

Implementasi Fuzzy C-Means Untuk Prediksi Perilaku Mahasiswa Berdasarkan Jumlah Ketidakhadiran

Ridwan Rismanto¹, Imam Fahrur Rozi², Arief Prasetyo³

Abstrak—Dalam kegiatan belajar mengajar di perguruan tinggi, mahasiswa dituntut tertib dalam melaksanakan perkuliahan. Sebagai syarat agar mahasiswa dapat melanjutkan ke semester berikutnya adalah jumlah kehadiran dan indeks prestasi. Dalam pelaksanaannya, selalu ada mahasiswa bermasalah yang melakukan pelanggaran yaitu jumlah ketidakhadiran yang melebihi batas sehingga harus diterbitkan Surat Peringatan, atau bahkan putus studi. Tujuan surat peringatan ini adalah agar mahasiswa mengetahui status akademiknya sekaligus memunculkan efek jera agar kesalahan yang serupa tidak diulangi lagi. Namun pada prakteknya, masih ada saja mahasiswa yang melanggar meskipun sudah pernah mendapatkan surat peringatan. Kecenderungan yang terjadi adalah mahasiswa yang memiliki jumlah ketidakhadiran cukup banyak selama beberapa semester, pada semester berikutnya jumlah ketidakhadirannya pun akan banyak lagi. Hal ini juga berdampak pada indeks prestasi semesternya. Dalam penelitian ini akan dibuat sebuah sistem yang dapat mengklasifikasikan mahasiswa berdasarkan jumlah ketidakhadiran mahasiswa selama dua semester. Metode klasifikasi yang digunakan adalah clustering dengan Fuzzy C-Means. Hasil klasifikasi digunakan untuk memprediksi perilaku mahasiswa di semester berikutnya. Diharapkan penelitian ini akan bermanfaat untuk proses pembinaan mahasiswa dalam menempuh kegiatan belajar mengajar di perguruan tinggi.

Kata Kunci—Klasifikasi, Data Mining, Clustering, Fuzzy C-Means.

I. PENDAHULUAN

Dalam dunia pendidikan, khususnya kegiatan belajar mengajar di perguruan tinggi, salah satu aspek penilaian mahasiswa selain indeks prestasi adalah jumlah kehadiran. Keduanya adalah aspek yang saling berkaitan. Menurut Nawawi [1], lembaga pendidikan formal sebagai suatu kesatuan unsur yang mengandung urutan secara berjenjang saling berhubungan dan saling menunjang satu dengan yang lainnya sehingga bergerak secara terpadu dalam mewujudkan mekanisme kerja yang terarah pada suatu tujuan.

Indeks prestasi semester dan tingkat kehadiran mahasiswa di Politeknik Negeri Malang menentukan apakah mahasiswa dapat meneruskan perkuliahan di semester selanjutnya atau tidak. Namun demikian, fenomena yang terjadi adalah masih saja ada mahasiswa yang tidak hadir dalam perkuliahan atau melakukan perilaku membolos. Bentuk sanksi yang diberikan atas ketidakhadiran mahasiswa adalah berupa Surat Peringatan (SP). Ketidakhadiran karena alpa mencapai 18 jam atau lebih akan mendapatkan SP1, 36 jam atau lebih SP2, 47 jam atau lebih SP3. Jumlah ketidakhadiran karena alpa mencapai 56 jam, akan diberikan sanksi putus studi (PS) atau terminal untuk mahasiswa tingkat akhir. Sedangkan jumlah kumulatif ketidakhadiran karena sakit, izin dan alpa mencapai 152 jam dalam satu tahun akademik, maka dinyatakan putus studi untuk mahasiswa tingkat 1 dan 2 (kecuali mengajukan cuti akademik) atau terminal untuk mahasiswa tingkat akhir. Status akademik berdasarkan indeks prestasi semester, mahasiswa dinyatakan lulus semester apabila indeks prestasi semester (IPS) minimal sama dengan 2.00.

Dalam pelaksanaannya, setiap semester selalu ada mahasiswa bermasalah yang mendapatkan SP baik SP1, SP2, SP3 ataupun yang putus studi dikarenakan jumlah ketidakhadiran. Di sinilah peran Dosen Pembina Akademik (DPA), salah satunya yaitu membantu dan membimbing mahasiswa sehingga dapat mengembangkan potensinya secara optimal dan dapat menyelesaikan studi sesuai dengan waktu yang ditentukan. Mahasiswa yang pernah mendapatkan SP dalam satu semester, dan rata-rata jumlah ketidakhadirannya tinggi, kecenderungannya untuk mendapatkan SP di semester selanjutnya adalah cukup tinggi. Oleh karena itu perlu dilakukan pengelompokan, atau klasifikasi terhadap mahasiswa bermasalah tersebut agar dapat dilakukan pembinaan yang lebih lanjut.

Pada penelitian ini akan diterapkan metode clustering yaitu Fuzzy C-Means untuk mengklasifikasikan perilaku mahasiswa berdasarkan jumlah alpa, izin dan sakit dalam dua semester. Hasil klasifikasi tersebut kemudian digunakan sebagai prediksi perilaku mahasiswa di semester selanjutnya.

Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, maka permasalahan dalam penelitian ini adalah:

¹Ridwan Rismanto adalah pengajar di Program Studi Teknik Informatika Politeknik Negeri Malang; email : rismanto@polinema.ac.id

²Imam Fahrur Rozi adalah pengajar di Program Studi Manajemen Informatika Politeknik Negeri Malang; email : imam.rozi@gmail.com

³Arief Prasetyo adalah pengajar di Program Studi Manajemen Informatika Politeknik Negeri Malang; email : ariefpras@gmail.com

1. Bagaimana menerapkan metode clustering Fuzzy C-Means untuk mengklasifikasikan perilaku mahasiswa berdasarkan jumlah ketidakhadiran.
2. Bagaimana memprediksi perilaku mahasiswa berdasarkan hasil klasifikasi.

Batasan Masalah

Batasan masalah dari penelitian ini adalah:

1. Studi kasus adalah mahasiswa Jurusan Teknologi Infomasi Polinema.
2. Metode pengklasifikasian menggunakan Fuzzy C-Means.
3. Platform aplikasi berbasis web.

Tujuan Penelitian

Tujuan penelitian ini adalah, adanya aplikasi untuk klasifikasi dan prediksi perilaku mahasiswa berdasarkan jumlah ketidakhadiran.

II. TINJAUAN PUSTAKA

Clustering

Data Mining dengan teknik clustering, berbeda dengan teknik *association rule mining* dan *classification* dimana kelas data telah ditentukan sebelumnya. *Clustering* melakukan pengelompokan data tanpa berdasarkan kelas data tertentu. Bahkan *clustering* dapat dipakai untuk memberikan label pada kelas data yang belum diketahui itu. Karena itu *clustering* sering digolongkan sebagai metode *unsupervised learning*. Prinsip dari *clustering* adalah memaksimalkan kesamaan antar anggota satu kelas dan meminimumkan kesamaan antar kelas atau *cluster*. *Clustering* dapat dilakukan pada data yang memiliki beberapa atribut yang dipetakan sebagai ruang multidimensi.

Fuzzy C-Means

Dalam [2] disebutkan bahwa Logika fuzzy pertama kali dikembangkan oleh Lotfi A. Zadeh, seorang ilmuwan Amerika Serikat berkebangsaan Iran dari universitas California di Berkeley, melalui tulisannya pada tahun 1965.

Fuzzy secara bahasa diartikan sebagai kabur atau samar-samar. Suatu nilai dapat bernilai benar atau salah secara bersamaan. Dalam fuzzy dikenal derajat keanggotaan yang memiliki rentang nilai 0 (nol) hingga 1 (satu). Berbeda dengan himpunan tegas yang memiliki nilai 1 atau 0 (ya atau tidak).

Logika fuzzy adalah suatu cara yang tepat untuk memetakan suatu ruang input kedalam suatu ruang output, mempunyai nilai kontinyu. Fuzzy dinyatakan dalam derajat dari suatu keanggotaan dan derajat dari kebenaran. Oleh sebab itu sesuatu dapat dikatakan sebagian benar dan sebagian salah pada waktu yang sama [3].

Fuzzy clustering adalah proses menentukan derajat keanggotaan, dan kemudian menggunakannya dengan memasukkannya kedalam elemen data kedalam satu kelompok cluster atau lebih.

Hal ini akan memberikan informasi kesamaan dari setiap objek. Satu dari sekian banyaknya algoritma fuzzy clustering yang digunakan adalah algoritma fuzzy

clustering c means. Vektor dari fuzzy clustering, $V = \{v_1, v_2, \dots, v_c\}$, merupakan sebuah fungsi objektif yang di definisikan dengan derajat keanggotaan dari data X dan pusat cluster V .

Fuzzy C-Means adalah metode clustering yang memungkinkan suatu data menjadi anggota satu atau lebih cluster. Metode yang dikembangkan oleh Dunn [4] dan Bezdek [5] ini banyak digunakan dalam pengenalan pola (*pattern recognition*).

Algoritma fuzzy clustering c means membagi data yang tersedia dari setiap elemen data berhingga lalu memasukkannya kedalam bagian dari koleksi cluster yang dipengaruhi oleh beberapa kriteria yang diberikan. Berikan satu kumpulan data berhingga. $X = \{x_1, \dots, x_n\}$ dan pusat data.

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

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 =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 = ξ
 - 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_k , dengan $k=1,2,\dots,c$ dan $j=1,2,\dots,m$. dimana X adalah variabel fuzzy yang digunakan dan w adalah bobot.

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w * X_{ij})}{\sum_{i=1}^n (\mu_{ik})^w} \dots \dots \dots (3)$$

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

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^w \right) \dots (4)$$
6. Perhitungan fungsi objektif P_t dimana nilai variabel fuzzy X di kurangi dengan dengan pusat cluster V kemudian hasil pengurangannya di kuadratkan lalu masing-masing hasil kuadrat di jumlahkan untuk dikali dengan kuadrat dari derajat keanggotaan μ untuk tiap cluster. Setelah itu jumlahkan semua

nilai di semua cluster untuk mendapatkan fungsi objektif P_i

7. Hitung perubahan matriks partisi:

$$\mu_{ik} = \frac{[\sum_{j=1}^m (x_{ij} - v_{kj})^2]^{-\frac{1}{w-1}}}{\sum_{k=1}^c [\sum_{j=1}^m (x_{ij} - v_{kj})^2]^{-\frac{1}{w-1}}} \dots \dots \dots (5)$$

8. dengan: $i=1,2,\dots,n$ dan $k=1,2,\dots,c$. Untuk mencari perubahan matriks partisi μ_{ik} , pengurangan nilai variabel fuzzy X_{ij} di lakukan kembali terhadap pusat cluster V_{kj} lalu dikuadratkan. Kemudian dijumlahkan lalu dipangkatkan dengan $-1/(w-1)$ dengan bobot, $w=2$ hasilnya setiap data dipangkatkan dengan -1 . Setelah proses perhitungan dilakukan, normalisasikan semua data derajat keanggotaan baru dengan cara menjumlahkan derajat keanggotaan baru $k=1,\dots,c$, hasilnya kemudian dibagi dengan derajat keanggotaan yang baru. Proses ini dilakukan agar derajat keanggotaan yang baru mempunyai rentang antara 0 dan tidak lebih dari 1.

