

IMPLEMENTASI DAN KOMPARASI ALGORITMA FUZZY C-MEANS DAN K-MEANS UNTUK MENGELOMPOKKAN SISWA BERDASARKAN NILAI AKADEMIK DAN PERILAKU SISWA (DATA SURVEY)

Hadian Artanto¹⁾, Istiadi²⁾, Fitri Marisa³⁾, Dwi Purnomo⁴⁾

¹⁾ Program Studi Teknik Informatika, Universitas Widyagama Malang
Email : hadian496@gmail.com

²⁾ Program Studi Teknik Informatika, Universitas Widyagama Malang
Email : istiadi@widyagama.ac.id

³⁾ Program Studi Teknik Informatika, Universitas Widyagama Malang
Email : fitrimarisa@widyagama.ac.id

ABSTRAK

Di dunia pendidikan seringkali nilai akademik selalu dijadikan acuan apakah sudah berhasil atau tidak dalam proses menyelenggarakan pembelajaran dan bimbingan. Akan tetapi prestasi yang diperoleh oleh siswa dipengaruhi oleh banyak faktor seperti tingkat kedisiplinan. Oleh karena itu penelitian ini dilakukan untuk mengelompokkan siswa tidak hanya berdasarkan nilai akademik saja tetapi juga perilaku siswa sehari-hari dengan menggunakan algoritma *K-Means* dan *Fuzzy C-Means* (FCM). Penelitian ini menghasilkan rata-rata tingkat akurasi dari proses perhitungan algoritma *K-Means* mencapai 91% dan membutuhkan waktu 4.4105 detik. Sedangkan rata-rata tingkat akurasi dari proses perhitungan algoritma *Fuzzy C-Means* mencapai 68% dan membutuhkan waktu 5.5416 detik.

Kata Kunci : *Clustering, Algoritma K-Means, Algoritma Fuzzy C-Means (FCM)*

ABSTRACT

In the world of education, academic value is often always used as a reference whether it has been successful or not in the process of organizing learning and guidance. However, the achievements obtained by students are influenced by many factors such as the level of discipline. Therefore this study was conducted to classify students not only based on academic grades but also daily student behavior using the K-Means and Fuzzy C-Means (FCM) algorithms. This research resulted in an average accuracy rate of the K-Means algorithm calculation process reaching 91% and takes 4.4105 seconds. While the average accuracy rate of the Fuzzy C-Means algorithm calculation process reaches 68% and takes 5.5416 seconds.

Keywords : *Clustering, Algoritma K-Means, Algoritma Fuzzy C-Means (FCM)*

PENDAHULUAN

Pendidikan adalah proses perubahan sikap dan tingkah laku seseorang melalui proses pembelajaran dan bimbingan. Keberhasilan pendidikan tidak hanya dipandang dari aspek akademik saja, tetapi juga dilihat dari aspek spiritual dan sosial. Dengan demikian, pendidikan memegang peranan penting dalam membentuk sifat dan tabiat peserta didik yang bermutu dan berdaya guna agar sesuai dengan cita-cita pendidikan (Alimaun, 2015). Kedisiplinan merupakan salah satu faktor yang mempengaruhi hasil belajar siswa. Penyebab ketidakterdisiplinan siswa adalah kurang jelasnya peraturan dan sanksi yang diberikan kepada siswa. Untuk itu perlu adanya pemberian sanksi dan pengawasan yang lebih jelas maupun pembinaan berupa konseling (Baskara, 2016).

Sehingga disetiap sekolah perlu adanya pengelompokan berdasarkan tingkat kedisiplinan dan prestasi dari masing-masing siswa, agar pihak sekolah terutama guru BK lebih mudah mengetahui kelompok siswa yang perlu diberi penanganan khusus.

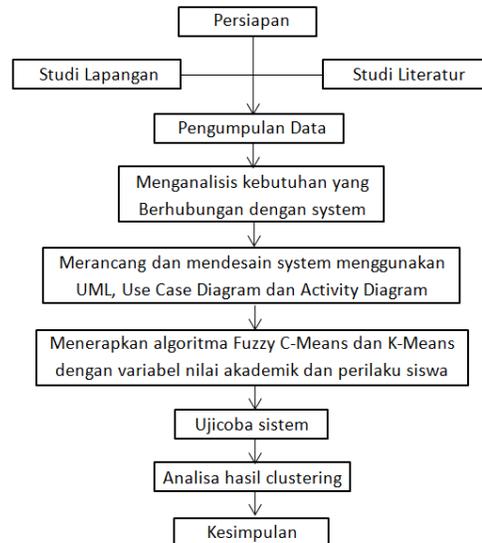
Berbagai penelitian telah dilakukan untuk mengelompokkan siswa. Nelson (Butarbutar, 2016) menggunakan algoritma *Fuzzy C-Means* dan *K-Means* dalam mengelompokkan data siswa berdasarkan nilai akademik. Maya (Bancin, 2014) menggunakan algoritma *K-Means* dan AHP dalam penilaian kedisiplinan siswa. Fitria (Febrianti, 2016) menggunakan algoritma *Fuzzy C-Means* dan *K-Means* dalam mengelompokkan data iris. Nisia (Yuanita, 2016) menggunakan algoritma *K-Means* dalam pembagian kelas siswa.

