkungan Online" mengadopsi metode gaya belajar Honey dan Mumford dengan menggunakan teknik pengelompokan *Fuzzy C-Means* untuk menentukan preferensi pembelajaran dari seorang mahasiswa dengan jumlah data yang diuji sebanyak 50 data mahasiswa sehingga membantu instruktur dan pendidik dalam proses pengajaran.

4. Penelitian Yadav & Singh (2012) menggunakan metode *Fuzzy C-Means* dengan data mahasiswa tes 20 mahasiswa dapat memodelkan prestasi akademik mahasiswa dalam lingkungan pendidikan yang digunakan sebagai patokan / dasar dalam memantau perkembangan mahasiswa di lingkungan pendidikan.

5. Penelitian yang dilakukan oleh Setiyono & Isnanto (2008) membahas pendekatan fuzzy clustering dengan membagi kelas berdasarkan nilai prestasi mahasiswa dalam

mata pelajaran yang merupakan syarat untuk mengambil mata pelajaran baru dimana jumlah data yang akan dikelompokkan adalah 70 mahasiswa. data. Hasil penelitian menunjukkan bahwa Fuzzy C-Means lebih cocok untuk alokasi mahasiswa ke dalam beberapa kelas berdasarkan masukan dari jumlah cluster yang diinginkan.

6. Munandar et al., (2013) menganalisis 126 data mahasiswa sesuai dengan bobot subjek tertentu dengan menggunakan konsep Fuzzy C-Means sehingga mereka mampu memberikan dukungan keputusan tentang pengelompokan data nilai mahasiswa untuk menentukan klasifikasi konsentrasi di mana seorang mahasiswa harus dimasukkan.

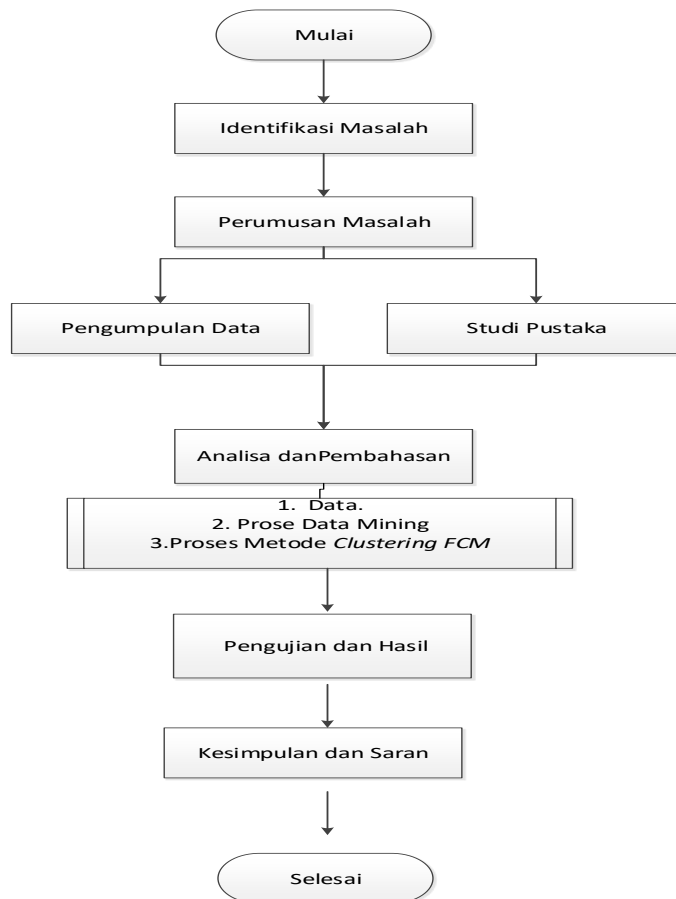
7. Penelitian yang dilakukan oleh Redjeki, Sri. (2017) melakukan pemodelan untuk mengklasifikasikan prestasi dosen menggunakan Fuzzy C-Means. Objek penelitian ini dilakukan di salah satu universitas swasta di Yogyakarta. Parameter input yang digunakan dalam metode FCM adalah mengajar kehadiran, penyelesaian nilai akhir dan hasil kuesioner dari mahasiswa. Masukan dari 3 variabel ini disebut variabel independen, sedangkan nilai indeks prestasi dosen (IPAD) adalah variabel dependen. Ada 3 kelompok pencapaian yang ditentukan dalam penelitian ini, yaitu baik, sedang dan kurang. Data yang digunakan adalah 59 dosen

BAB III

METODOLOGI PENELITIAN

3.1 Kerangka Pikir Penelitian

Pada bab ini akan dijelaskan tahap – tahap yang dilakukan dalam melakukan penelitian. Tahapan penelitian berguna agar pelaksanaan penelitian dapat berjalan dengan baik dan sistematis serta memenuhi tujuan yang diinginkan. Susunan tahapan penelitian dijelaskan pada gambar 3.1 *flowchart* dibawah ini :



Gambar 3.1 Flowchart Metodologi Penelitian

Dari *flowchart* Metode Penelitian pada gambar 3.1 terdapat beberapa proses penelitian diantaranya adalah :

1. Identifikasi Masalah

Pada tahapan ini dilakukan identifikasi terhadap masalah yang ditemukan pada saat penelusuran dalam kasus pengelompokan bidang keahlian mahasiswa dengan teknik data mining dan metode clustering dengan Algoritma Fuzzy C-Means.

2. Perumusan Masalah

Dari identifikasi masalah yang dilakukan maka ditemukan rumusan masalah Bagaimana bentuk penerapan data mining dan metode clustering dengan Algoritma Fuzzy C-Means dalam kasus pengelompokan mahasiswa berdasarkan data akademik mahasiswa untuk rekomendasi pemilihan bidang keahlian dalam upaya mendukung rencana strategis pengelolaan program studi Teknik Informatika.

3. Pengumpulan Data

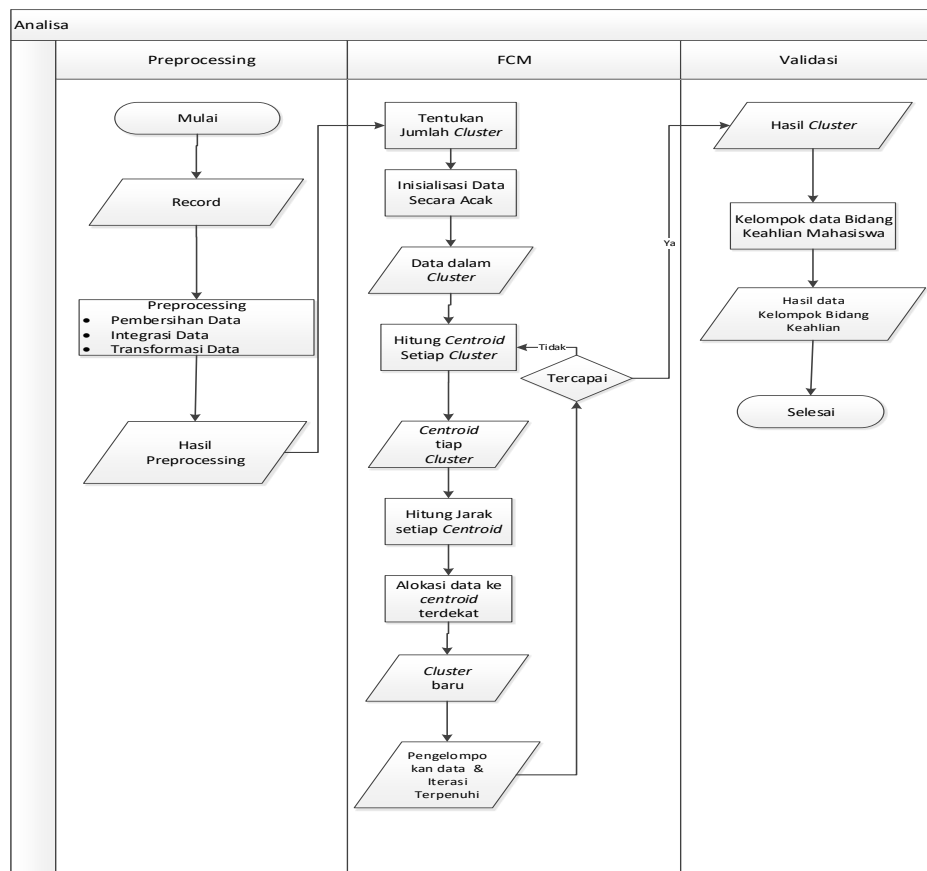
Pengumpulan data merupakan tahap ketiga dalam metodologi penelitian. Pada tahap pengumpulan data ditentukan jenis data apa yang digunakan untuk penelitian, bahan-bahan yang akan dibutuhkan selama penelitian. Tahap pengumpulan data merupakan tahap yang penting dari suatu penelitian. Data yang digunakan untuk penelitian harus benar-benar akurat dan jelas sumbernya. Pada tahap pengumpulan data ini, data-data di peroleh dari data akademis mahasiswa tahun 2014/2015.

4. Studi Pustaka

Tahapan ini merupakan tahapan awal dalam melakukan penelitian ini, studi pustaka ini di lakukan untuk melengkapi pengetahuan dasar dan teori-teori yang di gunakan dalam penelitian ini, informasi mengenai penelitian ini bisa ditemukan pada buku-buku ataupun jurnal, serta tentang penggunaan metode *clustering* algoritma fuzzy *C-means*.

5. Analisa dan Pembahasan

Pada tahapan ini gambaran tentang penelitian yang akan dilakukan telah diketahui. Analisa dilakukan terhadap data-data di peroleh dari data akademis mahasiswa Prodi Teknik Informatika tahun 2014/2015 dan di proses menggunakan data mining dan Algoritma *Fuzzy C-Means*. Tahapan ini merupakan tahapan proses analisa data mining terbagi atas beberapa bagian yaitu tahap *preprocessing* merupakan tahap penyeleksian data dan pembersihan *noisy* yang bertujuan sebagai data yang valid untuk dijadikan dan digunakan dalam penelitian ini. Pada tahapannya perubahan data data sehingga menghasilkan atribut *dataset* yang baik untuk diteliti dengan tujuan untuk pemahaman terhadap isi *record* pada data akademis dan melakukan beberapa tahap dari *preprocessing* dan *cleaning*. Berikut merupakan analisa data bagaimana proses awal data *inputan* dan data *output* dari sistem yang akan dibangun pada gambar 3.2 dibawah ini:



Gambar 3.2 Analisa Data

Langkah-langkah *Fuzzy C-Means* :

1. Input data yang akan dicluster X , berupa matriks berukuran $n \times m$ (n =jumlah sample data, m =atribut setiap data). X_{ij} =data sample ke- i ($i=1,2,\dots,n$), atribut ke- j ($j=1,2,\dots,m$).
2. Tentukan :
 - a. Jumlah cluster
 - b. Pangkat
 - c. Maksimum iterasi
 - d. Error terkecil yang diharapkan
 - e. Fungsi obyektif awal
 - f. Iterasi awal
3. Bangkitkan nilai acak μ_{ik} , $i=1,2,\dots,n$; $k=1,2,\dots,c$ sebagai elemen-elemen matriks partisi awal μ_{ik} . μ_{ik} adalah derajat keanggotaan yang merujuk pada seberapa besar kemungkinan suatu data bisa menjadi anggota ke dalam suatu cluster.
4. Hitung pusat *Cluster* ke- k : V_{kj} , dengan $k=1,2,\dots,c$ dan $j=1,2,\dots,m$. dimana X_{ij} adalah variabel fuzzy yang digunakan dan w adalah bobot
5. Hitung fungsi obyektif pada iterasi ke- t , P_t .
6. Hitung perubahan matriks partisi.
7. Periksa kondisi berhenti: a) jika $(|P_t - P_{t-1}| < \xi)$ atau $(t > \maxIter)$ kemudian berhenti. B) jika tidak, $t = t + 1$, ulangi langkah 4

Desain terhadap kebutuhan analisa dilakukan dengan menggunakan menggunakan konsep UML meliputi *usecase*, *sequence diagram*, *class diagram* dan *activity diagram*, mulai perancangan interface sampai kepada perancangan database dan pengkodean sistem. Tahap desain berdasarkan tahap analisis yang telah dilakukan sebelumnya yang terbagi menjadi 3 aspek, yaitu:

1. Desain Subsistem Data

Tahap ini merupakan desain dari analisis subsistem manajemen data sebelumnya. Tahapan ini berupa perancangan *usecase*, *sequence diagram*, *class*

diagram dan activity diagram. Dan berikutnya akan dibuat perancangan kamus data.