9. Cek kondisi berhenti:

- a) jika: $(|P_t - P_{t-1}| < \xi)$ atau $(t > \maxIter)$ maka berhenti.
b) jika tidak, $t=t+1$, ulangi langkah ke-4.

Dimana μ adalah derajat keanggotaan dari X dan pusat cluster adalah sebuah bagian dari keanggotaan matriks $[\mu]$ adalah jarak dari Euclidean distance dan m adalah parameter fuzzy yang rata-rata derajat keanggotaan dari setiap data derajat keanggotaan tidak lebih besar dari 1,0 [6].

Output dari Fuzzy C-Means merupakan deretan pusat cluster dan beberapa derajat keanggotaan untuk tiap-tiap titik data. Informasi ini dapat digunakan untuk membangun suatu fuzzy inference system.

Status Akademik di Politeknik Negeri Malang

Pelaksanaan prestasi belajar di tiap-tiap Perguruan Tinggi berbeda-beda antara satu dengan lainnya, begitu juga di Politeknik Negeri Malang. Prestasi belajar di Politeknik Negeri Malang dilakukan melalui Ujian Tengah Semester (UTS), nilai tugas serta kuis, dan Ujian Akhir Semester (UAS) yang dilakukan pada setiap semester dengan tujuan untuk menilai apakah mahasiswa telah memahami dan menguasai bahan yang disajikan dalam suatu matakuliah.

Menurut buku Pedoman Akademik Politeknik Negeri Malang tahun 2016, nilai akhir adalah nilai angka untuk masing-masing mata kuliah hasil dari beberapa kali evaluasi mata kuliah yang bersangkutan. Nilai akhir kemudian dikonversikan ke nilai mutu yang berupa nilai huruf dan nilai setara dengan ketentuan sebagai berikut:

Tabel 1. Kategori Nilai Huruf

Nilai Angka	Nilai Huruf	Nilai Setara	Kualifikasi
$80 < N \leq 100$	A	4	Sangat Baik
$73 < N \leq 80$	B+	3,5	Lebih dari Baik
$65 < N \leq 73$	B	3	Baik
$60 < N \leq 65$	C+	2,5	Lebih dari Cukup
$50 < N \leq 60$	C	2	Cukup
$39 < N \leq 50$	D	1	Kurang
$N \leq 39$	E	0	Gagal

Indeks prestasi semester (IPS) adalah nilai rata-rata akhir semester dari gabungan mata kuliah yang ditempuh pada semester yang bersangkutan. Indeks prestasi semester ditentukan dengan rumus sebagai berikut [7] :

$$IPS = \frac{\sum_{i=1}^k (N_i \times S_i)}{\sum_{i=1}^k S_i}$$

IPS : Indeks Prestasi Semester

N_i : Nilai yang dicapai dari mata kuliah ke- i pada semester tersebut.

S_i : Bobot dari mata kuliah ke- i yang ditempuh pada semester tersebut.

$\sum S_i$: Jumlah bobot SKS mata kuliah yang ditempuh pada semester tersebut.

k : Banyaknya mata kuliah yang ditempuh pada semester tersebut.

Output dari Fuzzy C-Means merupakan deretan pusat cluster dan beberapa derajat keanggotaan untuk tiap-tiap titik data. Informasi ini dapat digunakan untuk membangun suatu fuzzy inference system.

Ketidakhadiran Mahasiswa

Beberapa alasan ketidakhadiran mahasiswa antara lain adalah karena sakit, izin dan alpa. Sanksi ketidakhadiran menurut buku Pedoman Akademik Politeknik Negeri Malang tahun 2016 [7] adalah sebagai berikut:

1. Sanksi Akademis

Bentuk sanksi akademis dapat berupa peringatan tertulis hingga putus studi. Peringatan tertulis dapat diberikan secara berjenjang sesuai dengan jumlah waktu ketidakhadiran karena alpa dalam kurun waktu satu semester dengan ketentuan yaitu, jika alpa mencapai 18 jam atau lebih maka mendapatkan Surat Peringatan 1 (SP1). Alpa mencapai 36 jam atau lebih SP2. Alpa 47 jam atau lebih SP3. Jika alpa mencapai 56 jam atau lebih, maka putus studi, untuk mahasiswa tingkat akhir diberikan sanksi terminal.

