

CLUSTERING KINERJA AKADEMIS MAHASISWA MENGUNAKAN FUZZY C-MEANS

Nurissaidah Ulinnuha¹

Program Studi Matematika Fakultas Sains dan Teknologi
Universitas Islam Negeri Sunan Ampel Surabaya¹

nuris.ulinnuha@uinsby.ac.id¹

Abstrak

Dunia pendidikan khususnya di universitas sering mengalami masalah dengan tidak tercapainya visi dan misi institusi. Pemantauan kinerja akademik mahasiswa menjadi peranan yang sangat penting di lembaga-lembaga perguruan tinggi sehingga salah satu tolak ukurnya adalah sistem monitoring dan evaluasi yang bersifat obyektif dan subyektif. Salah satu cara untuk memonitor perkembangan akademik mahasiswa adalah dengan mengelompokkan mahasiswa berdasarkan kesamaan karakteristik dan kemampuan. Pada penelitian kali ini parameter-parameter yang digunakan berupa nilai-nilai indeks prestasi semester 1 sampai 3 dan nilai indeks prestasi kumulatif dari data mahasiswa program studi matematika. Dari data tersebut akan di cluster menggunakan Algoritma Fuzzy C-Means. Clustering menggunakan Fuzzy C-Means pada semua parameter-parameter nantinya akan terkelompok ke dalam mahasiswa dengan nilai sangat memuaskan, memuaskan dan kurang memuaskan. Hal ini bertujuan agar dapat mempermudah memonitoring mahasiswa yang berada pada cluster yang sama. Dengan adanya sistem monitoring akan membantu mahasiswa untuk memperbaiki hasil akademisnya. Dari hasil uji coba dengan menggunakan sistem Fuzzy C-Means Clustering ke dalam 3 cluster dengan 16 data yang di training didapatkan akurasi RMSE 2.76×10^{-15} dan 4 data yang di testing didapatkan akurasi RMSE 0.1165.

Kata kunci: Kinerja Akademik, Clustering, Fuzzy C-Means

Abstract

Education, especially at the university, often have problems with the vision and mission of the institution that not achieved. Monitoring students academic performance becomes a very important role in the institutions of higher education so that one of the benchmark is a monitoring and evaluation system that is both objective and subjective. One way to monitor the academic progress of students is to classify students based on similar characteristics and capabilities. In this research, the parameters used are the values of grade point average (GPA) from 1st to 3rd semester and the value of the cumulative GPA mathematics student's data. From these data will be in grouped using Fuzzy C-Means Algorithm. Clustering using Fuzzy C-Means on all parameters will be grouped into student with a value very satisfactory, satisfactory or unsatisfactory. It aims to facilitate the monitoring of students who are in the same cluster. This monitoring system will help students to improve academic results. From the test results by using a system Fuzzy C-Means Clustering into 3 clusters with 16 data, obtained in training RMSE accuracy of 2.76×10^{-15} and from four data in testing obtained the accuracy RMSE 0.1165.

Keywords : Academic performance, Clustering, Fuzzy C-Means

1. PENDAHULUAN

Pendidikan di era globalisasi sekarang ini merupakan hal yang penting karena dengan adanya pendidikan akan memperluas wawasan seseorang. Perkembangan ilmu pengetahuan sendiri harus disertai dengan kemampuan sumber daya manusianya. Salah satu faktor yang dapat mempersiapkan sumber daya manusia yang handal dalam menghadapi era teknologi dan globalisasi adalah universitas. Universitas merupakan lembaga pendidikan tertinggi yang mampu merubah peradaban manusia dengan keilmuan sehingga menjadikan sumber daya manusianya sendiri mampu dalam dunia industri,

lembaga-lembaga pemerintahan dan lembaga penelitian[1].

Dunia pendidikan khususnya di universitas sering mengalami masalah dengan tidak tercapainya visi dan misi institusi. Banyak faktor yang menyebabkan tidak tercapainya target output yang diinginkan oleh institusi. Faktor yang paling signifikan berasal dari faktor internal yang meliputi sumber daya manusia, metode pengajaran, dan terkadang kurikulum yang telah dirumuskan tidak memenuhi standarisasi [1].