2. Desain Subsistem Model

Tahap ini merupakan hasil dari analisis model yaitu metode yang digunakan dalam pembuatan sistem tersebut. Pada tahap ini akan dibuat suatu desain model berupa *flowchart* sistem dan flowchart perhitungan metode FCM.

3. Desain Subsistem Dialog

Tahap ini merupakan hasil dari analisis subsistem manajemen dialog. Tahapan ini akan menghasilkan sebuah perancangan struktur menu dan perancangan antarmuka (*interface*) sistem.

Tahap pengembangan adalah proses mewujudkan *blue-print* atau desain menjadi kenyataan. Setelah desain diperlukan pengembangan suatu perangkat lunak, dan proses pembuatan mengacu pada tahap desain

6. Pengujian dan Hasil.

Setelah dilakukan analisa dan pembahasan maka akan dilakukan pengujian dan menampilkan hasil dari penelitian yang telah dibuat. Tahap pengujian diperlukan sebagai ukuran bahwa penelitian dapat dijalankan sesuai dengan tujuan. Pada pengujian ini ada beberapa hal yang dijadikan acuan, yaitu :

1. Validasi, yaitu melakukan proses validasi hasil penelitian kepada validator/pakar.
2. Pengujian terhadap calon *user* dengan menggunakan *User Acceptence Test* dalam hal ini Prodi Teknik Informatika.

7. Kesimpulan dan Saran

Tahap ini berisikan tentang kesimpulan penelitian ini dan hasil yang didapatkan. Tahap ini juga berisikan hal yang disimpulkan dan disarankan penulis bagi pembaca untuk melakukan pengembangan terhadap penelitian ini kedepannya.

BAB IV

ANALISA DAN PEMBAHASAN

Pada bab ini membahas tentang analisa data dan perancangan sistem yang akan dibangun dengan menerapkan *Clustering* dengan algoritma *Fuzzy C-Means*. Analisa merupakan kegiatan dalam penentuan tindakan atau keputusan yang akan menentukan dari rancangan sistem. Sementara tahap perancangan merupakan kegiatan menentukan rincian sistem yang akan dibuat.

4.1 Analisa Sistem Lama

Pada saat ini proses menentukan kebidangan mahasiswa di jurusan Teknik informatika uin suska riau dilakukan secara manual dengan menggunakan rekomendasi dari pembimbing akademik mahasiswa yang bersangkutan. Selain itu pembimbing akademik juga memberikan pertimbangan langsung terhadap minat mahasiswa tersebut. Penilaian kebidangan yang ada saat ini secara garis besar dibagi menjadi 2 macam yaitu kebidangan teknologi informasi dan ilmu komputer. Namun yang menjadi sebuah permasalahan adalah dosen pembimbing akademik tidak dapat mempertimbangkan keseluruhan dari nilai mahasiswa di setiap semesternya. Selain itu penentuan kebidangan akademik mahasiswa tersebut tidak dapat dijadikan pedoman karena bersifat subjektif, tidak berdasarkan fakta nilai yang dimiliki oleh mahasiswa tersebut. Hal tersebut terjadi karna banyaknya jumlah mahasiswa yang dimiliki oleh setiap dosen pembimbing. Untuk mempermudah penentuan kebidangan mahasiswa dibutuhkan sebuah sistem atau aplikasi yang mampu menentukan pembagian kelompok dari semua mahasiswa berdasarkan kemiripan sifat dari setiap variabel kebidangan.

4.2 Analisa Sistem Baru

Penentuan bidang keahlian mahasiswa di sistem baru, dibangun menggunakan metode *fuzzy c-means* dengan bahasa pemrograman berbasis web(PHP). Hal ini bertujuan untuk memberikan kemudahan bagi setiap dosen bersangkutan dalam menentukan kebidangan mahasiswanya di waktu dan tempat

yang *flexible*. Metode pengembangan fungsional aplikasi dirancang menggunakan konsep *Unified Modeling Language (UML)*. Konsep *UML* yang digunakan adalah *usecase*, *sequence diagram*, *activity diagram* dan *class diagram* dan metode pengembangan aplikasi menggunakan konsep Object Oriented Programming (OOP), dengan output yang diberikan aplikasi adalah pengelompokan mahasiswa menjadi 2 kluster yang sesuai dengan tingkat kemiripan pola semua data nilai mahasiswa yang tersedia di dalam database aplikasi. Dengan adanya sistem ini, diharapkan akan mempermudah dosen pembimbing akademik dan mahasiswa untuk memilih kebidangan yang direkomendasikan sesuai dengan pencapaian yang dimiliki oleh mahasiswa tersebut.

Adapun untuk menganalisis data dalam penerapan data mining ini menggunakan proses tahapan *knowledge discovery in databases (KDD)* yang terdiri dari Database, Data Cleaning, Data integration, Task relevan data, Data transformation, Data mining, Pattern evolution, knowledge :

1. Database

Koleksi data yang saling berhubungan untuk dipergunakan secara bersama kemudian dirancang untuk memenuhi kebutuhan informasi organisasi

2. Data Cleaning

Pada umumnya, data yang diperoleh, baik dari database suatu perusahaan maupun eksperimen, memiliki isian-isian yang tidak sempurna seperti data yang hilang, data yang tidak valid atau hanya sekedar salah ketik. Data yang tidak relevan itu juga lebih baik dibuang karena keberadaannya bisa mengurangi mutu atau akurasi dari hasil data mining nantinya. Pembersihan data juga akan mempengaruhi performansi dari system data mining karena data yang akan ditangani akan berkurang jumlah dan kompleksitasnya.

3. Data integration

Integrasi data dilakukan pada atribut-attribut yang mengidentifikasi entitas-entitas yang unik seperti atribut nama, nomor pegawai, tempat lahir, agama dan lain sebagainya. Integrasi data perlu dilakukan secara cermat karena kesalahan pada integrasi data bisa menghasilkan hasil data yang menyimpang dan bahkan menyesatkan pengambilan aksi nantinya.

4. Task relevan data

Setelah semua sumber data digabung atau diintegrasikan menjadi satu keseluruhan database, maka tahap selanjutnya adalah melakukan task relevant data. Pada tahap ini melakukan relevansi atribut dari data yang relevant atau yang sesuai dengan target atau output yang akan dihasilkan.

5. Data transformation

Coding adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Proses coding dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam database.

6. Data mining

Data mining adalah proses mengeksplorasi dan menganalisa data dalam jumlah yang besar yang bertujuan untuk menemukan suatu pola atau informasi yang menarik dari data yang tersimpan dalam jumlah yang besar dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan. Tahap ini merupakan inti dari tahapan KDD yang dilakukan untuk menganalisis data yang telah dibersihkan.

7. Pattern evaluation

Dalam tahap ini, merupakan hasil dari teknik data mining berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada memang tercapai. Bila ternyata hasil yang diperoleh tidak sesuai dengan hipotesa, ada beberapa alternatif yang dapat diambil seperti menjadikannya umpan balik untuk memperbaiki data mining lain yang lebih sesuai, atau menerima hasilnya sebagai suatu hasil yang diluar dugaan yang mungkin bermanfaat.

8. Knowledge

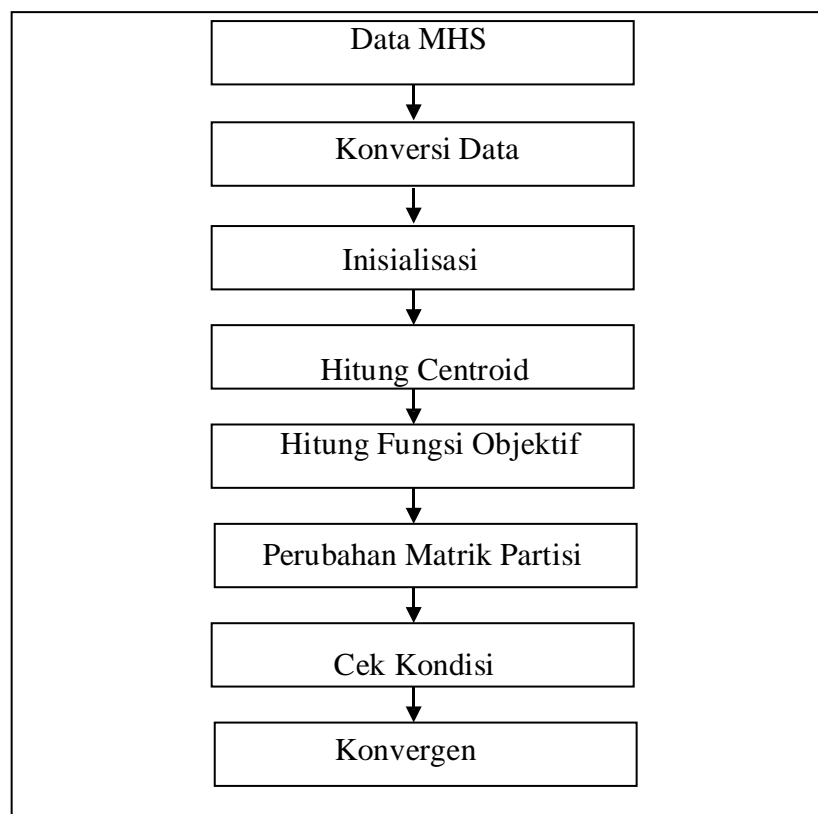
Tahap terakhir dari proses data mining adalah bagaimana memformulasikan keputusan atau aksi dari hasil analisis yang didapat.

4.2.1 Analisa Kebutuhan Data

Data yang digunakan pada penelitian ini adalah data sekunder. Pengumpulan data dilakukan di jurusan Teknik informatika Universitas Islam Negeri Sultan Syarif Kasim Riau. Data yang digunakan adalah data mata kuliah yang tersedia pada semester 1 sampai dengan semester 6. Selain data mata kuliah, data yang digunakan adalah data nilai dan tahun masuk mahasiswa mayang telah menyelesaikan semua mata kuliah sampai dengan semester 6.

4.2.1.1 Tahapan Proses Algoritma *Fuzzy C-Means*

Berdasarkan pada tahapan KDD berikutnya dimasukkan kedalam tahapan blok diagram, bagian ini akan membahas secara garis besar langkah-langkah penghitungan menggunakan metode algoritma *Fuzzy C-Means* yang dipaparkan pada gambar 4.1.