2. Kompensasi

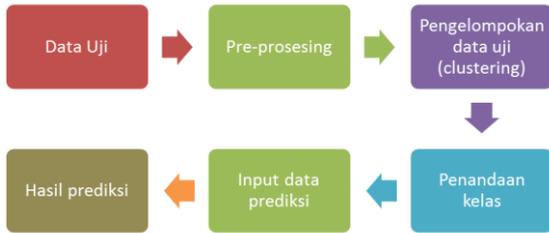
Kompensasi merupakan penggantian waktu yang dibebankan kepada mahasiswa atas ketidakhadirannya karena alpa. Kompensasi ini tidak menghapuskan jumlah ketidakhadiran yang sudah tercatat. Bentuk pelaksanaan kompensasi diatur oleh jurusan, yang diharapkan berimplikasi terhadap peningkatan pengetahuan, keterampilan, kedisiplinan dan loyalitas almamater.

III. METODE PENELITIAN

Penelitian ini menggunakan metode Fuzzy C-Means atau Soft K-Means untuk menentukan cluster dari data yang ada. Cluster ini digunakan untuk mengklasifikasi mahasiswa berdasarkan jumlah absensi (alpa, ijin, sakit) pada semester ganjil dan genap. Sehingga akan didapatkan hasil yaitu kelompok mahasiswa yang rajin mengikuti kuliah dan yang sering tidak mengikuti kuliah.

Tahapan Penelitian

Berikut ini adalah tahapan dari proses yang dilakukan pada penelitian ini untuk clustering data. Pada bagian ini diusulkan metode penelitian dengan fokus pada tahap identifikasi kebutuhan, perencanaan, dan perancangan prototipe seperti yang terlihat pada Gambar 1.



Gambar 1. Metode Penelitian

Tahapan penelitian dimulai dengan mengumpulkan dan merekap data absensi mahasiswa tahun ajaran 2015/2016 semester Ganjil dan Genap. Tahap Pre-prosesing dilakukan dengan menghilangkan data yang tidak lengkap, salah, dan tidak konsisten. Dilanjutkan dengan mengelompokkan data menggunakan algoritma K-Means. Berikutnya menandai kelas hasil pengelompokan. Menginputkan data prediksi pada perangkat lunak.

Analisis Data

Algoritma K-Means merupakan algoritma klasterisasi yang mengelompokkan data berdasarkan titik pusat klaster (centroid) terdekat dengan data. Tujuan dari K-Means adalah pengelompokan data dengan memaksimalkan kemiripan data dalam satu klaster dan meminimalkan kemiripan data antar klaster. Ukuran kemiripan yang digunakan dalam klaster adalah fungsi jarak. Sehingga pemaksimalan kemiripan data didapatkan berdasarkan jarak terpendek antara data terhadap titik centroid.

Tahapan awal yang dilakukan pada proses klasterisasi data dengan menggunakan algoritma K-Means adalah pembentukan titik awal centroid c_j . Pada umumnya pembentukan titik awal centroid dibangkitkan secara acak. Jumlah centroid c_j yang dibangkitkan sesuai dengan jumlah klaster yang ditentukan di awal. Setelah k centroid terbentuk kemudian dihitung jarak tiap data x_i dengan centroid ke- j sampai k , dinotasikan dengan $d(x_i, c_j)$. Terdapat beberapa ukuran jarak yang digunakan sebagai ukuran kemiripan suatu instance data, salah satunya adalah jarak Euclid. Perhitungan jarak Euclidean seperti pada Persamaan berikut ini:

Jika $d(X_i, C_j)$ semakin kecil, kesamaan antara dua unit pengamatan semakin dekat [8]. Syarat menggunakan jarak Euclid adalah jika semua fitur dalam dataset tidak saling berkorelasi. Jika terdapat fitur yang berkorelasi maka menggunakan konsep jarak Mahalanobis.

Kelanjutan dari jarak tersebut dicari yang terdekat sehingga data akan mengelompok berdasarkan centroid yang paling dekat [9]. Tahap berikutnya adalah update titik centroid dengan menghitung rata-rata jarak seluruh data terhadap centroid. Selanjutnya akan kembali lagi ke

proses awal. Iterasi ini akan diulangi terus sampai didapatkan centroid yang konstan artinya titik centroid sudah tidak berubah lagi. Atau iterasi dihentikan berdasarkan jumlah iterasi maksimal yang ditentukan.

Teknik clustering memiliki penggunaan yang luas dan saat ini memiliki kecenderungan yang semakin meningkat seiring dengan jumlah data yang terus berkembang [10]. K-means adalah teknik sederhana untuk analisis klastering. Tujuannya adalah untuk menemukan divisi terbaik entitas n ke dalam kelompok k (disebut cluster), sehingga total jarak antara anggota kelompok dan entroid sesuai, terlepas dari kelompok diminimalkan. Setiap entitas milik cluster dengan mean terdekat. Ini hasil ke partisi ruang data ke Voronoi Sel.

Penelitian dilakukan di Jurusan Teknologi Informasi Politeknik Negeri Malang dengan data sample adalah data absensi mahasiswa tahun ajaran 2015/2016 semester Ganjil dan Genap.

IV. HASIL DAN PEMBAHASAN

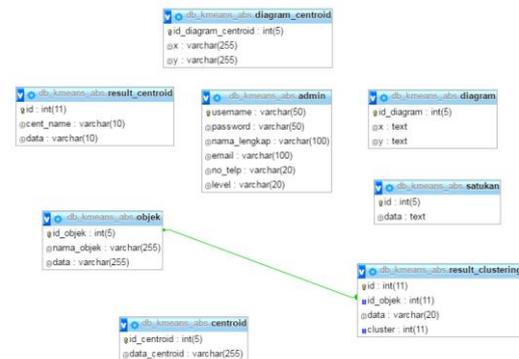
Data Pengujian

Data pengujian yang digunakan adalah berupa tabel yang memiliki komponen penyusun sebagai berikut:

1. Memiliki dua atribut yaitu jumlah AIS (Alpa, Izin, Sakit) pada semester genap dan jumlah AIS pada semester ganjil.
2. Jumlah instance adalah 30.