Di Universitas, pemantauan kinerja akademik mahasiswa menjadi peranan yang sangat penting di lembaga-lembaga perguruan tinggi. Metode monitoring

dan evaluasi yang bersifat obyektif dan subyektif menjadi tolak ukur dalam proses pembelajaran dengan hasil target output adalah dapat memberikan bimbingan konseling pada mahasiswa agar dapat memperbaiki nilai. Dengan banyaknya jumlah mahasiswa dan minimnya tenaga pengajar (dosen), membuat sistem monitoring adalah sistem yang terbaik untuk memantau dan mengevaluasi proses akademik mahasiswa.

Di setiap semesternya, mahasiswa mendapatkan nilai indeks prestasi semester yang nilai ini menjadi tolak ukur dari perkembangan mahasiswa dalam semester tersebut. nilai indeks prestasi semester ini di jadikan parameter tersendiri untuk menentukan kualitas dan pemantauan dini pada mahasiswa untuk memperbaiki nilai.

Pengelompokkan mahasiswa merupakan hal yang penting karena bertujuan untuk memonitor perkembangan akademik mahasiswa. Pengelompokkan atau yang disebut dengan analisa *clustering* adalah suatu analisa yang berguna untuk mengelompokkan objek-objek berdasarkan kemiripan peubah-peubah yang diamati [4][9].

Salah satu metode yang cocok untuk pemetaan mahasiswa berdasarkan kesamaan karakteristik dan kemampuan adalah metode *Fuzzy C-Means Clustering*, seperti dalam jurnal-jurnal sebelumnya *Fuzzy C-Means Clustering* digunakan sebagai pemetaan evaluasi penilaian dan kualifikasi mahasiswa, lalu digunakan untuk menentukan bidang tugas akhir mahasiswa berdasarkan nilai, digunakan sebagai penentuan nilai akhir kuliah, dan digunakan sebagai klasifikasi sinyal EEG[7][8][9]. Karena dalam penelitian-penelitian sebelumnya metode *Fuzzy logic* dan analisis regresi kurang mampu untuk mengatasi permasalahan tersebut [4][9]. Sehingga *Fuzzy C-Means* merupakan metode yang cerdas untuk mengclusterkan kinerja akademis mahasiswa berdasarkan kemampuan dan untuk dapat memantau mahasiswa agar dapat memperbaiki nilai akademisnya.

2. TINJAUAN PUSTAKA

2.1 Logika Fuzzy

Fuzzy adalah salah satu konsep logika matematika yang pertama kali dikenalkan oleh Lotfi A Zadeh pada tahun 1965 [3]. Zadeh memberikan definisi tentang himpunan *fuzzy* sebagai berikut:

Jika X adalah koleksi obyek yang di notasikan oleh x , maka suatu himpunan *fuzzy* A dalam X adalah suatu himpunan pasangan berurutan:

$$A = \{(x, \mu_A(x)) | x \in X\} \quad (1)$$

Dengan $\mu_A(x)$ adalah derajat keanggotaan x di A yang memetakan X ke ruang keanggotaan M yang terletak pada rentang $(0,1)$ [2].

2.2 Fuzzy Clustering

Clustering adalah salah satu dari teknik *Fuzzy* yang digunakan untuk menentukan *cluster* optimal dalam suatu ruang vector dengan berdasarkan kesamaan karakteristik diantara obyek-obyek tersebut [5].