Gambar 4.1 Tahapan Algoritma Fuzzy C-Means

4.2.1.2 Uraian Algoritma Fuzzy C-Means

Setelah penjelasan langkah-langkah penghitungan secara garis besar bagan tahapan, bagian ini akan menguraikan secara detail masing-masing prosesnya, berikut adalah uraian gambar 4.1 :

1. Data Mata Kuliah Mahasiswa

Penelitian ini dilaksanakan pada Program Studi Teknik Informatika UIN Suska Riau dengan memanfaatkan data nilai setiap mata kuliah bidang Teknik Informatika strata satu (S-1). Data nilai mata kuliah yang digunakan sebagai pertimbangan diantaranya adalah Pengantar Teknologi Informasi dan Komunikasi, Dasar Pemrograman, Sistem Digital, Aljabar Linear, Matematika Diskrit, Organisasi Komputer, Algoritma dan Pemrograman, Teori Bahasa dan Otomata, Basis Data, Arsitektur Komputer, Struktur Data, Teknologi Web, Sistem Operasi, Sistem Basis Data, Strategi Algoritma, Interaksi Manusia dan Komputer, Pemrograman Bergerak, Pengembangan Aplikasi Berbasis Web, Jaringan Komputer, Sistem Informasi, Rekayasa Perangkat Lunak Berorientasi Objek, Keamanan Informasi, Disain Interaksi dan Antarmuka, Metode Numerik, Probabilitas dan Statistik, Kecerdasan Buatan, Manajemen Proyek Teknologi Informasi, Data Mining, Tata Kelola Teknologi Informasi yang nantinya dapat digunakan untuk mengetahui bidang keahlian yang sesuai yaitu bidang Teknologi Informasi dan Ilmu Komputer sesuai dengan kompetensi yang ada di Prodi Teknik Informatika UIN Suska Riau.

2. Konversi Data

Data transkrip nilai mahasiswa yang berupa huruf akan menyusahkan pada saat tahap pemrosesan, maka dari itu data transkrip nilai harus dikonversi kedalam angka terlebih dahulu.

Klasifikasi	Nilai Angka	Nilai MK	Bobot
Sangat Kurang	E	0-40	0,2
Kurang	D	41-55	0,4
Cukup	C	56-65	0,6
Baik	B	66-80	0,8
Sangat Baik	A	81-100	1

3. Inisialisasi

Pemberian nilai sembarang pada *pseudo-partition* (derajat keanggotaan setiap data pada setiap cluster) dalam jangkauan $[0,1]$, seperti persamaan 2.1.

Nilai awal :

○ Jumlah cluster	= c	= 2;
○ Pangkat	= w	= 2;
○ Maksimum iterasi	= MaxIter	= 100;
○ Error terkecil yang diharapkan	= ξ	= 10^{-2} ;
○ Fungsi obyektif awal	= P_0	= 0;
○ Iterasi awal	= t	= 1;

Keterangan

X_1 = Pengantar Teknologi Informasi dan Komunikasi

X_2 = Dasar Pemrograman

X_3 = Sistem Digital

X_4 = Aljabar Linear

X_5 = Matematika Diskrit

X_6 = Organisasi Komputer

X_7 = Algoritma dan Pemrograman

X_8 = Teori Bahasa dan Otomata

X_9 = Basis Data

X_{10} = Arsitektur Komputer

X_{11} = Struktur Data

X_{12} = Teknologi Web

X_{13} = Sistem Operasi

X_{14} = Sistem Basis Data

X_{15} = Strategi Algoritma

X_{16} = Interaksi Manusia dan Komputer

X_{17} = Pemrograman Bergerak

- X₁₈ = Pengembangan Aplikasi Berbasis Web
- X₁₉ = Jaringan Komputer
- X₂₀ = Sistem Informasi
- X₂₁ = RPLBO
- X₂₂ = Keamanan Informasi
- X₂₃ = Disain Interaksi dan Antarmuka
- X₂₄ = Metode Numerik
- X₂₅ = Probabilitas dan Statistik
- X₂₆ = Kecerdasan Buatan
- X₂₇ = Manajemen Proyek Teknologi Informasi
- X₂₈ = Data Mining
- X₂₉ = Tata Kelola Teknologi Informasi

4. Hitung Centroid

Penghitungan centroid dilakukan pada tiap-tiap cluster dengan menggunakan rumus persamaan 2.3 dimana n adalah jumlah data, w adalah bobot pangkat, μ_{jk} adalah nilai derajat keanggotaan data, dan X_{ij} adalah data yang dicluster.

MATRIKS PARTISI AWAL

	0.753	0.606
	0.671	0.306
	0.286	0.556
	0.893	0.898
	0.822	0.239
	0.742	0.222
	0.170	0.869
	0.098	0.498
U=	0.881	0.814
	0.151	0.378
	0.308	0.389

0.463	0.610
0.822	0.528
0.252	0.793
0.455	0.884
0.586	0.389
0.237	0.767
0.468	0.716
0.431	0.707
0.734	0.479

5. Hitung Fungsi Objektif

Selanjutnya dilakukan penghitungan fungsi objektif menggunakan rumus persamaan 2.4.

PENGHITUNGAN FUNGSI OBJEKTIF

L1	L2	L1+L2
588.536	484.8192	1073.355214
17753515	3527720	21281234.26
3227446	11693993	14921438.7
31394552	30433272	61827823.27
26589019	2156477	28745496.35
21675312	1866911	23542223.12
1138043	28542349	29680392.03
379783.8	9351183	9730966.366
30569634	25000737	55570370.45
896148.8	5390996	6287144.549
3733312	5713345	9446657.734
8444403	14019100	22463502.47
26536218	10525504	37061722.69
2501363	23722842	26224205.03
8146111	29552604	37698715.18

13512279	5719432	19231711.59
2216277	22225607	24441883.71
8613149	19348765	27961913.99
7307677	18870054	26177730.99
21205224	8677095	29882319.04

6. Perubahan Matrik Partisi

Menghitung perubahan matrik partisi digunakan rumus persamaan 2.5 dimana V adalah pusat *cluster* ke- k untuk atribut ke- j , X adalah data ke- i atribut ke- j , sedangkan μ adalah derajat keanggotaan untuk data sample ke- i pada *cluster* ke- k .

MATRIKS PARTISI U

L1	L2	LT=L1+L2
0.000963654	0.000757053	0.001720708
0.000000025	0.000000026	0.000000052
0.000000025	0.000000026	0.000000052
0.000000025	0.000000026	0.000000052
0.000000025	0.000000026	0.000000052
0.000000025	0.000000026	0.000000052
0.000000025	0.000000026	0.000000052
0.000000025	0.000000026	0.000000052
0.000000025	0.000000026	0.000000052
0.000000025	0.000000026	0.000000052
0.000000025	0.000000026	0.000000052
0.000000025	0.000000026	0.000000052
0.000000025	0.000000026	0.000000052
0.000000025	0.000000027	0.000000052
0.000000025	0.000000027	0.000000052
0.000000025	0.000000026	0.000000052
0.000000025	0.000000026	0.000000052
0.000000025	0.000000026	0.000000052

0.000000025	0.000000026	0.000000052
0.000000025	0.000000026	0.000000052
0.000000025	0.000000026	0.000000052
0.000000025	0.000000026	0.000000052

7. Cek Kondisi

Tahapan ini adalah tahap dilakukannya pengecekan kondisi berhenti pada nilai fungsi objektif dengan cara sebagai berikut :

- a. $|P_t - P_{t-1}| < \zeta$) atau ($t > \text{MaxIter}$) maka berhenti;
- b. Jika tidak, iterasi dinaikkan $t=t+1$, ulangi langkah ke-4.

Pengecekan kondisi dilakukan sesuai kebutuhan, karena semakin kecil selisih maka semakin tinggi akurasinya dan cluster yang didapat dinyatakan sudah convergen.

P1	=	512178524.9
P0	=	0
$ P1-P0 $		512178524.9

8. Konvergen

Tahap terakhir dari proses clustering yaitu mencari nilai terbesar dari derajat keanggotaan yang diperoleh dari iterasi terakhir. Derajat keanggotaan terbesar menunjukkan bahwa kecenderungan tertinggi lulusan untuk masuk menjadi anggota *cluster* tersebut.

Derajat keanggotaan tiap data pada setiap cluster dengan FCM

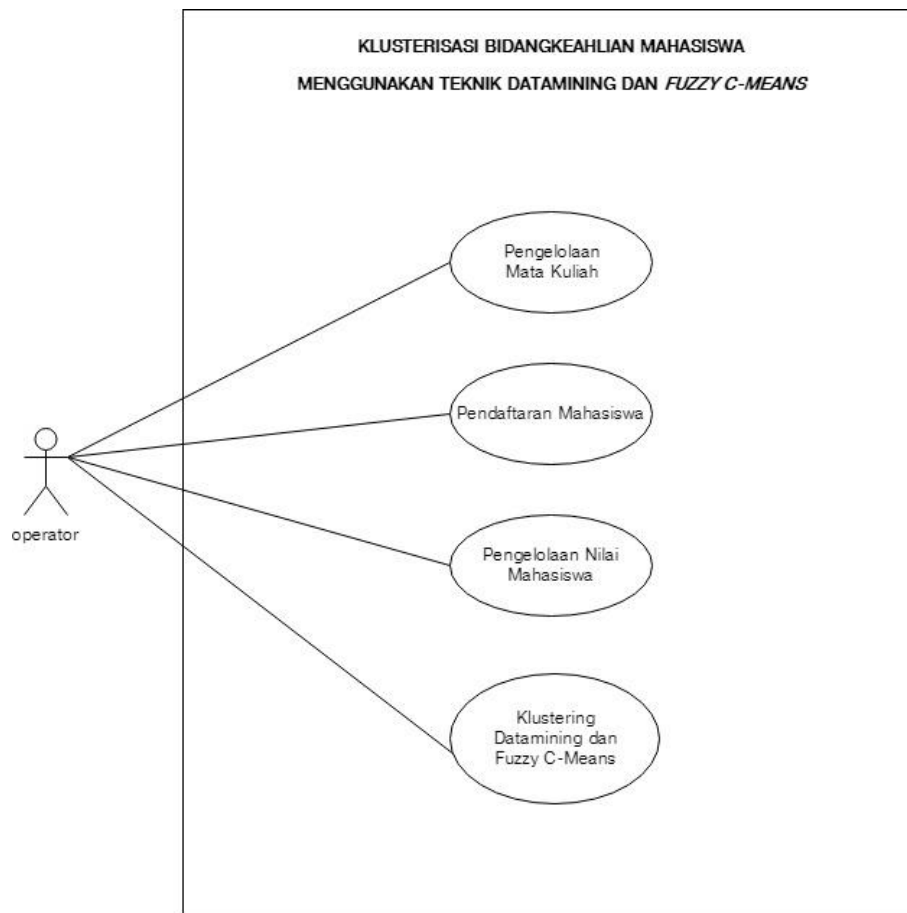
Data Ke	Derajat keanggotaan (m) data pada Cluster ke-			Data Cenderung Masuk ke Cluster dengan Derajat keanggotaan	
	1	2			
1	0.56	0.44		0.560033596	C1
2	0.49	0.51		0.510423704	C2
3	0.49	0.51		0.510425641	C2
4	0.49	0.51		0.510426211	C2
5	0.49	0.51		0.510427617	C2
6	0.49	0.51		0.510430653	C2
7	0.49	0.51		0.510427693	C2
8	0.49	0.51		0.51042968	C2
9	0.49	0.51		0.510425605	C2
10	0.49	0.51		0.510423997	C2
11	0.49	0.51		0.510425769	C2
12	0.49	0.51		0.510431531	C2
13	0.49	0.51		0.510432443	C2
14	0.49	0.51		0.510427494	C2
15	0.49	0.51		0.510424182	C2
16	0.49	0.51		0.510427207	C2
17	0.49	0.51		0.510425461	C2
18	0.49	0.51		0.510425983	C2
19	0.49	0.51		0.510424499	C2
20	0.49	0.51		0.510425616	C2

4.2.2 Analisa Fungsional Sistem

Pada tahap analisa fungsional sistem akan dijelaskan mengenai perancangan sistem *Clustering* kebidangan mahasiswa berdasarkan nilai – nilai yang dimiliki. Analisa yang digunakan menggunakan konsep UML meliputi *usecase*, *sequence diagram*, *class diagram* dan *activity diagram*.

