

# **CLUSTERING KINERJA AKADEMIS MAHASISWA PASCASARJANA UIN KHAS JEMBER MENGGUNAKAN FUZZY C- MEANS**

**Elok Rahmawati<sup>1</sup>, Deni Arifianto<sup>2</sup>, Amalina Maryam Zakiyyah<sup>3</sup>**

<sup>1</sup>Mahasiswa Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember  
email: [elokrahmawati894@gmail.com](mailto:elokrahmawati894@gmail.com)

<sup>2</sup>Dosen Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember  
email: [deniarifianto@unmuhjember.ac.id](mailto:deniarifianto@unmuhjember.ac.id)

<sup>3</sup>Dosen Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember  
email: [amalinamaryamzakiyyah@unmuhjember.ac.id](mailto:amalinamaryamzakiyyah@unmuhjember.ac.id)

## **ABSTRAK**

*Clustering mahasiswa program pascasarjana dilakukan untuk mengelompokkan mahasiswa ke dalam cluster tertentu yang memiliki kemiripan data menggunakan algoritma Fuzzy C-Means. Data yang diambil yaitu data umur dan ips (index prestasi mahasiswa). Metode validitas cluster dari hasil algoritma Fuzzy C-Means menggunakan DBI (Davies Bouldin Index) untuk mengukur seberapa bagus cluster yang dihasilkan oleh algoritma Fuzzy C-Means. Dari penelitian yang telah dibuat peneliti berikut hasil DBI cluster 2 adalah cluster yang menghasilkan nilai DBI terkecil yaitu 1,9570. Dan berdasarkan hasil dari identifikasi profil mahasiswa pada 2 cluster didapatkan bahwa cluster 1 memiliki data mahasiswa dengan rentang umur 21 sampai 45, nilai ips (index prestasi semester) 1 rentang 2,00 sampai 4,00 dan rata-rata jenis kelamin adalah perempuan. Sedangkan, cluster 2 memiliki data mahasiswa dengan rentang umur 24 sampai 45, nilai ips (index prestasi semester) 1 rentang 1,70 sampai 3,80 dan rata-rata jenis kelamin adalah laki-laki.*

**Kata kunci:** DBI (Davies Bouldin Index), Fuzzy C-Means

## **ABSTRACT**

*Clustering graduate students is done to group students into certain clusters that have similar data using the Fuzzy C-Means algorithm. The data taken are age data and ips (student achievement index). The cluster validity method of the Fuzzy C-Means algorithm results uses DBI (Davies Bouldin Index) to measure how good the cluster produced by the Fuzzy C-Means algorithm is. From the research that has been made by the following researchers, the DBI results of cluster 2 are the clusters that produce the smallest DBI value, namely 1.9570. And based on the results of identifying student profiles in 2 clusters, it is found that cluster 1 has student data with an age range of 21 to 45, the value of ips (semester achievement index) 1 ranges from 2.00 to 4.00 and the average gender is female. Meanwhile, cluster 2 has student data with an age range of 24 to 45, an ips value (semester achievement index) 1 range of 1.70 to 3.80 and the average gender is male.*

**Keyword :** DBI (Davies Bouldin Index), Fuzzy C-Means

## **1. PENDAHULUAN**

Salah satu indikator kemajuan suatu negara adalah meningkatnya kualitas pendidikan yang dimilikinya. Dalam hal ini, bidang pendidikan memiliki peran penting untuk meningkatkan kualitas diri. Tergantung pada kualitas tenaga kerja akan dipengaruhi oleh beberapa hal diantaranya pendidikan dan kesehatan. Dengan mengenyam pendidikan tinggi, manusia dapat meningkatkan kualitasnya. Salah satu perguruan tinggi dapat memperoleh pelatihan tambahan selain dari pendidikan yang diwajibkan pemerintah, yaitu dengan mengikuti kuliah.

Di Indonesia banyak sekali tempat menuntut ilmu dan jurusan yang berbeda-beda, salah satunya adalah di kota Jember yang merupakan kota yang memiliki tempat belajar yang cukup banyak dan berbagai jurusan yang bisa dipelajari untuk menimba ilmu. .bangun.

Untuk mencapai efek pembelajaran yang berkualitas, UIN KHAS Jember telah menerapkan berbagai cara untuk pengembangan yang telah dilakukan seperti dengan merumuskan kurikulum, membangun gedung dengan fasilitas yang sesuai dan melengkapi dan melengkapi fasilitas khusus untuk siswa penyandang cacat, mengevaluasi kinerja fakultas dan mengubah pengajaran metode.