2.2.1 Fuzzy C-Means (FCM)

Fuzzy C-Means (FCM) adalah suatu teknik *pengclusteraan* data yang dalam suatu *cluster* keberadaan tiap-tiap titik data ditentukan oleh derajat keanggotaan. Algoritma *Fuzzy C-Means* adalah sebagai berikut [2]:

- Menginputkan data yang dicluster X , dalam bentuk matrix berukuran $n \times m$ dengan n = jumlah sampel data dan m = atribut setiap data. X_{ij} yang merupakan data sampel ke- i ($i=1,2,\dots,n$) dan atribut ke- j ($j=1,2,\dots,m$).
- Menentukan parameter yang dibutuhkan, meliputi:
 - Jumlah *cluster* : c
 - Pangkat : w
 - Maksimum iterasi : $MaxIter$
 - Error terkecil : ϵ
 - Fungsi objektif awal : $P_0 = 0$
 - Iterasi awal : $t=1$
- Membangkitkan bilangan *random* μ_{ik} , dengan $i = 1,2,\dots,n$ dan $k = 1,2,\dots,c$ yang merupakan elemen-elemen matriks partisi awal U . setelah itu, Menghitung jumlah setiap kolom dengan persamaan :

$$Q_i = \sum_{k=1}^c \mu_{ik} \quad (2)$$
 dengan $j = 1,2,\dots,n$ sehingga untuk menghitung matriks *random* menggunakan persamaan:

$$\mu_{ik} = \frac{\mu_{ik}}{Q_i} \quad (3)$$
- Menghitung pusat *cluster* ke- k dengan persamaan:

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w * X_{ij})}{\sum_{i=1}^n (\mu_{ik})^w} \quad (4)$$
- Menghitung fungsi objektif pada iterasi ke- t , P_t dengan persamaan:

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^w \right) \quad (5)$$
- Menghitung perubahan matriks partisi dengan persamaan:

$$\mu_{ik} = \frac{\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{-1}{w-1}}}{\sum_{k=1}^c \left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{-1}{w-1}}} \quad (5) \text{ dimana } i = 1,2,\dots,n; \text{ dan } k = 1,2,\dots,c.$$

- Mengecek kondisi sampai berhenti dengan persamaan:

$$(|P_t - (P_t - 1)| < \epsilon) \quad (6)$$
 atau $(t > maxIter)$ apabila berhenti. dan apabila tidak : $t = t + 1$, ulangi langkah ke-d.

2.3 Monitoring

Monitoring adalah proses rutin pengumpulan data dan pengukuran kemajuan atas objektif program atau memantau perubahan yang fokus pada proses dan keluaran. Monitoring melibatkan perhitungan atas apa yang kita lakukan dan pengamatan atas kualitas dari layanan yang kita berikan [10].

2.4 Evaluasi

Evaluasi adalah suatu proses sistemik untuk mengetahui tingkat keberhasilan suatu program.

Dalam bidang pendidikan, evaluasi merupakan sebuah proses pengumpulan data untuk menentukan sejauh mana, dalam hal apa, dan bagian mana tujuan pendidikan sudah tercapai [10].

3. METODE PENELITIAN

3.1 Metode Pengumpulan Data

a) Jenis dan Sumber Data

Data yang digunakan untuk pembuatan penelitian ini berasal dari nilai-nilai Indeks Prestasi Semester Mahasiswa Program Studi Matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Ampel Surabaya. Jumlah data keseluruhan yang diperoleh sebanyak $20 \times 4 = 80$ data, dari 20 mahasiswa dengan 4 parameter yaitu indeks prestasi semester satu sampai tiga dan indeks prestasi kumulatif. Data inilah yang nanti akan digunakan sebagai data *training* dan data *testing*.

b) Teknik Pengumpulan Data

Data diperoleh dari akademik Fakultas Sains dan Teknologi. Sampel data sebagai berikut.