A. *Usecase Diagram*

Pada analisa fungsional *sistem usecase* diagram ini menggambarkan proses yang dilakukan oleh aplikasi secara keseluruhan. Proses tersebut secara garis besar terbagi atas 4 pengelompokan, yaitu pengelolaan mata kuliah , pendaftaran mahasiswa, pengelolaan nilai dan proses klustering semua kebidangan mahasiswa seperti pada Gambar (4.2)



Gambar 4.2 usecase diagram sistem klustering bidang keahlian mahasiswa

Pada proses pengelolaan mata kuliah mencakup penambahan, sunting, dan penghapusan data mata kuliah yang ada pada aplikasi. Pengelolaan mata kuliah didalam nya menggunakan *attribute* nama mata kuliah dan jenis kebidangan mata kuliah tersebut. Proses penambahan data mata kuliah dilakukan sebelum menambahkan data mahasiswa dan nilai kedalam aplikasi.

Pendaftaran mahasiswa secara garis besar menggambarkan proses penambahan data mahasiswa baru di sistem. Proses ini harus dilakukan sebelum

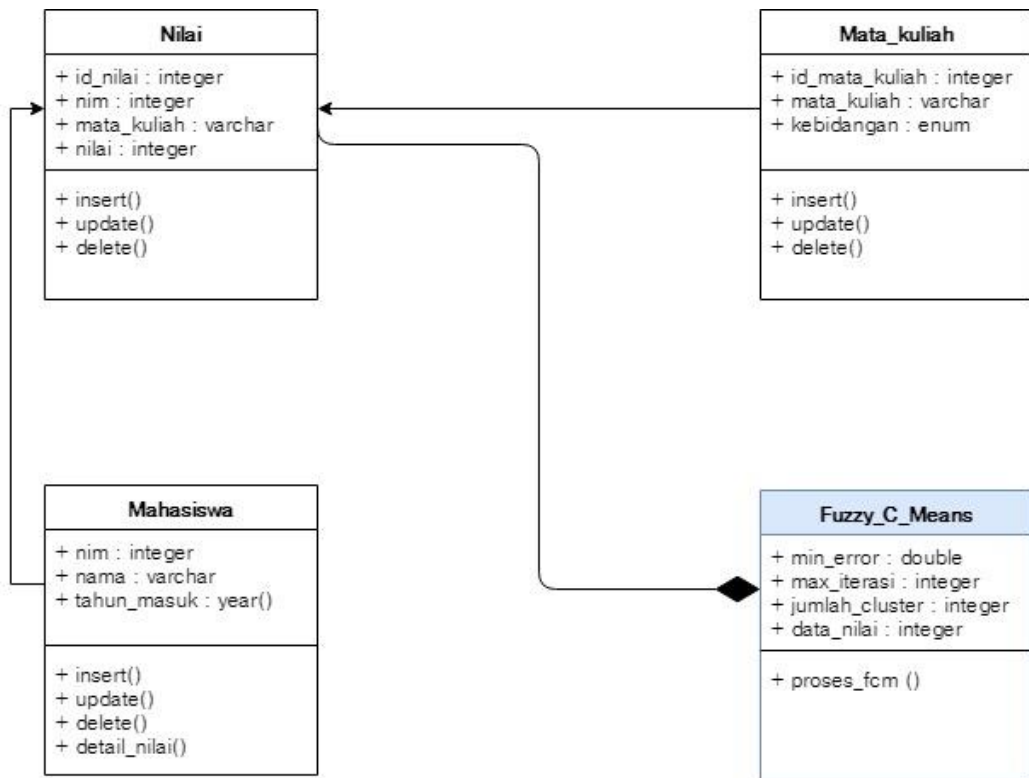
melakukan penambahan data nilai mahasiswa. Setelah mahasiswa ditambahkan kedalam aplikasi, operator baru bisa menambahkan nilai nilai mahasiswa tersebut sesuai dengan semua mata kuliah yang tersedia di aplikasi.

Proses utama penentuan kebidangan dilakukan di tahap klustering sesuai dengan semua data yang tersedia diaplikasi dengan menggunakan metode *fuzzy c-means*.

B. *Class Diagram*

Pada analisa fungsional sistem *class diagram* ini menggambarkan secara detail semua *attribute* dan fungsi – fungsi (*function*) yang digunakan dalam aplikasi. Hal ini untuk mempermudah proses pengembangan aplikasi yang dibangun.

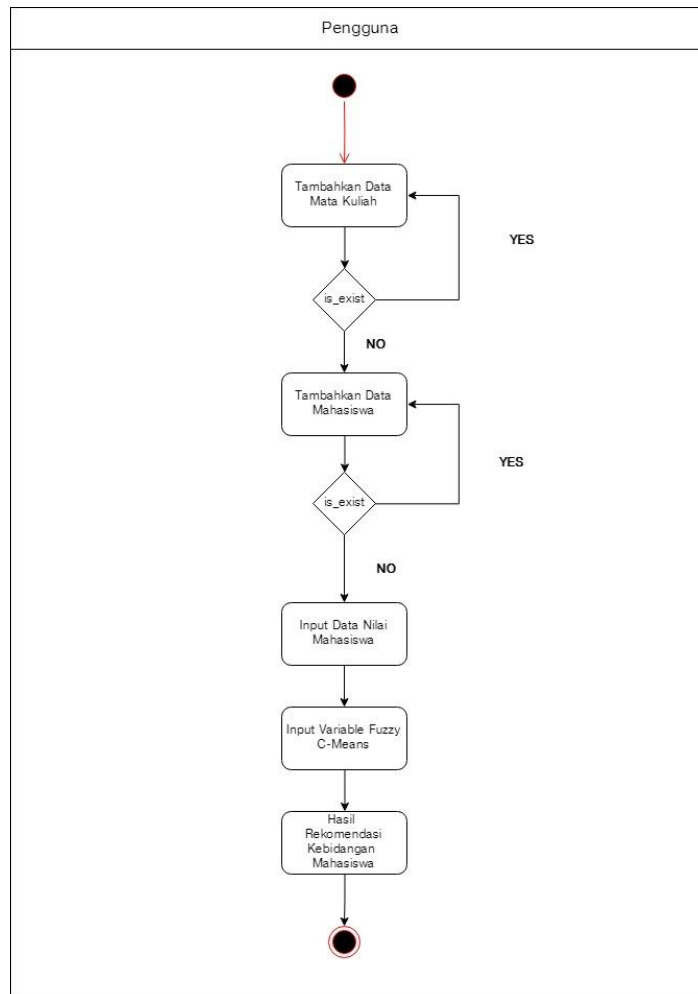
Sesuai dengan proses yang telah digambarkan pada analisa fungsional usecase diagram, prosedur penggunaan aplikasi diawali dengan pengelolaan mata kuliah yang tersedia di jurusan Teknik informatika. Lalu dilanjutkan dengan menambahkan data mahasiswa. Setelah data mata kuliah dan mahasiswa selesai di inputkan, data yang harus ditambahkan adalah nilai mahasiswa yang sesuai dengan data hasil pencapaian mahasiswa berupa nilai dari *range* lebih atau sama dengan 0 sampai dengan kurang atau sama dengan 100. Setelah ketiga data tersebut diinputkan proses selanjutnya adalah klustering data dengan output berupa pengelompokan mahasiswa sesuai dengan kemiripan pola yang ada pada data data di dalam sistem.



Gambar 4.3 Class Diagram

C. Activity Diagram

Pada tahap analisa fungsional *activity diagram* menggambarkan secara sistematis proses penggunaan aplikasi klustering data nilai untuk menentukan kebidangan mahasiswa yang telah ada datanya didalam sistem.



Gambar 4.4 Activity Diagram

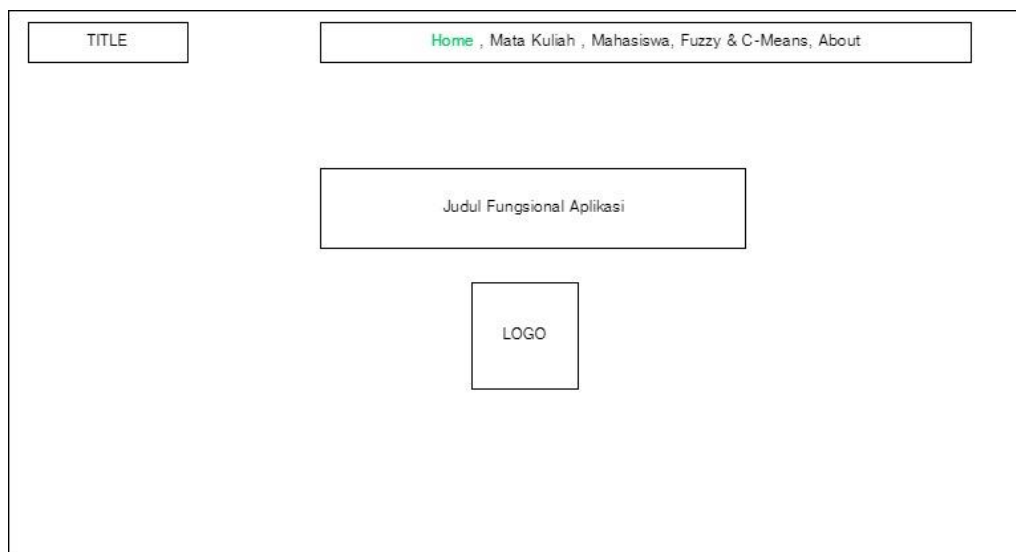
Pada analisa fungsional yang digambarkan, diawali dengan menambahkan data mata kuliah dengan syarat mata kuliah tersebut belum ada didalam sistem. Jika mata kuliah tersebut tersedia, sistem akan otomatis menolak input data mata kuliah tersebut. Setelah dipastikan data tersebut belum ada di sistem, proses selanjutnya adalah menambahkan data mahasiswa yang belum pernah terdaftar disistem. Dengan opsi yang sama, yaitu data akan ditambahkan jika data mahasiswa tersebut belum pernah dimasukkan kedalam sistem. Setelah itu secara otomatis setiap mahasiswa yang didaftarkan disistem akan memiliki data nilai dengan nilai *default* 0. Setelah nilai mahasiswa diinputkan maka proses kluster dapat dilakukan dengan output pengelompokan reomendasi kebidangan mahasiswa berdasarkan kemiripan pola data yang ada di aplikasi.

4.3 Perancangan Sistem

Pada sub bab ini akan dibahas tentang perancangan sistem *Clustering* kebidangan mahasiswa berdasarkan variable nilai dari semester 1 hingga 6. Perancangan ini dibagi menjadi 3 bagian antara lain: perancangan basis data, perancangan struktur menu dan yang terakhir perancangan antar muka (*interface*).

4.3.1 Perancangan Struktur Menu

Perancangan struktur menu menggambarkan hubungan antara suatu menu dengan menu yang lain. Menu merupakan salah satu bagian yang penting dalam merancang sebuah sistem, karena dalam menu dapat dilihat bagaimana struktur dari sebuah sistem. Struktur menu yang dirancang ini bertujuan untuk kemudahan proses pengembangan aplikasi.