Rasio mahasiswa pascasarjana merupakan salah satu ukuran kualitas kampus UIN KHAS Jember karena rasio mahasiswa pascasarjana dapat mempengaruhi nilai akreditasi kampus. Sebagai contoh, studi kasus pada program Pascasarjana yang masa studi normalnya adalah 2 tahun studi atau 4 semester menunjukkan bahwa masih banyak mahasiswa yang tidak dapat mengikuti perkuliahan dalam batas studi normal. Di bawah ini adalah persentase total data untuk mahasiswa pascasarjana lembaga agama Islam Jember 2015-2018.

Berdasarkan banyaknya mahasiswa program Pascasarjana UIN KHAS Jember terdapat berbagai macam karakter serta kondisi mahasiswa yang berbeda-beda. Sering kali mengakibatkan mahasiswa tersebut mengulang mata kuliah yang terjadi karena mendapat nilai yang tidak sesuai dengan ketentuan ke lulusan. Peneliti di sini berkeinginan untuk mengelompokkan data mahasiswa melalui proses metode *clustering*.

Dalam proses pengelompokan data mahasiswa peneliti menggunakan metode clustering dikarenakan metode tersebut mampu memperoleh output kelompok-kelompok data. Menurut Satriyanto (2011), *Clustering* adalah metode pengelompokan berdasarkan ukuran kedekatan. Metode *clustering* memiliki bermacam-macam algoritma pada proses pengolahannya. *Fuzzy C-means* merupakan salah satu dari macam-macam algoritma metode *clustering*. 1 yang disusun oleh Fitria dkk, dengan judul Perbandingan Pengclusteran Data Iris Menggunakan Metode *K-Means* dan *Fuzzy C-Means*. Menyimpulkan bahwa hasil dari pengclusteran data user *knowledge* modeling dengan menggunakan metode *K-Means* dan *Fuzzy C-Means* ternyata metode *Fuzzy C-Means* yang menghasilkan nilai validitasnya yang mendekati 1 yaitu 0,2854 sedangkan menggunakan metode *K-Means* menghasilkan nilai 0,1866 yang artinya metode *Fuzzy C-Means* lebih baik dari pada metode *K-Means* untuk melakukan *clustering*.

Oleh karenanya peneliti ingin menggunakan metode *Fuzzy C-Means* untuk penelitian peneliti, dikarenakan metode *Fuzzy C-means* menghasilkan *output* yang lebih baik untuk proses pengklasteran data. Maka peneliti mengajukan judul penelitian "*Clustering* Kinerja Akademis Mahasiswa Pascasarjana UIN KHAS Jember Menggunakan *Fuzzy C-Means*".

## 2. TINJAUAN PUSTAKA

Penelitian ini tidak terlepas dari literatur yang digunakan sebagai landasan, dasar serta pedoman dalam menyelesaikan penelitian ini.

### A. Data Mining

Data *mining* adalah proses pengumpulan data yang mengandung sejumlah besar informasi yang akan digunakan untuk tujuan tertentu.

## B. Clustering

*Clustering* adalah metode dalam pengolahan data mining yang menghasilkan output dari proses pengolahan data nya adalah sebuah kelompok atau kelas-kelas yang berkarakteristik sama atau mirip.

## C. Fuzzy C-Means

Output dari *Fuzzy C-Means* adalah deretan pusat *cluster*, dan setiap titik data memiliki beberapa derajat keanggotaan. Informasi ini dapat digunakan untuk membangun sistem *inferensi fuzzy*. Berikut Persamaan 1 - 5 perhitungan *fuzzy c-means*.

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

$$\mu_{ik} = \frac{Q_i}{\sum_{i=1}^n ((\mu_{ik})^w * x_{ij})} \quad (2)$$

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

$$P_t = \sum_{i=1}^n \sum_{k=1}^c ([\sum_{j=1}^m (X_{ij} - V_{kj})])^2 (\mu_{ik})^w \quad (4)$$

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

## D. Normalisasi Min-Max

*Normalisasi Min-Max* adalah proses yang digunakan untuk menormalkan semua variabel data menjadi data dengan bobot yang sama (Shofiani, 2017: 51). Dapat dilihat pada Persamaan 6

$$X_n = \frac{x_0 - x_{min}}{x_{max} - x_{min}} \quad (6)$$

## E. Davies Bouldin Index

Tujuan dari DBI adalah menentukan nilai atau value dari beberapa cluster yang diperoleh untuk menjadikan salah satu nilai cluster yang paling valid untuk di ambil atau digunakan. Berikut Persamaan 7