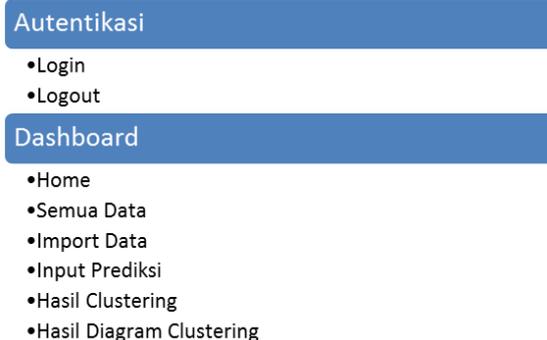
Desain Aplikasi

1. Desain database



Gambar 2. Desain Database

2. Sitemap aplikasi



Gambar 3. Sitemap Aplikasi

Pembahasan

Berikut ini adalah langkah-langkah penerapan metode clustering pada data pengujian:

1. Menentukan koordinat titik tengah setiap cluster. Penentuan cluster dibuat 2 buah obyek dengan 2 atribut. Metode Clustering dengan algoritma K-Means akan menghasilkan 2 cluster berdasarkan class jumlah AIS semester genap dan ganjil. Pengesetan nilai awal tengah dengan menentukan titik tengah (centroid) dari cluster secara acak seperti pada table berikut ini:

Tabel 2. Cluster centroids

Cluster 0		Cluster 1	
4	14	49	51

2. Penentuan nilai dari cluster-cluster tersebut untuk dijadikan acuan untuk melakukan perhitungan pada setiap baris tabel data pengujian. Contoh penentuan jarak obyek ke centroid, pengujian dilakukan pada data mahasiswa Yenny Chandrawati Salim (objek 1) dengan jumlah \sum AIS genap = 6, dan \sum AIS ganjil = 27, dan mahasiswa Mahrus Ali (objek 2) dengan jumlah \sum AIS genap = 65, dan \sum AIS ganjil = 45 dengan titik centroid:

Objek 1:

$$\text{centroid 0: } \sqrt{(4 - 6)^2 + (14 - 27)^2} = 13.153$$

$$\text{centroid 1: } \sqrt{(49 - 6)^2 + (51 - 27)^2} = 49.244$$

Objek 2:

$$\text{centroid 0: } \sqrt{(4 - 65)^2 + (14 - 45)^2} = 68.425$$

$$\text{centroid 1: } \sqrt{(49 - 65)^2 + (51 - 45)^2} = 17.088$$

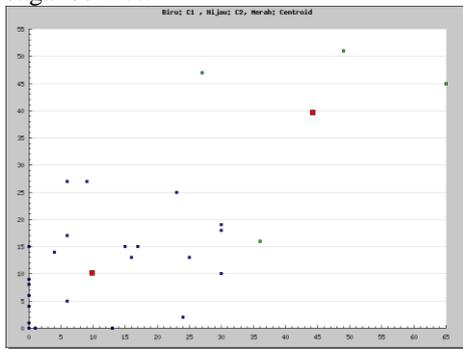
Dari hasil perhitungan maka didapatkan hasil seperti pada tabel berikut ini:

Tabel 3. Penentuan jarak obyek

Objek	Cluster 0	Cluster 1
1	13.153	49.244
2	68.425	17.088

Dari hasil maka diperoleh, objek 1 lebih dekat dengan cluster 0 sedangkan objek 2 lebih dekat dengan cluster 1. Proses yang sama akan dilakukan untuk semua data pengujian, kemudian akan diubah titik centroid yaitu rata-rata jarak dari seluruh objek, kemudian dihitung kembali jarak dari objek ke centroid tersebut. Begitu seterusnya hingga pada iterasi tertentu, objek-objek tersebut tidak berubah lagi keanggotaan clusternya.