Gambar 0.5 Perancangan Struktur Menu Aplikasi

4.3.2 Perancangan Basis Data

Merancang basis data merupakan tahapan setelah perancangan struktur menu dari sebuah perancangan sistem aplikasi. Dalam tahapan ini dilakukan analisa terhadap tabel-tabel yang dibutuhkan pada sistem pengelompokkan data kebidangan mahasiswa.

A. Tabel Mata Kuliah

No.	Field	Type	Keterangan
1	id_mata_kuliah	Integer	Primary_key, auto increment
2	mata_kuliah	Varchar	Unique
3	bidang	Enum	TI / IK

B. Tabel Mahasiswa

No.	Field	Type	Keterangan
1	nim	Big Integer	Primary_key
2	mata_kuliah	Varchar	Foreign Key
3	tahun_masuk	Year()	

C. Tabel Nilai

No.	Field	Type	Keterangan
1	id_nilai	Integer	Primary_key, Auto Increment
2	nim	Big Integer	Foreign Key
3	mata_kuliah	Varchar	Foreign Key
4	Nilai	Integer	

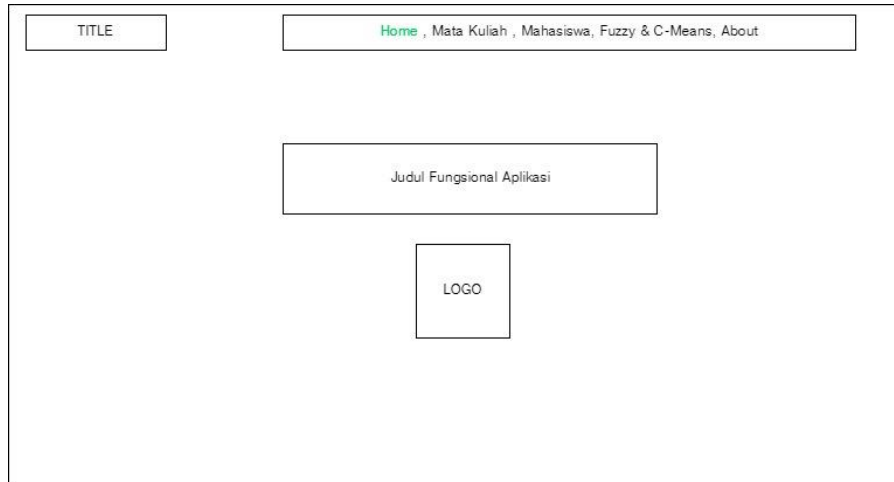
4.3.3 Perancangan Antarmuka (*Interface*)

Antarmuka (*Interface*) sistem adalah sarana pengembangan sistem yang digunakan untuk membuat komunikasi yang lebih mudah antara sistem dengan pemakainya. Penekanan *interface* meliputi tampilan yang baik, mudah dipahami, dan tampilan menu-menu yang mudah dimengerti. Interface sistem yang akan dibangun adalah sebagai berikut:

A. Perancangan Antarmuka Halaman Awal (*home*)

Halaman awal merupakan halaman yang di akses saat aplikasi ini baru diakses atau dibuka. Pada halaman ini akan menampilkan rancangan berupa judul

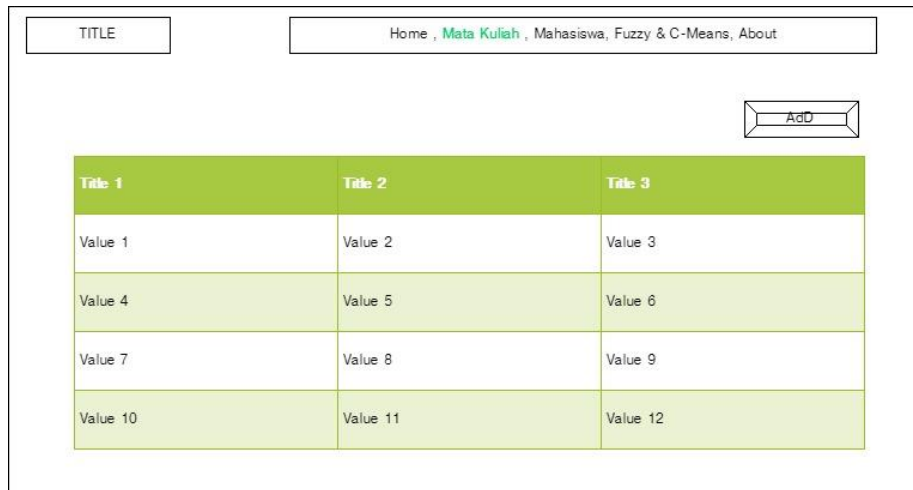
dengan konten garis besar fungsional aplikasi. Perancangan antarmuka halaman awal adalah sebagai berikut



Gambar 4.6 Rancangan antarmuka halaman awal

B. Perancangan Antarmuka Menu Mata Kuliah

Pada perancangan antarmuka mata kuliah berisi tentang pengelolaan data mata kuliah seperti tambah , edit, lihat dan hapus seperti pada gambar dibawah ini.



Gambar 4.7 Rancangan antarmuka halaman menu mata kuliah

C. Perancangan Antarmuka Menu Mahasiswa

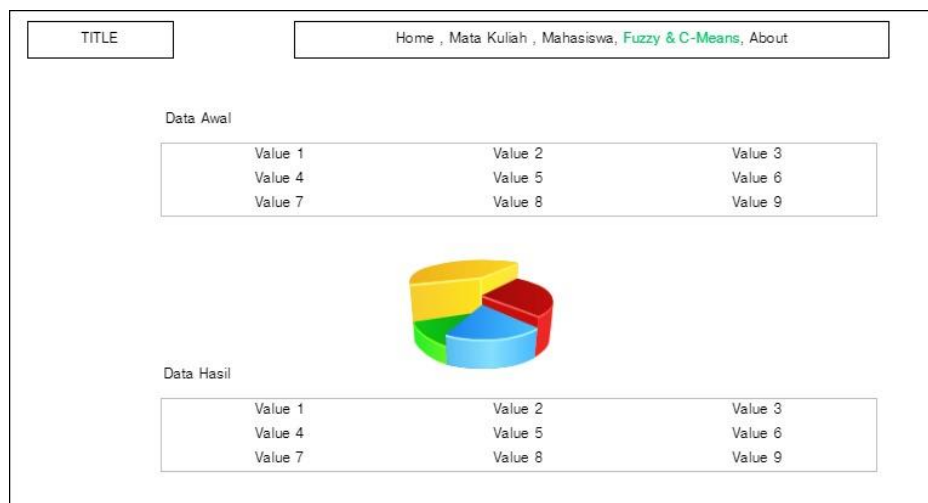
Halaman mahasiswa perancangannya berisi tentang pengelolaan mahasiswa kuliah seperti tambah , edit, lihat , hapus dan melihat detail nilai mahasiswa seperti pada gambar dibawah ini.



Gambar 4.8 Rancangan antarmuka halaman menu mahasiswa

D. Perancangan Antarmuka Proses Fuzzy C-Means

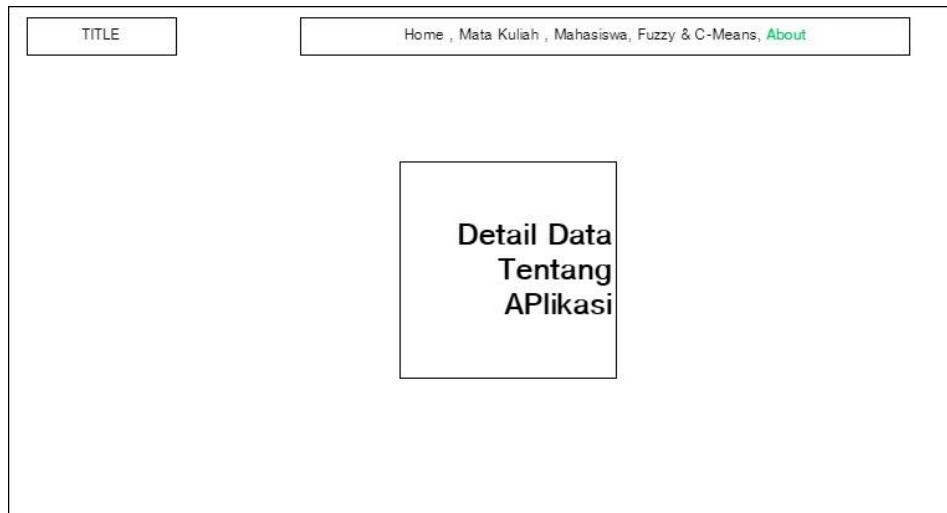
Halaman ini digunakan untuk melihat hasil pengelompokan kebidangan mahasiswa berdasarkan pencapaian nilai dari semester 1 sampai dengan semester 6.



Gambar 4.9 Rancangan antarmuka halaman menu proses fuzzy c-means

E. Perancangan Antarmuka tentang aplikasi (about)

Halaman tentang aplikasi (about) perancangannya berisi tentang detail informasi detail aplikasi



Gambar 4.10 Rancangan antarmuka halaman menu tentang aplikasi

BAB V

PENGUJIAN DAN HASIL

5.1 Implementasi

Implementasi merupakan tahap development dan evaluasi sistem berdasarkan hasil analisa dan perancangan sistem sebelumnya sehingga sistem yang dibuat dapat difungsikan dalam keadaan sebenarnya dan sesuai dengan tujuan yang diharapkan. Adapun tujuan dari implementasi dan pengujian adalah sebagai berikut:

1. Menyelesaikan desain sistem yang telah dibuat dalam dokumen perancangan yang telah disetujui.
2. Menguji dan mendokumentasikan program atau prosedur dari dokumen perancangan sistem yang telah disetujui.
3. Mempertimbangkan bahwa sistem memenuhi permintaan pemakai, yaitu dengan menguji sistem secara keseluruhan.
4. Memastikan bahwa konversi ke sistem yang baru berjalan dengan benar yaitu dengan membuat rencana, mengontrol dan perancangan sistem secara benar.

5.1.1 Batasan Implementasi

Batasan implementasi dari penelitian ini adalah:

1. Bahasa pemrograman yang digunakan dalam pengimplementasian sistem ini yaitu PHP dengan DBMS MySQL pada sistem operasi *Microsoft Windows 8*.
2. Adapun algoritma yang digunakan dalam sistem adalah dengan Algoritma *Fuzzy Clustering Means* (FCM)

5.1.2 Lingkungan Implementasi

Lingkungan implementasi adalah lingkungan dimana aplikasi ini dikembangkan. Untuk mengimplementasikan sistem ini maka dibutuhkan beberapa komponen pendukung diantaranya perangkat keras dan perangkat lunak. Pada penelitian ini, implementasi dilakukan menggunakan perangkat dengan minimum spesifikasi sebagai berikut:

A. Perangkat Keras

1. Processor : Intel® Celeron® CPU 1007U @ 1.50 GHz
2. Memori : 2.00 GB
3. Hardisk : 250 GB

B. Perangkat Lunak

1. Sistem Operasi : WINDOWS 10 / 8 / 7
2. Web Browser : *Google Chrome , Microsoft Edge , Safari, Mozilla Firefox*
3. Dbms : Mysql
4. Web server : Apache
5. Perangkat Pendukung : Xampp

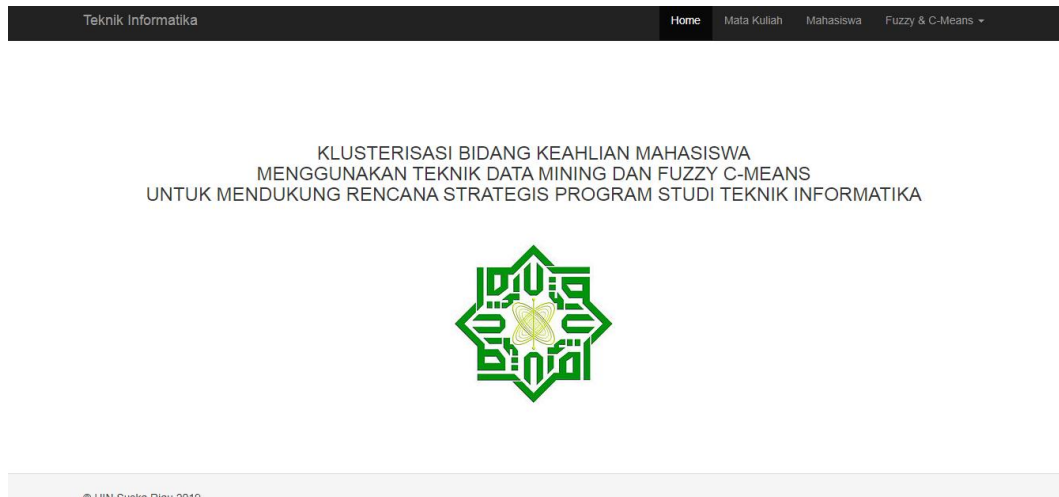
Perangkat Lunak dapat di jalankan dengan menggunakan *browser Google Chrome, Microsoft Edge, Safari, Mozilla Firefox* dengan mengaktifkan <http://fcm.reproment.com/> yang berisikan aplikasi untuk klusterisasi bidang keahlian mahasiswa Menggunakan teknik data mining dan *fuzzy c-means*.