Pada saat ini sekolah sudah berusaha untuk meningkatkan kualitas siswanya baik itu nilai akademik maupun perilaku siswa. Penelitian ini bertujuan untuk mengelompokkan data siswa berdasarkan nilai akademik (Matematika, Bahasa Inggris, Bahasa Indonesia) dan perilaku siswa (Disiplin Waktu, Disiplin Tata Tertib, Disiplin Mengerjakan Tugas dan Disiplin Berpakaian). Penelitian ini menggunakan metode *clustering* dengan algoritma *Fuzzy C-Means* dan *K-Means*. Setelah hasil diperoleh, selanjutnya akan di analisis.

Clustering merupakan salah satu metode *data mining* yang digunakan untuk mengelompokkan data ke dalam suatu kelompok dengan karakteristik yang mirip.

METODE

Penelitian ini menggunakan pendekatan *Waterfall* dengan konsep dasar SDLC (*System Development Life Cycle*) seperti yang ditampilkan pada Gambar 1.



Gambar 1 Metode Penelitian

Terdapat beberapa tahapan antara lain: perencanaan, analisis dan implementasi. Penelitian ini dilakukan di SMK NU Sunan Ampel Poncokusumo yang mencakup 134 siswa. Dimana data penilaiannya berupa nilai akademik (Matematika, Bahasa Indonesia dan Bahasa Inggris) dan perilaku siswa sehari-hari (Disiplin Waktu, Disiplin Tata Tertib, Disiplin Mengerjakan Tugas dan Disiplin Berpakaian). Untuk data nilai akademik diperoleh dari operator sekolah sedangkan data perilaku siswa diperoleh dari kuesioner dan catatan guru BK.

Dalam metode *clustering* konsep utama yang ditekankan adalah pencarian pusat *cluster* secara iteratif, dimana pusat *cluster* ditentukan berdasarkan jarak minimum setiap data pada pusat *cluster*.

Algoritma K-Means

Langkah- langkah penyelesaian algoritma *K-Means* dapat diuraikan sebagai berikut:

1. Menentukan jumlah *cluster* yang akan dibentuk.
2. Menentukan nilai *centroid* atau rata-rata data yang ada di masing-masing *cluster*.
3. Menghitung jarak setiap data yang ada terhadap masing-masing *centroid* menggunakan rumus *Euclidian* hingga ditemukan jarak yang paling dekat dari setiap data dengan *centroid*.

$$D_{ik} = \sqrt{\sum_{j=1}^p (X_{ij} - X_{kj})^2} \tag{1}$$

Dimana :

D(i,k) = Perhitungan jarak data ke-i dengan pusat klaster ke-k

p = maksimal data

j = banyak data

4. Kelompokkan data berdasarkan jarak minimum data terhadap pusat *cluster*.
5. Kembali ke tahap 2, lakukan perulangan hingga nilai *centroid* yang dihasilkan tetap dan anggota *cluster* tidak berpindah ke *cluster* lain.

Algoritma Fuzzy C-Means

Langkah- langkah penyelesaian algoritma *Fuzzy C-Means* dapat diuraikan sebagai berikut :

1. Tentukan jumlah *cluster* dengan komponen seperti Tabel 1.

Tabel 1 Komponen jumlah *cluster*

Komponen Jumlah Cluster	Keterangan
Jumlah Cluster	3
Pangkat untuk matriks partisi	2
Maximum iterasi	100
Error terkecil yang diharapkan	0.000001
Fungsi objektif awal	P ₀ = 0

2. Bangkitkan Bilangan Random
3. Hitung Pusat *Cluster*

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w * X_{ij})}{\sum_{i=1}^n (\mu_{ik})^w} \tag{2}$$

4. Hitung Fungsi Objektif

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^2 \right) \tag{3}$$

5. Hitung Perubahan Matriks

$$\mu_{ik} = \frac{\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{-1}{(w-1)}}}{\sum_{k=1}^c \left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{-1}{(w-1)}}} \tag{4}$$

6. Cek kondisi berhenti, apabila error terkecil sudah mencapai ketentuan maka iterasi berhenti jika tidak, ulangi langkah ke-4.

HASIL DAN PEMBAHASAN

Dari penelitian ini diperoleh beberapa hasil yang diperoleh, diantaranya: Halaman algoritma *K-Means* ditunjukkan seperti pada Gambar 2.