3. Hasil clustering setelah dilakukan iterasi ke-3 sebagai berikut:



Gambar 4. Hasil clustering

4. Titik centroid akhir dari clustering adalah sebagai berikut:

Tabel 4. Centroids

Cluster	x	y
0	9.81	10.12
1	44.25	39.75

Implementasi Perangkat Lunak

1. Halaman home



Gambar 5. Halaman home

2. Import data dan manajemen data



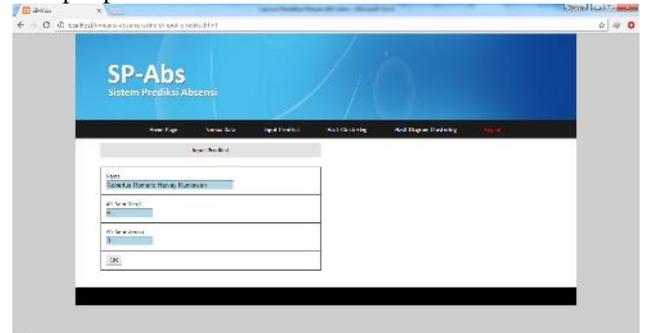
Gambar 6. Manajemen data

3. Hasil clustering



Gambar 7. Hasil clustering

4. Input prediksi



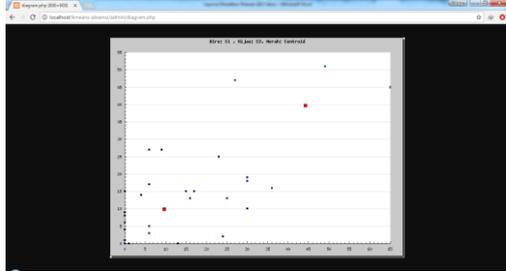
Gambar 8. Input prediksi

5. Hasil prediksi



Gambar 9. Hasil prediksi

8. Grafik hasil clustering



Gambar 10. Grafik hasil clustering

Hasil pengujian

1. Data uji yang digunakan pada penelitian ini adalah data absensi mahasiswa pada semester genap dan ganjil tahun ajaran 2015/2016 dapat dilihat pada tabel 4.

Tabel 5. Data uji

NO.	NIM.	N A M A	REKAP. ABSENSI											
			SEMESTER GANJIL 2015/2016				AKHIR GENAP 2015/2016				SEM 1 & 2			
			A	I	S	ST1	A	I	S	ST1	A	I	S	ST1
1	1431140141	Ahmad Syah Rizal	6,0	0,0	0,0	6,0	-	10,0	0,0	17,0	27,0	33,0	-	
2	1431140107	Billy Yoga Rinaldy	0,0	6,0	0,0	6,0	-	1,0	0,0	4,0	5,0	11,0	-	
3	1431140066	Davin Yusuf Wicaksana	0:15	4:0	26:0	30:15	-	12:0	0:0	6:0	18:0	48:15	-	
4	1431140109	Dimas Andre Iswahyudi	1:0	27:0	21:0	49:0	-	5:0	15:0	31:0	51:0	100:0	-	
5	1431140096	Dodi Subakti	0:0	0:0	0:0	0:0	-	4:0	0:0	0:0	4:0	4:0	-	
6	1431140040	Dwi Inayatus'diyah	12:0	19:0	20:0	51:0	-	1:0	13:0	34:0	48:0	99:0	-	
7	1431140126	Eko Putra Syah Warman	1:15	0:0	0:0	1:15	-	0:0	0:0	0:0	0:0	1:15	-	
8	1431140004	Evita Ajeng Pratiwi	1:0	0:0	3:0	4:0	-	10:0	0:0	4:0	14:0	18:0	-	
9	1431140061	Fatimah Nur Arofah Jumrotin	0:0	1:0	12:0	13:0	-	0:0	0:0	0:0	0:0	13:0	-	
10	1431140100	Furqon Hanif	22:0	0:0	5:0	27:0	SP1	31:0	3:0	13:0	47:0	74:0	SP1	
11	1431140081	Idham Syahrir Metareum	0:15	0:0	0:0	0:15	-	1:0	0:0	0:0	1:0	1:15	-	
12	1431140032	Ilsa Putri Nahar	0:0	0:0	0:0	0:0	-	2:0	0:0	6:0	8:0	8:0	-	
13	1431140034	Ilham Akbar	1:15	15:0	0:0	16:15	-	7:0	6:0	0:0	13:0	29:15	-	
14	1431140154	Intan Finda Ameilia	0:0	0:0	0:0	0:0	-	0:0	6:0	9:0	15:0	15:0	-	
15	1431140028	Laras Putrianti Nabila	0:0	0:0	0:0	0:0	-	0:0	0:0	0:0	0:0	0:0	-	
16	1431140085	Mahrus Ali	0:0	18:0	47:0	65:0	-	0:0	0:0	45:0	110:0	-		
17	1431140155	Mas Agus Setyo Purwanto	0:0	0:0	0:0	0:0	-	0:0	0:0	0:0	0:0	0:0	-	
18	1431140013	Maulidya Yulianingtyas	0:0	3:0	6:0	9:0	-	10:0	0:0	17:0	27:0	36:0	-	
19	1431140058	Moch. Rosid Noviansyah	0:0	0:0	6:0	6:0	-	5:0	6:0	6:0	17:0	23:0	-	
20	1331140037	Muhamad Aris Putra Utama	0:0	3:0	27:0	30:0	-	0:0	0:0	19:0	19:0	49:0	-	
21	1431140130	Muhammad Rifat	12:0	0:0	5:0	17:0	-	15:0	0:0	0:0	15:0	32:0	-	
22	1431140033	Mutiara Heka Widyaningtyas	0:0	0:0	0:0	0:0	-	0:0	9:0	0:0	9:0	9:0	-	
23	1431140118	Prasetya Putra Pratama	0:0	0:0	0:0	0:0	-	0:0	0:0	0:0	0:0	0:0	-	
24	1431140079	Qiqi Maghfirah	0:0	0:0	0:0	0:0	-	2:0	4:0	0:0	6:0	6:0	-	
25	1431140099	Rama Adi Nugroho	12:0	6:0	7:0	25:0	-	6:0	0:0	7:0	13:0	38:0	-	
26	1431140010	Reza Dwi Saputra	0:0	0:0	0:0	0:0	-	0:0	0:0	0:0	0:0	0:0	-	
27	1431140123	Rifqi Dwi Putranto	0:0	12:0	18:0	30:0	-	4:0	0:0	6:0	10:0	40:0	-	
28	1431140082	Robiatul Maulidiah	0:0	0:0	23:0	23:0	-	1:0	6:0	18:0	25:0	48:0	-	
29	1431140023	Rochman Syahrul Huda	5:0	12:0	7:0	24:0	-	2:0	0:0	0:0	2:0	26:0	-	
30	1431140074	Yenny Chandrawati Salim	0:0	6:0	0:0	6:0	-	3:0	0:0	24:0	27:0	33:0	-	
31	1431140146	Yoga Setya Putra	4:15	6:0	5:0	15:15	-	2:0	1:0	12:0	15:0	30:15	-	
Nilai Bobot Rata-rata Kelas (NBRK)														