5.1.3 Implementasi Antar Muka Sistem

Pada implementasi sistem akan ditunjukkan menu utama, sub menu dan proses lainnya yang dapat dilakukan pada Perangkat Lunak untuk klusterisasi bidang keahlian mahasiswa Menggunakan teknik data mining dan *fuzzy c-means*.

1. Implementasi Halaman Awal

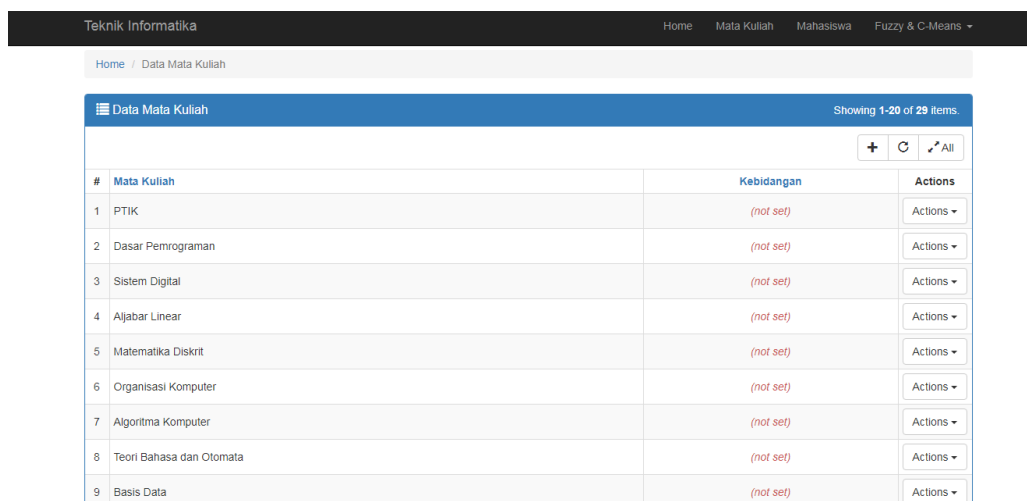
Halaman ini merupakan halaman pembuka atau halaman home pada perangkat lunak yang berisikan menu Home, Mata kuliah, dan Fuzzy C Means.



Gambar 5.1 Halaman Home

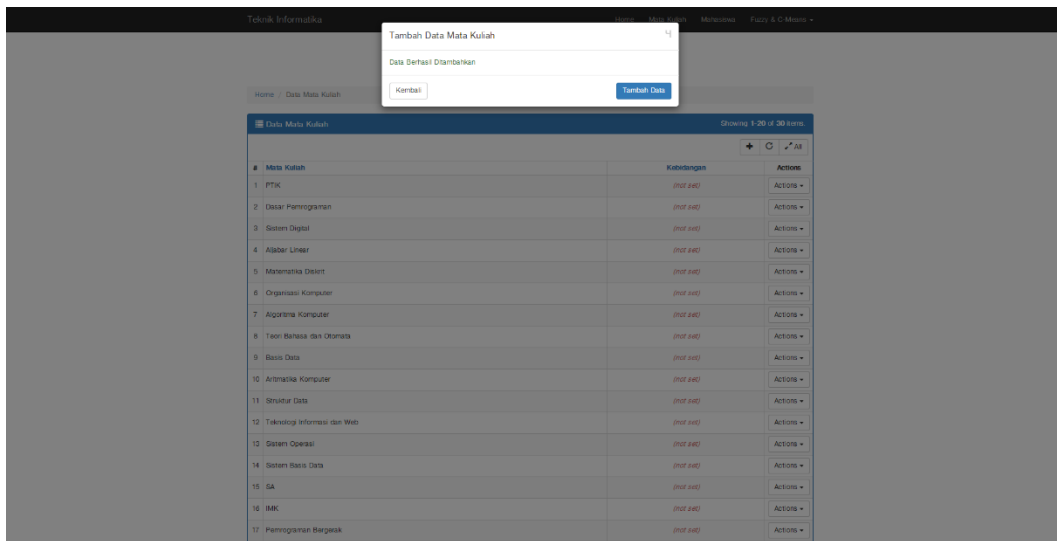
2. Implementasi Halaman Data Master.

Halaman ini merupakan halaman data master tempat menginputkan nama mata kuliah pada perangkat lunak yang berisikan menu action untuk *view*, *update delete*, pada Perangkat Lunak untuk klusterisasi bidang keahlian mahasiswa Menggunakan teknik data mining dan *fuzzy c-means*.



#	Mata Kuliah	Kebidangan	Actions
1	PTIK	(not set)	Actions ▾
2	Dasar Pemrograman	(not set)	Actions ▾
3	Sistem Digital	(not set)	Actions ▾
4	Aljabar Linear	(not set)	Actions ▾
5	Matematika Diskrit	(not set)	Actions ▾
6	Organisasi Komputer	(not set)	Actions ▾
7	Algoritma Komputer	(not set)	Actions ▾
8	Teori Bahasa dan Otomata	(not set)	Actions ▾
9	Basis Data	(not set)	Actions ▾

Gambar 5.2 Halaman Data Master



Setelah data berhasil ditambahkan maka akan muncul pesan data berhasil ditambahkan.

3. Implementasi Halaman Data Mahasiswa.

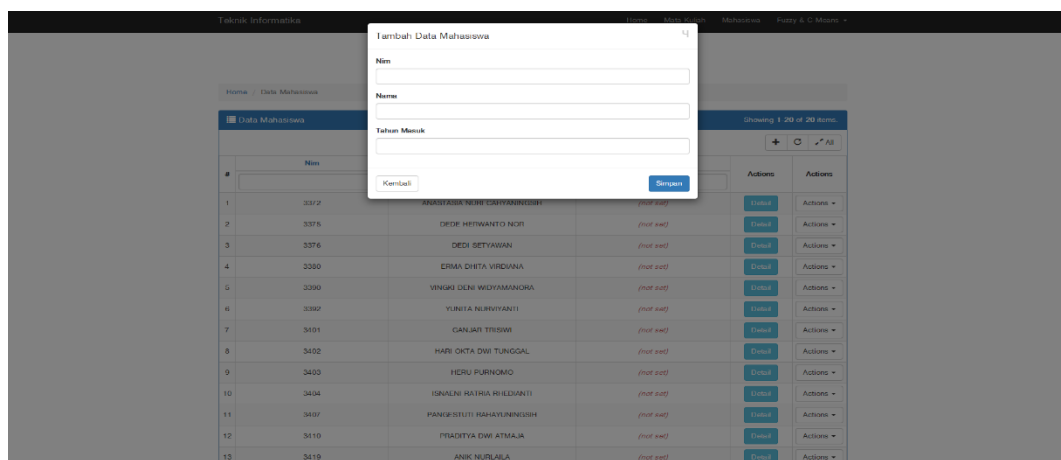
Halaman ini merupakan halaman data mahasiswa tempat menginputkan Nim, Nama, Nilai setiap mata kuliah pada perangkat lunak yang berisikan menu action untuk *view*, *update* *delete*, pada Perangkat Lunak untuk klusterisasi bidang keahlian mahasiswa Menggunakan teknik data mining dan *fuzzy c-means*.

#	Nim	Nama	Tahun Masuk	Actions	Actions
1	3372	ANASTASIA NURI CAHYANINGSIH	(not set)	Detail	Actions
2	3375	DEDE HERWANTO NOR	(not set)	Detail	Actions
3	3376	DEDI SETYAWAN	(not set)	Detail	Actions
4	3380	ERMA DHITA VIRDIANA	(not set)	Detail	Actions
5	3390	VINGKI DENI WIDYAMANORA	(not set)	Detail	Actions
6	3392	YUNITA NURVIYANTI	(not set)	Detail	Actions
7	3401	GANJAR TRISWI	(not set)	Detail	Actions
8	3402	HARI OKTA DWI TUNGGAL	(not set)	Detail	Actions
9	3403	HERU PURNOMO	(not set)	Detail	Actions

Gambar 5.3 Data Mahasiswa

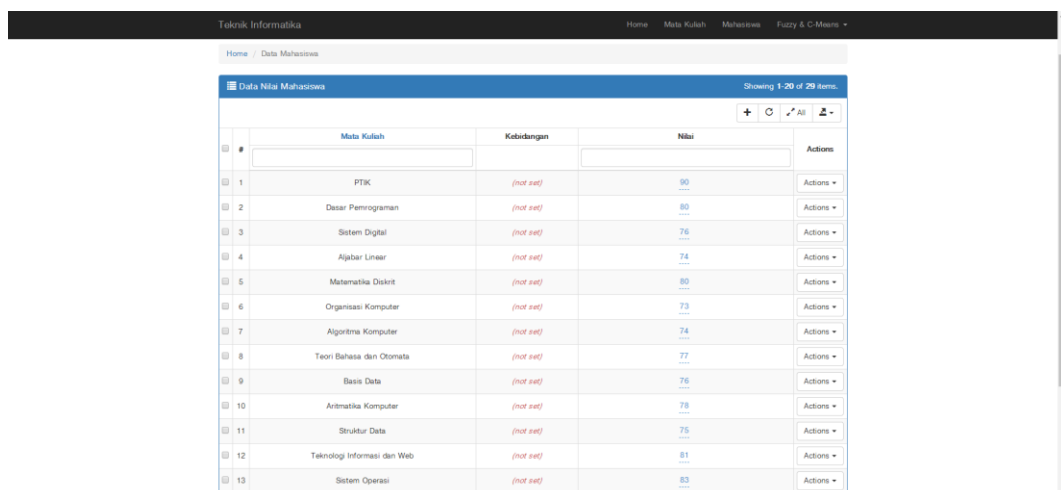
Halaman ini merupakan tempat untuk melihat data mahasiswa dan nilai mata kuliah. Pada halaman ini juga ada sebuah aksi untuk melihat dan mengedit data nilai mahasiswa sesuai dengan semua mata kuliah yang tersedia di data master mata kuliah.