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} (R_{i,j}) \quad (7)$$

Untuk menghitung nilai DBI nilai yang harus di cari terlebih dahulu adalah:

1. *SSW (Sum of Square Within-cluster)*. Dengan menggunakan Persamaan 8.

$$SSW_i = \frac{1}{m_i} \sum_{j=i}^{m_i} d(x_j, c_i) \quad (8)$$

2. *SSB (Sum of Square Between-cluster).*

Dengan menggunakan Persamaan 9.

$$SSB_{i,j} = d(ci, cj) \quad (9)$$

3. Ratio. Dengan menggunakan Persamaan 10

$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_{i,j}} \quad (10)$$

#### 4. METODE PENELITIAN

##### A. Pendekatan Penelitian

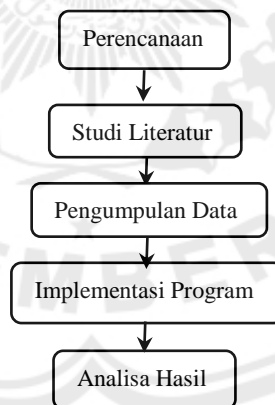
Metode penelitian kuantitatif mengacu pada observasi yang mengukur skala nilai sampai batas tertentu. Penelitian kuantitatif mencakup semua jenis penelitian berdasarkan persentase, rata-rata, atau kalkulasi statistik lainnya. Alat yang digunakan misalnya tabel dan diagram.

##### B. Jenis Penelitian

Riset terapan menekankan pada manfaat praktis dari hasil riset dalam memecahkan masalah spesifik dan menemukan produk baru yang bermanfaat bagi kehidupan.

##### C. Langkah-langkah Penelitian

Langkah-langkah penelitian peneliti ditunjukkan pada Gambar 1, serta penjelasannya sebagai berikut:



Gambar 1. Metode Penelitian  
Sumber: Langkah Penelitian

##### D. Perencanaan

Pada tahap perencanaan terdapat 3 uraian yang akan dilakukan untuk tahap ini yaitu merumuskan masalah, survey, observasi.

## E. Studi Literatur

Adapun buku dan jurnal yang diambil referensi dalam studi literatur ini adalah Buku ciptaan Kusuma, S., & Purnomo, H. (2010). Aplikasi Logika Fuzzy.

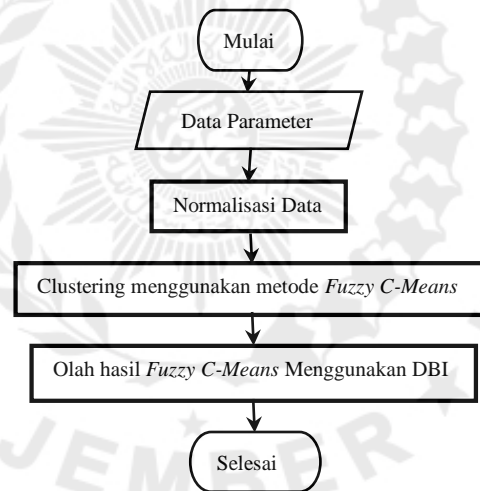
## F. Analisis Data

Pada proses penumpulan data terdapat 4 hal yang diperhatikan yaitu :

1. Seleksi data Bertujuan untuk menyeleksi data parameter yang diperlukan dalam proses data *mining*.
2. *Cleaning* Bertujuan untuk menjamin kualitas data yang dipilih, pada titik ini masalah yang akan diselesaikan adalah duplikasi data.
3. Integrasi data bertujuan mengintegrasikan data ke dalam database, karena datanya lebih dari satu tabel, maka perlu dilakukan tahapan penggabungan data dari berbagai tabel yang berbeda.
4. Normalisasi parameter yang bertujuan untuk menyeimbangkan data nilai parameter dan menyederhanakan data nilai parameter dengan skala value 0 sampai 1 menggunakan Persamaan 6

## G. Implementasi Program

Pada gambar 2 diagram alur dibawah ini akan dijelaskan bagaimana langkah perancangan sistem.