2. Dari data uji tersebut selanjutnya dilakukan tahap pre-prosesing agar dapat diproses oleh aplikasi seperti yang terlihat pada tabel 5.
3. Dari data tersebut kemudian dilakukan import kedalam aplikasi yang kemudian dilakukan clustering dengan hasil yang didapatkan sebagaimana diperlihatkan pada tabel 6.
4. Dari data hasil clustering dapat ditentukan kelompok:

Cluster 1: rata-rata jumlah AIS semester genap dan ganjil: 19.51851852

Cluster 2: rata-rata jumlah AIS semester genap dan ganjil: 84

Maka cluster 1 adalah kelompok mahasiswa yang relatif aman, tidak memerlukan pengawasan.

Sedangkan cluster 2 adalah kelompok mahasiswa yang memerlukan pengawasan.

Tabel 6. Hasil pre-prosesing

No	Nama	Data	x	y
1	Ahmad Syah Rizal	6,27	6	27
2	Billy Yoga Rinaldy	6,5	6	5
3	Davin Yusuf Wicaksana	30,18	30	18
4	Dimas Andre Iswahyudi	49,51	49	51
5	Dodi Subakti	0,4	0	4
6	Dwi Inayatus'diyah	51,48	51	48
7	Eko Putra Syah Warman	1,0	1	0
8	Evita Ajeng Pratiwi	4,14	4	14
9	Fatimah Nur Arofah Jumrotin	13,0	13	0
10	Furqon Hanif	27,47	27	47
11	Idham Syahrir Metareum	0,1	0	1
12	Ilsa Putri Nahar	0,8	0	8
13	Ilham Akbar	16,13	16	13
14	Intan Finda Ameilia	0,15	0	15
15	Laras Putrianti Nabila	0,0	0	0
16	Mahrus Ali	65,45	65	45
17	Mas Agus Setyo Purwanto	0,0	0	0
18	Maulidya Yulianingtyas	9,27	9	27
19	Moch. Rosid Noviansyah	6,17	6	17
20	Muhamad Aris Putra Utama	30,19	30	19
21	Muhammad Rifat	17,15	17	15
22	Mutiara Heka Widyaningtyas	0,9	0	9
23	Prasetya Putra Pratama	0,0	0	0
24	Qiqi Maghfirah	0,6	0	6
25	Rama Adi Nugroho	25,13	25	13
26	Reza Dwi Saputra	0,0	0	0
27	Rifqi Dwi Putranto	30,10	30	10
28	Robiatul Maulidiah	23,25	23	25
29	Rochman Syahrul Huda	24,2	24	2
30	Yenny Chandrawati Salim	6,27	6	27
31	Yoga Setya Putra	15,15	15	15

Tabel 7. Hasil clustering

Objek	Data 1	Data 2	Cluster 1	Cluster 2
Objek 1	6	3	OK	Null
Objek 2	36	16	Null	OK
Objek 3	15	15	OK	Null
Objek 4	6	27	OK	Null
Objek 5	24	2	OK	Null
Objek 6	23	25	OK	Null
Objek 7	30	10	OK	Null
Objek 8	0	0	OK	Null
Objek 9	25	13	OK	Null
Objek 10	0	6	OK	Null
Objek 11	0	0	OK	Null
Objek 12	0	9	OK	Null
Objek 13	17	15	OK	Null
Objek 14	30	19	OK	Null
Objek 15	6	17	OK	Null
Objek 16	9	27	OK	Null
Objek 17	0	0	OK	Null
Objek 18	65	45	Null	OK
Objek 19	0	0	OK	Null
Objek 20	0	15	OK	Null
Objek 21	16	13	OK	Null
Objek 22	0	8	OK	Null
Objek 23	0	1	OK	Null
Objek 24	27	47	Null	OK
Objek 25	13	0	OK	Null
Objek 26	4	14	OK	Null
Objek 27	1	0	OK	Null
Objek 28	0	4	OK	Null
Objek 29	49	51	Null	OK
Objek 30	30	18	OK	Null
Objek 31	6	5	OK	Null

5. Dari hasil pengelompokan, maka dapat dilakukan uji coba untuk prediksi perilaku mahasiswa (dalam pengawasan atau tidak dalam pengawasan).

Untuk ujicoba ini dilakukan dengan mengambil sampel data 4 mahasiswa kelas 3 prodi Teknologi Informasi, pada semester ganjil dan genap tahun ajaran 2015/2016 dengan menginputkan AIS semester ganjil dan genap.

Hasil prediksi menempatkan mahasiswa kedalam dua kemungkinan kelompok, yaitu:

- Kelompok 1 (Tidak dalam pengawasan)
- Kelompok 2 (Dalam pengawasan)

Hasil prediksi kemudian dibandingkan dengan data komparasi yaitu data AIS mahasiswa tersebut pada tahun berikutnya, yaitu semester ganjil tahun ajaran 2016/2017.