Tabel 1
Sampel Data Mahasiswa

Mahasiswa Ke-	IPS 1	IPS 2	IPS 3	IPK
1	3.50	3.53	3.33	3.45
2	3.20	3.16	2.99	3.11
3	3.18	3.03	2.88	3.03
4	3.37	3.41	3.76	3.51
5	3.45	3.25	3.21	3.30

3.2 Metode Analisa Data

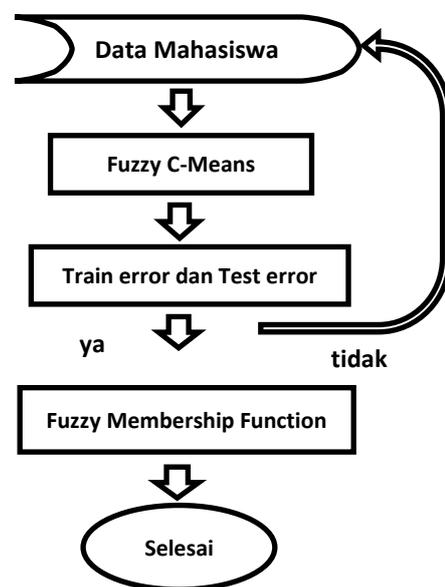
Input variabel yang digunakan dalam penelitian ini adalah indeks prestasi semester satu sampai tiga dan indeks prestasi kumulatif dengan rentang nilai [0.00 – 4.00] dan target yang ingin dicapai adalah pengelompokan mahasiswa berdasarkan kualitas dan kemampuan mahasiswa. *FuzzyClustering* adalah salah satu teknik dalam *Fuzzy* yang berfungsi untuk menentukan *cluster* optimal dalam suatu ruang vector yang didasarkan pada bentuk normal *Euclidian* untuk jarak antar vector [1].

Dalam analisa data dilakukan beberapa satu tahap yaitu :Penggunaan *FuzzyC-Means Clustering* untuk mengelompokkan mahasiswa dengan menggunakan MATLAB.

Langkah-langkah dari proses ini dijelaskan sebagai berikut:

- Inputkan data mahasiswa. Data nantinya dibagi menjadi dua bagian, yaitu data *training* dan data *testing*. Data yang akan di *training* kurang lebih 80 % dan data yang akan di *testing* kurang lebih 20 %.
- Fuzzy C-Means* (FCM) digunakan untuk menentukan *cluster* optimal dari data Mahasiswa Program Studi Matematika. Tahapan-tahapan dari *Fuzzy C-Means* (FCM) sebagai berikut:

- Tahap pertama menginputkan data yang akan di *cluster* berupa sebuah matriks $n \times m$ seperti dalam persamaan 2.2.
- Tahap kedua, dengan menggunakan function dari MATLAB, MATLAB akan memproses persamaan 2.4, 2.5, 2.6, 2.7, dan 2.8. persamaan ini akan diproses secara berulang-ulang sampai kondisi berhenti.
- Tahap ketiga, setelah kondisi berhenti akan dihasilkan berapa kali iterasi pada proses *fuzzy C-Means*, nilai pusat *cluster* dan fungsi keanggotaan dari data mahasiswa Program Studi Matematika.
- Tahap Keempat, akan ditampilkan grafik dari *fuzzy C-Means* berupa titik-titik *cluster* dan pusat-pusat *cluster*.



Gambar 1. Diagram Aliran Data

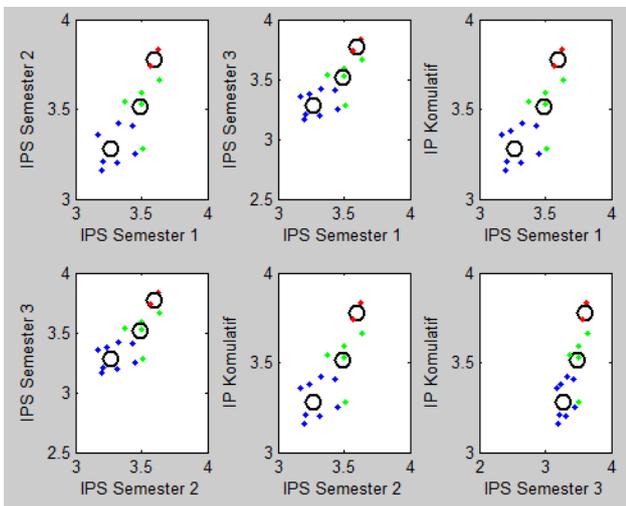
- Setelah data-data diproses dengan menggunakan *fuzzy C-Means*, data-data dijadikan sebagai data pelatihan dan data pengujian dengan melakukan perhitungan RMSE.
- Perbandingan akurasi
Dari hasil pengujian memperhatikan beberapa kombinasi jumlah data yang berbeda-beda, kemudian dilakukan proses perbandingan RMSE. Error terkecil menunjukkan pengclasteran data mendekati akurat.