Gambar 2. Diagram Alur Sistem  
Sumber: Implemtasi Program

## H. Clustering data menggunakan metode *Fuzzy C-Means*

Menentukan nilai parameter awal berdasarkan data dari tabel 1 ditentukan beberapa komponen perhitungan sebagai berikut:

- a. banyaknya *cluster* yang diinginkan =  $c$
- b. pangkat (pembobot) =  $w$
- c. maksimum iterasi =  $T$
- d. error terkecil =  $10^{-5}$
- e. fungsi objektif awal =  $P_0$
- f. iterasi awal = 1

Untuk menentukan pusat cluster menggunakan Persamaan 3.

Fungsi objektif pada iterasi pertama P1 dihitung dengan menggunakan Persamaan 4.

Menghitung perubahan Matriks Partisi dengan menggunakan Persamaan 5.

#### I. Validitas Cluster dengan Metode *Davies Bouldin Index* (DBI)

Pada Tabel 1. adalah hasil nilai anggota cluster dan Tabel 2. Adalah hasil pusat cluster

Tabel 1. Hasil nilai anggota cluster

Data cenderung masuk ke cluster ke-	
1	2
0,9932	0,9189
0,9932	0,8950
	0,7914

Sumber: Hasil DBI

Tabel 2. Hasil Pusat Cluster

cluster	Pusat cluster		
1	0,0745	0,7622	0,9218
2	0,0662	0,6139	0,0080

Sumber: Hasil Pusat Cluster

Setelah memperoleh nilai anggota cluster dan nilai pusat cluster maka dapat dihitung SSW, SSB, dan Ratio. Tabel 3. Hasil nilai DBI

Tabel 3. Nilai Rmax dan DBI

R	Data cluster ke-i		Rmax	DBI
	1	2		
Data cluster ke-	1	0	2,5558	2,5558
	2	2,5558	0	

Sumber: Hasil Perhitungan DBI

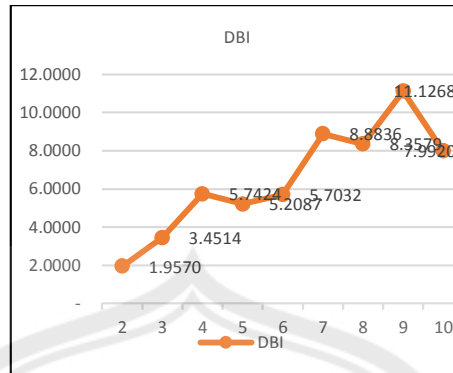
#### J. Analisa Hasil

Analisa hasil adalah tahapan terakhir dari penelitian ini. Dengan di ujinya data yang didapatkan kemudian data tersebut diolah dengan algoritma *fuzzy c-means* dan DBI (*Davies Boildin Index*).

Dalam analisa hasil ini berikut lingkungan uji coba dan jadwal proyek penelitian:

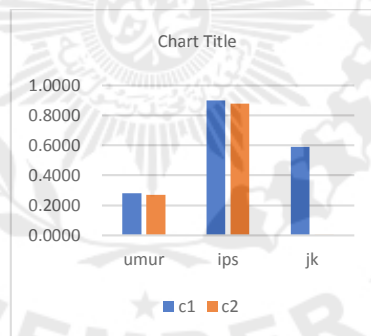
1. Lingkungan Uji Coba Pascasarjana UIN KHAS Jember
2. Alat dan bahan yang digunakan pada penelitian ini adalah :
  - a. Laptop core i5
  - b. Aplikasi text editor vs code
  - c. Xampp
  - d. Data parameter

## 5. PEMBAHASAN DAN HASIL



Gambar 3. Grafik Nilai DBI  
Sumber: Hasil Penelitian

Berdasarkan pada Gambar 1. Adalah Grafik nilai DBI menunjukkan bahwa pada di titik 1,9570 adalah nilai DBI terkecil dari 10 titik nilai DBI, dimana titik 1,9570 adalah 2 cluster. Maka 2 cluster adalah cluster yang paling optimal dari pengelompokan cluster yang lain. Dimana, evaluasi hasil DBI yang bernilai paling kecil menunjukkan bahwa nilai DBI yang paling baik karena semua anggota cluster relatif dekat satu dengan yang lain. Dan hasil pengelompokan pada Gambar 4. adalah data profil mahasiswa berdasarkan parameter umur, ips dan jenis kelamin dengan cluster 1 sebanyak 457 mahasiswa dan cluster 2 sebanyak 270 mahasiswa.