Pada halaman data mahasiswa terdapat 4 aksi utama yaitu tambah data, edit, hapus dan mengedit data nilai mata kuliah setiap siswa. Tahap awal saat menambahkan data mahasiswa, pengguna wajib mengisi detail data mahasiswa



Gambar 5.4 Halaman Form Input Data Mahasiswa baru

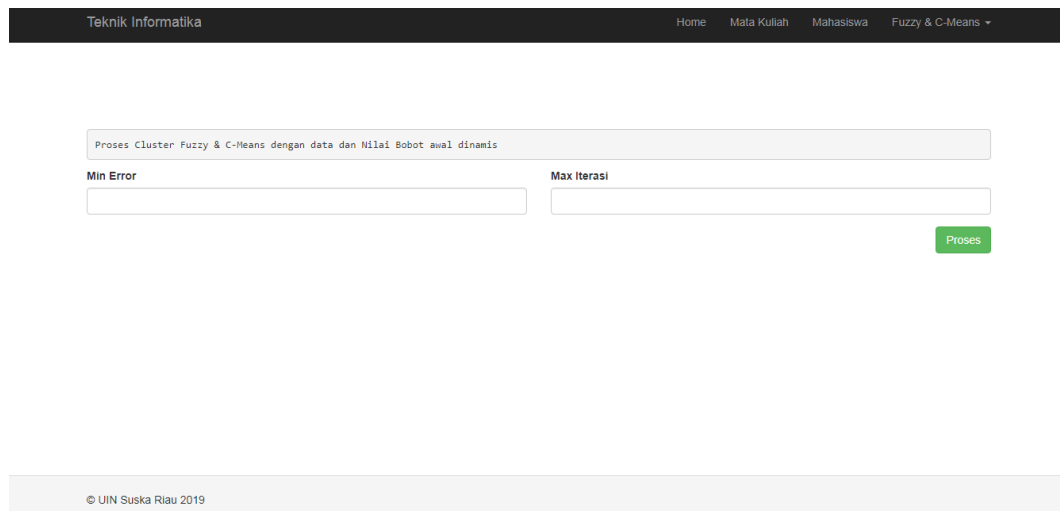
Setelah menambahkan data mahasiswa, selanjutnya sistem akan mengalihkan secara otomatis aplikasi ke halaman detail nilai mahasiswa. Setelah itu pengguna melakukan edit data nilai mahasiswa



Setelah menambahkan data mahasiswa baru dan sistem menampilkan tampilan detail nilai. Maka pengguna disarankan untuk menyesuaikan data nilai yg tertera diubah sesuai dengan nilai nilai yang didapatkan oleh mahasiswa yang telah didaftarkan sebelumnya

4. Implementasi Halaman Proses FCM.

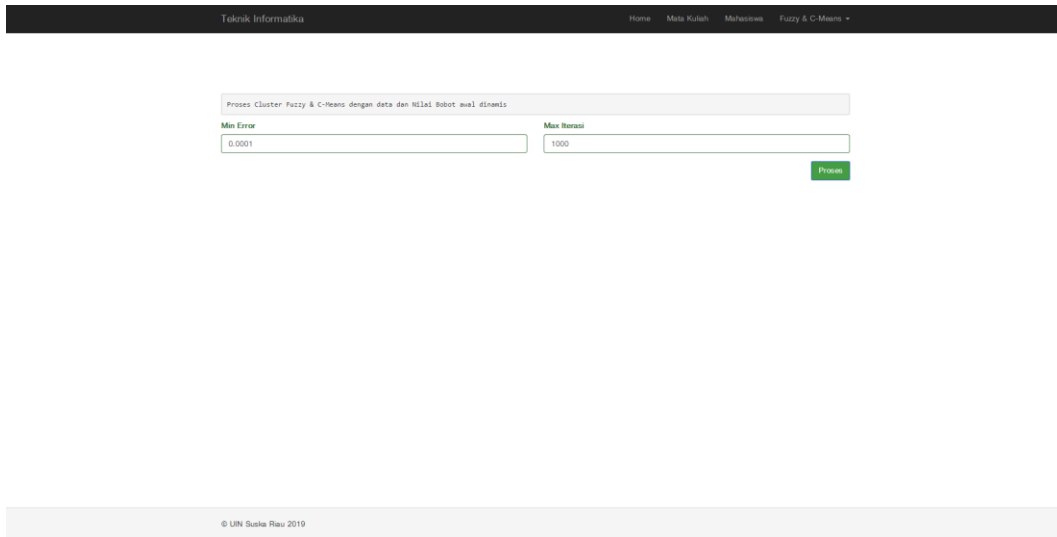
Halaman ini merupakan halaman proses FCM untuk memproses pengelompokan bidang keahlian mahasiswa pada Perangkat Lunak untuk klusterisasi bidang keahlian mahasiswa menggunakan teknik data mining dan *fuzzy c-means*.



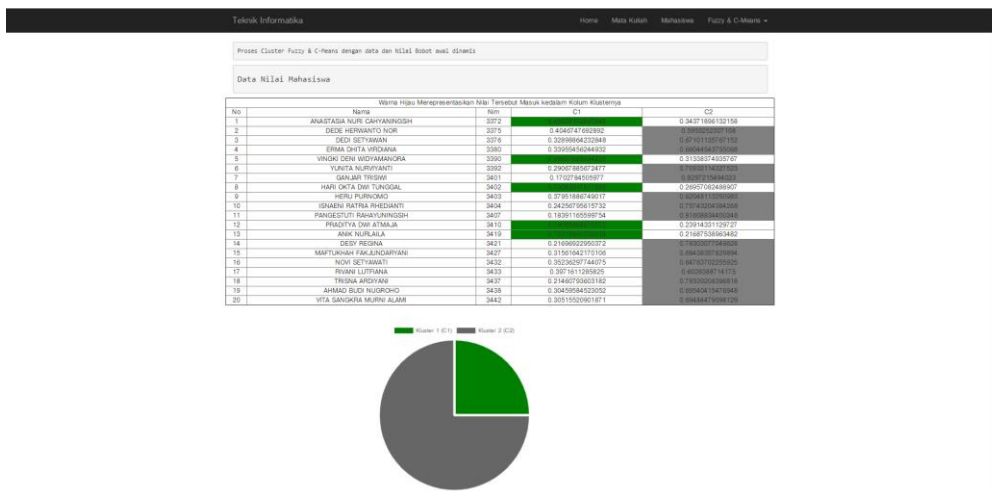
The screenshot shows a web interface for the FCM process. At the top, there is a navigation bar with the text 'Teknik Informatika' on the left and 'Home', 'Mata Kuliah', 'Mahasiswa', and 'Fuzzy & C-Means' on the right. Below the navigation bar, there is a title bar that reads 'Proses Cluster Fuzzy & C-Means dengan data dan Nilai Bobot awal dinamis'. Underneath the title bar, there are two input fields: 'Min Error' and 'Max Iterasi'. To the right of these input fields is a green button labeled 'Proses'. At the bottom of the page, there is a footer that reads '© UIN Suska Riau 2019'.

Gambar 5.5 Halaman Proses FCM

Halaman ini digunakan untuk melakukan proses *Clustering*. Data yang sudah ditambahkan sebelumnya akan diproses menggunakan metode *fuzzy c-means*. Semua data yang telah ditambahkan seperti data mahasiswa dan matakuliah akan di proses dengan tampilan awal yaitu menampilkan form yang berisi nilai minimal error dan nilai maksimal iterasi. Dua variable input ini akan menjadi acuan proses cluster berhenti pada salah satu kondisi tersebut. Selain itu data yang akan diolah adalah data mahasiswa dan mata kuliah yang tersedia di database aplikasi.



Setelah mengisi variable minimal error dan maksimal iterasi, maka pengguna harus memilih tombol proses. Lalu aplikasi akan mengolah semua data sesuai dengan proses alur algoritma *fuzzy c-means*.



Gambar 5.6 Hasil Data Clustering

Setelah proses itu selesai sistem akan menampilkan hasil data clustering menggunakan metode fuzzy c-means.

5.2 Pengujian

Pengujian sistem dilakukan terhadap sistem yang telah dibangun. Pengujian sistem bertujuan untuk menguji sistem yang telah dibangun sesuai dengan hasil analisa dan

perancangan yang telah dilakukan, sehingga dapat dibuat satu atau lebih kesimpulan akhir. Pengujian pada penelitian ini terdiri dari :

5.2.1 Pengujian Sistem dengan *Black Box*

A. Mata Kuliah

Pengujian menu halaman mata kuliah pada sistem dapat dilihat pada tabel 5.1 dibawah ini:

Tabel 5.1 Pengujian Form Mata Kuliah

Kasus uji	Langkah pengujian	Masukkan	Keluaran	Hasil Uji
Pengujian Form mata kuliah	Menambahkan data mata kuliah baru	Mata kuliah, kebidangan	Aplikasi akan menampilkan data semua mata kuliah beserta mata kuliah yang baru ditambahkan	Sukses

B. Menu Mahasiswa

Berikut pada tabel 5.2 dibawah ini merupakan pengujian form Data Menambahkan data mahasiswa beserta data nilai

Tabel 5.2 Pengujian Menu Data Mahasiswa dan Nilai

Kasus uji	Langkah Pengujian	Masukkan	Keluaran	Hasil Uji
Pengujian form data mahasiswa	Pengguna memilih menu mahasiswa, lalu menambahkan data data mahasiswa	Data mahasiswa	Data mahasiswa berhasil ditampilkan dan disimpan	Sukses
Pengujian form data	Pengguna menggunakan	Data Nilai Mahasiswa	Data berhasil disimpan	Sukses

Kasus uji	Langkah Pengujian	Masukkan	Keluaran	Hasil Uji
nilai mahasiswa	menu detail pada mahasiswa atau setelah proses menambahkan data mahasiswa langsung melakukan edit data pada variable nilai mahasiswa tersebut	sesuai dengan mata kuliah yang tersedia		

C. Menu *Clustering*

Berikut ini merupakan pengujian pada form data Data Menu *Clustering* dibawah ini:

Tabel 5.3 Pengujian menu *Clustering*

Kasus uji	Langkah pengujian	Masukkan	Keluaran	Hasil Uji
Pengujian form clustering	Pada tahap ini pengguna memilih menu clustering dinamis dan memasukkan data variable dan memilih menu proses	Nilai minimal error dan maksimal iterasi	Sistem dapat membagi data yang tersedia menjadi 2 cluster dan menampilkan statistic data cluster	Sukses

Pengujian terhadap perangkat lunak yang di bangun dengan melakukan validasi kepada Pakar Teknologi Informasi.

5.2.2 Uji Validasi Perangkat Lunak.

Tujuan validasi ini adalah untuk mendapatkan masukan mengenai kekurangan perangkat lunak yang di bangun. Masukan tersebut kemudian dianalisis dan digunakan untuk merevisi aplikasi dalam pengembangan model seleksi konsentrasi dengan inventori minat berbasis sistem cerdas sehingga dapat meningkatkan kualitas produk yang akan digunakan dalam penelitian. Data validasi ahli teknologi ini terbagi menjadi 3 aspek yaitu: pengoperasian aplikasi, desain aplikasi dan bahasa.