4. PEMBAHASAN DAN HASIL

Fungsi keanggotaan *Fuzzy* dibuat dari *cluster-cluster* mahasiswa berdasarkan Nilai Indeks Prestasi Semester dari semester satu sampai semester tiga. Algoritma *FuzzyC-Means Clustering* digunakan untuk mengelompokkan mahasiswa berdasarkan kesamaan karakteristik.

4.1 FuzzyC-Means Clustering

Dalam tinjauan pustaka telah dibahas mengenai algoritma *FuzzyC-Means Clustering*. Peranan MATLAB dengan algoritma *FuzzyC-Means Clustering* mengelompokkan mahasiswa ke dalam tiga *cluster* diantaranya mahasiswa dengan nilai kurang memuaskan, mahasiswa dengan nilai memuaskan, dan mahasiswa dengan nilai sangat memuaskan. Untuk *training* dan *testing* diambil sampel secara random dari 20 data mahasiswa dengan 4 parameter sebanyak kurang lebih 80%. Selain data *training* sisa kurang lebih 20% dari data set digunakan sebagai data *testing*. Pada penelitian ini nilai konstanta $m = 2, \epsilon = 0.01$ dan $k = 200$. Dengan parameter-parameter tersebut hasil empat kali *training* dan empat kali *testing* (uji data), telah didapatkan salah satu pengujian yang mendekati akurat. Salah satu tampilan *cluster-cluster* dan pusat *cluster* data direpresentasikan dalam bentuk grafik sebagai berikut :



Gambar 2. Plot Fuzzy C-Means

Hasil dari *FuzzyC-Means Clustering* direpresentasikan pada Tabel2.

Tabel 2 Deskripsi *FuzzyC-Means Clustering* berdasarkan nilai

No Cluster	1	2	3	
Nama Cluster	SM	M	KM	
Pusat Cluster	IPS 1	3.5984	3.4864	3.2656
	IPS 2	3.7747	3.5160	3.2790
	IPS 3	3.7346	3.3667	3.0354
	IPK	3.7007	3.4517	3.1888
Jumlah Cluster	2	5	9	

Keterangan :

- SM = Sangat Memuaskan
- M = Memuaskan
- KM = Kurang Memuaskan

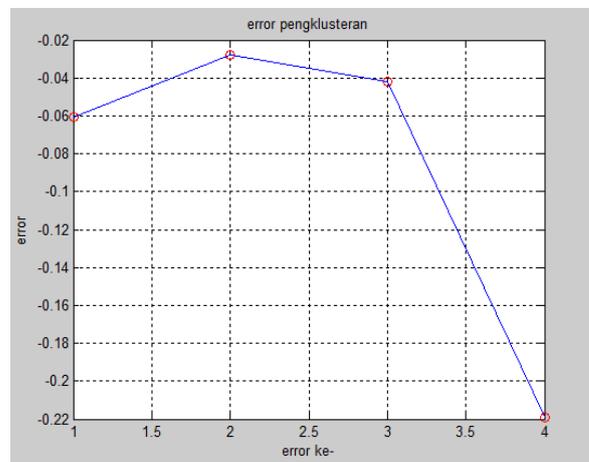
Berdasarkan representasi Gambar 2, dari 16 mahasiswa yang ditraining menghasilkan 2

mahasiswa dalam *cluster* mahasiswa dengan nilai sangat memuaskan, 5 mahasiswa dalam *cluster* mahasiswa dengan nilai memuaskan, dan 9 mahasiswa dalam *cluster* mahasiswa dengan nilai kurang memuaskan. Dan dengan menggunakan fismat 1 bertipe sugeno telah diketahui nilai *Training Error* (trnRMSE) data dan *Checking Error* (chkRMSE) data dengan tujuh kali *training* dan *testing* (uji data) direpresentasikan pada Tabel 3.