Gambar 4. grafik rata-rata kedekatan  
Sumber: Hasil Penelitian

## 6. PENUTUP

### A. Kesimpulan

Berdasarkan penelitian yang dilakukan oleh peneliti dalam pengelompokan seleksi kinerja akademis mahasiswa program pascasarjana UIN KHAS Jember menggunakan algoritma *fuzzy c-means*. Diperoleh hasil perhitungan DBI dari hasil masing-masing DBI maka dapat disimpulkan bahwa cluster 2 merupakan cluster paling baik karena nilai DBI cluster 2 menunjukkan nilai yang paling mendekati 0.

Dan mengidentifikasi profil di masing-masing cluster. Pada Gambar 4 adalah hasil identifikasi dari profil 2 cluster. Berdasarkan grafik identifikasi 2 cluster didapatkan bahwa cluster 1 memiliki data mahasiswa dengan umur 21 sampai 45, nilai ips 2,00 sampai 4,00

dan rata-rata jenis kelamin adalah perempuan. Sedangkan pada *cluster* 2 memiliki data mahasiswa dengan umur 24 sampai 45, nilai ips 1,70 sampai 3,80 dan rata-rata jenis kelamin adalah laki-laki.

## B. Saran

Adapun saran agar aplikasi ini dapat berfungsi dengan lebih optimal yaitu adalah:

1. Metode yang digunakan untuk perhitungan pengelompokannya dapat menggunakan metode selain fuzzy c-means.
2. Menentukan nilai validitas pada cluster dapat menggunakan perhitungan selain perhitungan davies bouldin index (DBI).
3. Menyeimbangkan bobot data parameter agar data parameter memiliki bobot yang sama dapat menggunakan perhitungan selain perhitungan normalisasi min-max.

## 7. REFERENSI

1. Kusumadewi, S., dan Hari Purnomo. 2010. Aplikasi Logika Fuzzy. Yogyakarta: Graha Ilmu.
2. Fenny, D. Analisis perbandingan cosine normalization dan min-max normalization pada pengelompokan terjemahan ayat Al Quran menggunakan algoritma k-means clustering (Bachelor's thesis, Fakultas Sains dan Teknologi Universitas Islam Negeri Syarif Hidayatullah Jakarta).
3. Guney, Y. (2009). Exogenous and endogenous factors influencing students' performance in undergraduate accounting modules. *Accounting Education*, 18(1), 51- 73
4. Haryati, D. F., Abdillah, G., & Hadiana, A. I. (2016, March). Klasifikasi Jenis Batubara Menggunakan Jaringan Syaraf Tiruan Dengan Algoritma Backpropagation. In Seminar Nasional Teknologi Informasi dan Komunikasi (SENTIKA) (pp. 18-19)
5. Jaedun, A. (2011). Metodologi penelitian eksperimen. Fakultas Teknik UNY, 12.
6. Mardi, Y. (2017). Data Mining: Klasifikasi Menggunakan Algoritma C4. 5. *Edik Informatika*, 2(2), 213-219.
7. Metisen, B. M., & Sari, H. L. (2015). Analisis clustering menggunakan metode K-Means dalam pengelompokan penjualan produk pada Swalayan Fadhila. *Jurnal media infotama*, 11(2).
8. Muhammad, A. F. (2018). Clusterisasi Proses Seleksi Pemain Menggunakan Algoritma K-Means (Study Kasus: Tim Hockey Kabupaten Kendal). *Jurnal Sistem Komputer*, 7(2).
9. Nugroho, S. B. M. (2016). Pengaruh Pendidikan Terhadap Pertumbuhan Ekonomi. *Media Ekonomi dan Manajemen*, 29(2).
10. Rianto, A. (2017). Implementasi Metode K-Means Clustering Dalam Mengelompokkan Emosi Senang, Marah, Dan Netral Berdasarkan Vokal Manusia (Doctoral dissertation, Universitas Multimedia Nusantara).
11. Ridwan, M., Suyono, H., & Sarosa, M. (2013). Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier. *jurnal EECCIS*, 7(1), 59-64.



12. Rohmah, R. L. (2019). Zonasi daerah terdampak bencana angin Puting Beliung menggunakan K-Means Clustering dengan analisis Silhouette Coefficient, Davies Bouldin Index dan Purity (Doctoral dissertation, UIN Sunan Ampel Surabaya).
13. Santosa, P. (2015). Metode Penelitian Sastra. Yogyakarta: azzagrafika.
14. Satriyanto, E. (2011). Clustering. Surabaya: Institut Teknologi Surabaya.
15. Shofiani, N. (2017). Segmentasi Supplier Menggunakan Metode K-Means Clustering (Studi Kasus: PTPN X PG Meritjan) (Doctoral dissertation, Institut Teknologi Sepuluh Nopember).