Tabel 5.4 Hasil Validasi Perangkat Lunak dari Aspek Pengoperasian Aplikasi

No.	Indikator	Rata-Rata	Kategori
1.	Kelayakan pengelompokkan bidang keahlian dengan menggunakan Perangkat Lunak ini untuk diakses dengan tepat dan cepat di berbagai web browser (Mozilla Firefox, Opera, Chrome, dan Internet Explorer).	0,92	Valid
2.	Kemudahan untuk menggunakan navigasi menu yang ada.	0,75	Valid
3.	Kemudahan petunjuk yang tersedia untuk dilihat dan dibaca.	0,83	Valid
4.	Kemudahan menginput data dan nilai.	0,83	Valid
5.	Kelayakan perhitungan yang dihasilkan untuk membantu Prodi menentukan bidang keahlian mahasiswa.	0,92	Valid
6.	Kemudahan untuk melakukan login ke sistem dengan tepat dan cepat.	0,75	Valid
Rata-Rata		0,83	Valid

Tabel 5.5 Hasil Validasi Perangkat Lunak dari Aspek Desain Aplikasi

No.	Indikator	Rata-Rata	Kategori
1.	Penggunaan huruf, baik jenis huruf, warna dan ukuran.	0,75	Valid
2.	Pemilihan gambar dan ilustrasi.	0,75	Valid

3.	Pemilihan kombinasi warna (<i>theme</i>).	0,75	Valid
4.	Keserasian tampilan untuk bagian-bagian dari aplikasi.	0,83	Valid
5.	Penempatan <i>Link</i> dan Navigasi.	0,75	Valid
6.	Pemilihan Desain <i>Backgorund</i> .	0,75	Valid
Rata-Rata		0,76	Valid

Tabel 5.6 Hasil Validasi Perangkat Lunak dari Aspek Bahasa

No.	Indikator	Rata-Rata	Kategori
1.	Bahasa yang digunakan pada Perangkat Lunak mudah dipahami	0,75	Valid
2.	Bahasa yang digunakan dalam Perangkat Lunak bersifat komunikatif	0,83	Valid
3.	Bahasa pada Perangkat Lunak menggunakan bahasa Indonesia yang baik dan benar	0,75	Valid
4.	Bahasa yang digunakan pada aplikasi sistematis	0,83	Valid
Rata-Rata		0,79	Valid

Secara keseluruhan, hasil validasi untuk perangkat lunak ini, dapat dilihat pada tabel.

Tabel 5.7 Hasil Uji Validasi Perangkat Lunak Secara Keseluruhan

No.	Aspek	Rata-Rata	Kategori
1.	Aspek Pengoperasian Aplikasi	0,83	Valid
2.	Aspek Desain Aplikasi	0,76	Valid
3.	Aspek Bahasa	0,83	Valid
Rata-rata		0,80	Valid

Berdasarkan penilaian di atas, validasi perangkat lunak memperoleh nilai 0,80. Merujuk kepada Azwar (2014) bahwa jika rentang angka V yang didapat $\geq 0,667$ dapat diinterpretasikan sebagai koefisien yang cukup tinggi, sehingga dapat dikategorikan bahwa kategori validitasnya berada dalam kategori “valid”.

5.2.3 Uji Praktikalitas kepraktisan Perangkat Lunak oleh Jurusan.

Angket praktikalitas oleh Jurusan ini diberikan kepada 2 orang yang bertugas sebagai pengelola di Jurusan, yaitu : ketua Jurusan, sekteraris Jurusan.

Instrumen uji praktis oleh Jurusan terhadap perangkat lunak diperoleh nilai praktis sebesar 85,33 dengan kategori praktis. Merujuk kepada Purwanto (2009) bahwa rentang nilai 80-89 diinterpretasikan praktis. Dengan kuisisioner sebagai berikut:

Aspek Kepraktisan	No	Pernyataan	Skor				
			1	2	3	4	5
Kualitas Isi dan Tujuan	1	Perangkat Lunak ini dapat melakukan pengelompokkan bidang keahlian mahasiswa					
	2	Langkah-langkah dalam melakukan penyeleksian dalam aplikasi ini dapat dilaksanakan dengan jelas.					
	4	Kualitas isi dan hasil yang ada dalam Perangkat Lunak ini ini bagus.					
	5	Petunjuk dalam Perangkat Lunak ini jelas dan mudah dipahami.					
Kualitas Teknik	6	Tampilan dari Perangkat Lunak ini menarik.					
	7	Komposisi warna dari Perangkat Lunak ini menarik.					
	8	Warna background dalam Perangkat Lunak ini sesuai dan menarik.					
	9	Ukuran tombol yang digunakan dalam Perangkat Lunak ini sesuai dengan tampilan.					
	10	Pemilihan font (huruf) dalam Perangkat Lunak ini sesuai tampilan.					
	11	Teks dan kalimat yang digunakan dalam Perangkat Lunak ini mudah dibaca.					
	12	Kemudahan membaca teks dan kalimat yang ada dalam Perangkat Lunak ini.					

Penggunaan	13	Kemudahan memahami bahasa yang digunakan dalam Perangkat Lunak ini.				
	14	Kemudahan menggunakan tombol-tombol yang ada dalam Perangkat Lunak ini.				
	15	Perangkat Lunak ini bersifat <i>user friendly</i> sehingga mudah digunakan.				
	16	Proses pengelompokkan bidang keahlian melalui model ini membuat proses lebih mudah dan cepat.				

Tabel 5.8 Hasil Uji Kepraktisan Perangkat Lunak

R	NO ITEM															Jumlah	Skor Maks	%
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15			
1	4	4	4	4	5	4	5	4	4	4	5	5	5	4	4	65	75	86,67
2	4	4	4	4	4	3	4	3	4	5	5	5	4	5	5	63	75	84,00

BAB VI

KESIMPULAN DAN SARAN

6.1 Kesimpulan

Bedasarkan hasil penelitian dan pembahasan yang mengacu pada tujuan penelitian ini, maka dapat disimpulkan bahwa:

1. Pengembangan Perangkat Lunak untuk klusterisasi bidang keahlian mahasiswa Menggunakan teknik data mining dan *fuzzy c-means* ini dapat membantu Program Studi dalam proses pengelompokan bidang keahlian mahasiswa sehingga membantu memberikan rekomendasi kepada mahasiswa dalam memilih bidang keahlian yang sesuai dengan kemampuan mereka sehingga mendukung rencana strategis program studi Teknik informatika
2. Telah dihasilkan Perangkat Lunak untuk klusterisasi bidang keahlian mahasiswa Menggunakan teknik data mining dan *fuzzy c-means* dengan hasil pengujian untuk pengujian validitas perangkat lunak dengan nilai rata-rata 0,80 dan praktikalitas perangkat lunak dengan nilai rata-rata 85,33.

6.2 Saran

Bedasarkan hasil penelitian dan pembahasan yang mengacu pada batasan penelitian ini, maka dapat disarankan bahwa:

1. Bagi peneliti perlunya kajian secara lebih mendalam agar model yang dikembangkan dapat memenuhi kebutuhan mahasiswa dan kebutuhan Jurusan dan penekanannya bukan pada produk semata, tetapi dapat dilihat dalam proses yang berlangsung
2. Hasil penelitian yang telah di lakukan dapat menjadi rujukan, kerangka kerja dan pedoman dalam melaksanakan pengembangan model pengelompokan lain nya.

DAFTAR PUSTAKA

- Agbonifo & Catherine, O., 2013. “*Fuzzy C-Means Clustering Model for Identification of Students Learning Preferences in Online Environment*”, International Journal of Computer Applications & Information Technology, IV(1),pp.15-21.
- Ahmed, N., Rasheed, A., and Jehanzeb, K. (2012). “*An Exploration of Predictors of Organizational Citizenship Behaviour and Its Significant Link to Employee Engagement*”, International Journal of Business, Humanities and Technology, Vol 2, No 4, pp. 99-106
- Bora, D.J., dan Gupta, A.K., 2014, “*A Comparative Study Between Fuzzy Clustering Algorithm and Hard Clustering Algorithm*”, International Journal of Computer Trends and Technology (IJCTT) -volume 10 number 2 –Apr 2014
- Cox, Earl. 2001. “*The Fuzzy System Handbook*”. AP Professional
- Davies, and Paul Beynon, 2004, “*Database System Third Edition*”, Palgrave macmillan, New York.
- Han, J. And Kamber, M, 2006, “*Data Mining Concept and Techniques Second Edition*”, Morgan Kauffma, San Fransisco
- Hari dan Kusumadewi. 2004. “*Aplikasi Logika Fuzzy untuk Pendukung Keputusan*”. Yogyakarta: Graha Ilmu.
- Hastuti, Khafiizh. 2012. “*Analisis Komparasi Algoritma Klasifikasi Data Mining Untuk Prediksi Mahasiswa Non Aktif*”. Seminar Nasional Teknologi Informasi & Komunikasi Terapan 2012. ISBN 979 -26 –0255 -0.
- Hurlock, E. B. (1999). “*Psikologi Perkembangan Suatu Pendekatan Sepanjang Rentang kehidupan*”. (E. 5, Ed.). Jakarta: Erlangga.
- Inyang, U.G. & Joshua, E.E., 2013. “*Fuzzy Clustering of Students Data Repository for At-Risks Students Identification and Monitoring*”, Journal International Computer and Information Science, VI(4), pp.37-50.

- Kusumadewi, Sri. 2002. “*Analisis & Desain Sistem Fuzzy Menggunakan Tool Box Matlab*”, Yogyakarta: Graha Ilmu
- Kusumadewi, Sri. 2003. “*Artificial Intelligence: Teknik dan Aplikasinya*”, Yogyakarta: Graha Ilmu.
- Larose, D. T., 2006, “*Data Mining Methods And Models*”, John Wiley & Sons, Inc., Hoboken, New Jersey.
- Munandar, T.A., Widyarto, W.O. & Harsiti, 2013. “*Clustering Data Nilai Mahasiswa untuk pengelompokan Konsentrasi Jurusan Menggunakan Fuzzy cluster Means*”. In Seminar Nasional Aplikasi Teknologi Informasi (SNATI). Yogyakarta, 2013.
- Munir, Rinaldi. 2005. “*Matematika Diskrit*”. Bandung: Informatika
- Pramudiono, I. 2007. “*Pengantar Data Mining : Menambang Permata Pengetahuan di Gunung Data*”.<http://www.ilmukomputer.org/wp-content/upload/2006/iko-datamining.zip> Diakses pada tanggal 5 September 2014
- Ravichandran, and Dinakaran, K. 2009. “*Hybrid Fuzzy C-Means Clustering Technique for Gene Expression Data*”, Volume 1 Dept. of Computer Science and Engineering Hindustan Institute of Tech., Coimbatore, India: International Journal of Research and Reviews in Applied Sciences, ISSN: 2076-734X, EISSN: 2076-7366
- Redjeki, Sri. (2017). “*Pemodelan Pengelompokan Prestasi Dosen Menggunakan Metode Fuzzy C-Means*”. Jiko (Jurnal Informatika dan Komputer). 2. 67. 10.26798/jiko.2017.v2i2.63.
- Setiyono, B. & Isnanto, R.R., 2008. “*Pembagian Kelas Kuliah Mahasiswa menggunakan algoritma pengklusteran Fuzzy C-Means*”. In Prosiding Seminar Nasional Teknoin 2008 Bidang Teknik Elektro. Yogyakarta, 2008
- Turban, E., Aronson, J. E., dan Liang, T. P. 2005. “*Sistem Pendukung Keputusan dan Sistem Cerdas*”. Terjemahan Dwi Prabantini. Yogyakarta : ANDI
- Varghese, B.M., J, J.T., A, U. & K, J.P., 2011. “*Clustering Student Data to Characterize Performance Patterns*”. International Journal of Advanced Computer Science and Applications, I(3), pp.138-40.

- Yadav, S.R. & Singh, V.P., 2012. Modeling Academic Performance Evaluation using Fuzzy C-Means Clustering Techniques. *International Journal of Computer Applications*, LX(8), pp.15-23.
- Yusuf, A., Ginardi, H. & Arieshanti, I., 2012. Pengembangan Perangkat Lunak Prediktor Nilai Mahasiswa Menggunakan Metode Spectral Clustering dan Bagging Regresi Linier. *Jurnal Teknik ITS*, I(1), pp.246-50.