Tabel 3 Hasil trnRMSE dan chkRMSE

Train Ke-	Train Data	Test Data	TrnRMSE	ChkRMS E
1.	16	4	2.76×10^{-15}	0,1165
2.	16	4	6.90×10^{-15}	0,1396
3.	17	3	3.72×10^{-15}	0,3897
4.	15	5	7.08×10^{-15}	0,3299
5.	17	3	7.04×10^{-15}	0,2653
6.	15	5	2.96×10^{-15}	1,0862
7.	16	4	2.00×10^{-15}	0,2083

Berdasarkan Tabel3, hasil *Training Error* (trnRMSE) data dan *Checking Error* (chkRMSE) data terkecil ada pada *training* data pertama dan *testing* data pertama dengan menunjukkan nilai trnRMSE sebesar 2.76×10^{-15} dan chkRMSE sebesar 0.1165 dengan grafik nilai error direpresentasikan sebagai berikut:



Gambar 3. Grafik Error Pengclusteran

Berdasarkan representasi Gambar 3. Hasil dari empat data yang di *testing* menunjukkan bahwa nilai error data *testing* pertama sebesar 0.0607, nilai error data *testing* kedua sebesar 0.0278, nilai error data *testing* ketiga sebesar 0.0421, dan nilai error data *testing* keempat sebesar 0.2192.

Kemudian dari *cluster-cluster Fuzzy C-Means* menghasilkan derajat keanggotaan sebagaimana ditunjukkan Tabel 4.

Tabel 4

Derajat Keanggotaan Mahasiswa Matematika			
MHS	Derajat Keanggotaan		
	KM	M	SM
1	0.0058	0.9880	0.0063
2	0.0173	0.0532	0.9296
3	0.0604	0.3701	0.5695
4	0.0604	0.0346	0.9556
5	0.0070	0.0243	0.9687
6	0.9691	0.0239	0.0070
7	0.0415	0.8987	0.0598
8	0.0438	0.2087	0.7475
9	0.0256	0.1047	0.8697
10	0.0345	0.1950	0.7706
11	0.0620	0.9109	0.0271
12	0.0154	0.0694	0.9151
13	0.9808	0.0156	0.0035
14	0.0240	0.0703	0.9056
15	0.2094	0.7151	0.0754
16	0.1302	0.7060	0.1638

Tabel 4 menunjukkan bahwa mahasiswa 6 dan 13 masuk dalam mahasiswa dengan nilai sangat memuaskan, mahasiswa 1, 7, 11, 15, dan 16 masuk dalam mahasiswa dengan nilai memuaskan, mahasiswa 2, 3, 4, 5, 8, 9, 10, 12, dan 14 masuk dalam mahasiswa dengan nilai kurang memuaskan. Sehingga sistem monitoring sangat di perlukan untuk mahasiswa yang masuk dalam *clus* mahasiswa dengan nilai kurang memuaskan agar dapat memperbaiki nilai dan untuk mengetahui kendala-kendala apa saja yang menghambat perkembangan mahasiswa tersebut. Tetapi, mahasiswa yang lain pun juga memerlukan adanya sistem monitoring. Untuk mengetahui dan mengevaluasi hasil perkembangan akademisnya.

5. KESIMPULAN

Pengelompokan mahasiswa merupakan hal yang penting karena bertujuan untuk memonitor perkembangan akademik mahasiswa. Salah satu metode yang cocok untuk pemetaan atau pengelompokan mahasiswa berdasarkan kesamaan karakteristik dan kemampuan adalah metode *FuzzyC-Means Clustering*. Dengan parameter-parameter yang digunakan berupa nilai-nilai indeks prestasi semester 1 sampai 3 dan nilai indeks prestasi kumulatif dari data mahasiswa program studi matematika, yang parameter-parameter ini dikelompokkan ke dalam 3 *cluster* yaitu *cluster* mahasiswa dengan nilai sangat memuaskan, memuaskan dan kurang memuaskan.