Tabel 8. Hasil uji coba prediksi

NIM	Uji Coba		Hasil prediksi	Komparasi
	AIS Ganjil 15/16	AIS Genap 15/16		AIS Mid Ganjil 16/17
1341180027	49	24	Kelompok 2 (Dalam Pengawasan)	52
1341180123	0	18	Kelompok 1 (Tidak Dalam Pengawasan)	1
1341180144	64	76	Kelompok 2 (Dalam Pengawasan)	55
1341180117	6	3	Kelompok 1 (Tidak Dalam Pengawasan)	11

Dari tabel diatas dapat disimpulkan bahwa mahasiswa Evan Cahya Septian Pratama dengan AIS ganjil 49 dan AIS genap 24, masuk dalam kelompok 2 (Dalam pengawasan). Setelah dilakukan komparasi pada tahun ajaran berikutnya, ternyata memang jumlah AIS nya relatif besar, yaitu 52. Maka benar adanya mahasiswa tersebut membutuhkan pengawasan.

Sedangkan mahasiswa Afwika Chori Quwattika dengan AIS ganjil 0 dan AIS genap 18, masuk dalam kelompok 1 (Tidak dalam pengawasan). Setelah dilakukan komparasi pada tahun ajaran berikutnya, ternyata memang jumlah AIS nya relatif kecil, yaitu 1. Maka benar adanya mahasiswa tersebut tidak membutuhkan pengawasan.

V. KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan penelitian yang dilakukan, dapat disimpulkan bahwa penggunaan metode clustering Fuzzy C-Means atau Soft K-Means dapat digunakan untuk mengelompokkan mahasiswa berdasarkan jumlah AIS (Alpa, Izin, Sakit) selama dua semester yaitu:

- Kelompok 1, dengan rata-rata jumlah AIS 19.51851852 merupakan kelompok yang tidak memerlukan pengawasan.
- Kelompok 2, dengan rata-rata jumlah AIS 84 merupakan kelompok yang memerlukan pengawasan.

Dari hasil pengelompokkan diatas dapat dilakukan prediksi perilaku mahasiswa (memerlukan atau tidak memerlukan pengawasan) dengan hasil ujicoba 4 mahasiswa:

- Mahasiswa 1: AIS genap 49, ganjil 24 (T.A. 2015/2016) → Kelompok 2 (dalam pengawasan). Hasil komparasi T.A. 2016/2017: AIS: 52 (tepat).
- Mahasiswa 2: AIS genap 0, ganjil 18 (T.A. 2015/2016) → Kelompok 1 (tidak dalam pengawasan). Hasil komparasi T.A. 2016/2017: AIS: 1 (tepat).
- Mahasiswa 3: AIS genap 64, ganjil 76 (T.A. 2015/2016) → Kelompok 2 (dalam pengawasan). Hasil komparasi T.A. 2016/2017: AIS: 55 (tepat).
- Mahasiswa 4: AIS genap 6, ganjil 3 (T.A. 2015/2016) → Kelompok 1 (tidak dalam pengawasan). Hasil komparasi T.A. 2016/2017: AIS: 11 (tepat).

Saran

Saran pengembangan sistem untuk di masa yang akan datang untuk semakin meningkatkan fitur maupun pelayanan yang ditawarkan, antara lain data uji coba yang lebih banyak, variabel yang lebih bervariasi (misal, status Lulus/Lulus Percobaan/SP/ dan sebagainya) agar hasil yang didapatkan lebih akurat.

DAFTAR PUSTAKA

- [1] H. Nawawi, *Organisasi Sekolah dan Pengelolaan Kelas Sebagai Lembaga Pendidikan*. Jakarta: Gunung Agung, 1982.
- [2] R. Munir, *Matematika Diskrit*. Bandung: Informatika, 2005.
- [3] S. Kusumadewi, *Analisis dan Desain Sistem Fuzzy Menggunakan Tool Box Matlab*. Yogyakarta: Graha Ilmu, 2003.
- [4] J. C. Dunn, "A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact, Well Separated Clusters," *J. Cybern.*, vol. 3, no. 3, pp. 32–57, 1973.
- [5] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York: Plenum Press, 1981.
- [6] P. Valarmathie, M. Srinath, T. Ravichandran, and K. Dinakaran, "Hybrid Fuzzy C-Means Clustering Technique for Gene Expression Data," *Int. J. Res. Rev. Appl. Sci.*, vol. 1, no. 1, pp. 33–37, 2009.
- [7] Polinema, *Pedoman Akademik Tahun 2016*. Malang: Politeknik Negeri Malang, Jurusan Teknologi Informasi, 2016.
- [8] B. S. Duran and P. L. Odell, *Cluster Analysis: A Survey*. New York: Springer-Verlag, 1974.
- [9] Y. Agusta, "K-Means–Penerapan, Permasalahan dan Metode Terkait," *J. Sist. dan Inform.*, vol. 3, no. 1, pp. 47–60, 2007.
- [10] R. Sharma, M. A. Alam, and A. Rani, "K-Means Clustering in Spatial Data Mining using Weka Interface," in *IJCA Proceedings on International Conference on Advances in Communication and Computing Technologies 2012*, 2012, pp. 26–30.