Berdasarkan analisa *FuzzyC-Means Clustering* dari 16 mahasiswa program studi matematika diperoleh mahasiswa 6 dan 13 masuk dalam mahasiswa dengan nilai sangat memuaskan, mahasiswa 1, 7, 11, 15, dan 16 masuk dalam mahasiswa dengan nilai memuaskan, mahasiswa 2, 3, 4, 5, 8, 9, 10, 12, dan 14 masuk dalam mahasiswa

dengan nilai kurang memuaskan. Sehingga sistem monitoring sangat di perlukan untuk mahasiswa yang nilainya kurang memuaskan, tidak hanya mahasiswa dengan nilai kurang memuaskan, mahasiswa yang lain pun juga memerlukan sistem monitoring ini. Dari hasil uji coba dengan menggunakan sistem *FuzzyC-Means Clustering* ke dalam 3 *cluster* dengan 16 data yang di *training* didapatkan akurasi RMSE 2.76×10^{-15} dan 4 data yang di *testing* didapatkan akurasi RMSE 0.1165.

DAFTAR PUSTAKA

- [1] Setianingrum, A. Hanifa. Model Pemetaan Evaluasi Penilaian Kualifikasi Lulusan Berbasis Metode Fuzzy *C-Means Clustering*. Jurnal Teknik Informatika Vol. 7 No. 2. 2014.
- [2] Kusumadewi, Sri dan Hari Purnomo. 2013. “Aplikasi Logika Fuzzy Edisi 2”. Yogyakarta: Graha Ilmu.
- [3] Kusumadewi, Sri. 2002. “Analisis & Desain Sistem Fuzzy menggunakan Tool Box Matlab”. Yogyakarta: Graha Ilmu.
- [4] Inyang, Udoinyang G. dan Enobog E. Joshua. *Fuzzy Clustering of Students' Data Repository for At-Risk Students Identification and Monitoring*. Jurnal Computer and Information Science Vol. 6 No. 4. 2013.
- [5] Khoiruddin, Arwan Ahmad. Menentukan Nilai Akhir Kuliah Dengan Fuzzy C-Means. Seminar Nasional dan Informatika SNS107-041. 2007.
- [6] Martiana, Entin, dkk. Penggunaan Metode Pengclasteran untuk Menentukan Bidang Tugas Akhir Mahasiswa Teknik Informatika PENS Berdasarkan Nilai. Politeknik Elektronika Negeri Surabaya.
- [7] Novitasari, Dian Candra Rini. Klasifikasi Sinyal EEG Menggunakan Metode *Fuzzy C-Means Clustering* (FCM) dan *Adaptive Neighborhood Modified Backpropagation* (ANMBP). Jurnal Mantik Vol. 1 No. 1 2015.
- [8] Novitasari, Dian Candra Rini. Klasifikasi Sinyal EEG Menggunakan Metode *Fuzzy C-Means Clustering* (FCM) dan *Adaptive Neuro Fuzzy Inference System* (ANFIS). Thesis Teknik Informatika ITS. 2013.
- [9] Novitasari, Dian Candra Rini. Pengklasteran Lahan Sawah di Indonesia sebagai Evaluasi Ketersediaan Produksi Pangan Menggunakan *Fuzzy C-Means*. Jurnal Mantik Vol. 2 No. 1 2016.
- [10] Burhanuddin, Afid. *Monitoring dan Evaluasi Satuan Pendidikan*. Di akses pada tanggal 30 Juni 2016 jam 07.24 di <https://afidburhanuddin.wordpress.com/2014/01/20/%C2%AD%C2%AD%C2%ADmonitoring-dan-evaluasi-pengelolaan-satuan-pendidikan/>