Analisa Menggunakan Data Mining Metode K-Means
 Data Alternatif, Kriteria dan Nilai
 Yang Akan Dihitung Keanggotaannya dalam Cluster

Jumlah Cluster Dicari	3
Maksimum Iterasi	100
<input type="button" value="Proses"/>	

Gambar 2 Tampilan Halaman Algoritma *K-Means*

Berdasarkan Gambar 2, admin melakukan analisa data yang berasal dari walikelas. Untuk melakukan analisa, admin harus menentukan jumlah *cluster* dan maximum iterasi, selanjutnya tekan tombol proses.

Halaman Algoritma *Fuzzy C-Means*

Halaman algoritma *Fuzzy C-Means* ditunjukkan seperti pada Gambar 3.

Analisa Menggunakan Data Mining Metode Fuzzy C-Means
 Data Alternatif, Kriteria dan Nilai
 Yang Akan Dihitung Keanggotaannya dalam Cluster

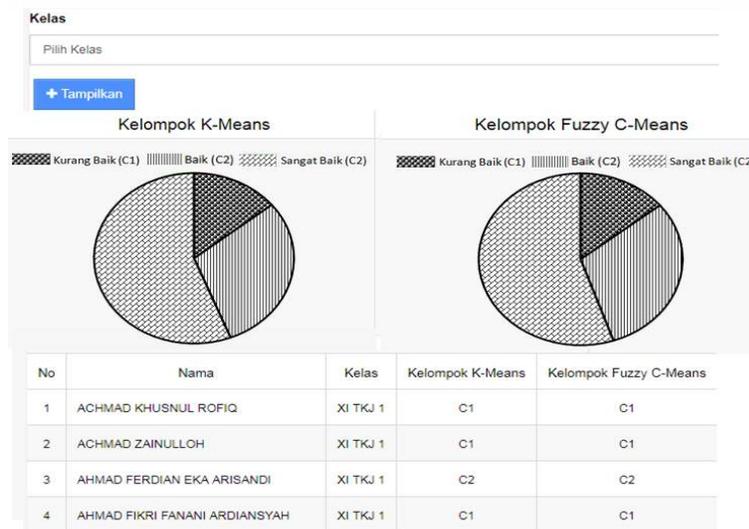
Jumlah Cluster Dicari	3
Maksimum Iterasi	100
Nilai Pembobot (Pangkat)	2
Nilai Error Terkecil	0.000001
<input type="button" value="Proses"/>	

Gambar 3 Tampilan Halaman Algoritma *Fuzzy C-Means*

Berdasarkan Gambar 3, admin melakukan analisa data yang berasal dari walikelas. Untuk melakukan analisa, admin harus menentukan jumlah *cluster*, maximum iterasi, nilai pangkat dan nilai error terkecil, selanjutnya tekan tombol proses.

Halaman Perbandingan Hasil *Clustering*

Halaman perbandingan hasil *clustering* ditunjukkan seperti pada Gambar 4.



Gambar 4 Halaman Perbandingan Hasil *Clustering*

Berdasarkan Gambar 4, perbandingan hasil *clustering* ditampilkan sesuai kelas masing-masing, di mana disebelah kiri untuk hasil pengelompokkan *K-Means* dan disebelah kanan untuk hasil pengelompokkan *Fuzzy C-Means*. Setelah mendapat hasil *clustering*, hasil tersebut dianalisa untuk mengetahui akurasi dari algoritma *K-Means* dan *Fuzzy C-Means*, seperti ditampilkan pada Tabel 2.

Tabel 2 Perbandingan Hasil *Clustering*

		K-Means	Fuzzy C-Means
		Jumlah Data	134 siswa
		Jumlah Cluster	3 kelompok
Percobaan ke-1		Jumlah Iterasi	5 iterasi
		Jumlah Error	12 siswa
		Akurasi	91%
		Waktu	4.8332 detik
			32 iterasi
Percobaan ke-2		Jumlah Error	63 siswa
		Akurasi	53%
		Waktu	5.1462 detik
			6.1533 detik
Percobaan ke-3		Jumlah Iterasi	29 iterasi
		Jumlah Error	11 siswa
		Akurasi	92%
		Waktu	3.2521 detik
Rata-Rata		Waktu	4.5472 detik
		Jumlah Iterasi	31 iterasi
		Jumlah Error	45 siswa
		Akurasi	68 %
	Waktu	4.4105 detik	5.5416 detik

Dari Tabel 2, dapat diketahui bahwa telah dilakukan 3 kali percobaan terhadap sistem untuk mengelompokkan 134 siswa yang dikelompokkan menjadi 3 *cluster*. Rata-rata tingkat akurasi dari proses perhitungan algoritma *K-Means* mencapai 91% dan membutuhkan waktu 4.4105 detik. Sedangkan rata-rata tingkat akurasi dari proses perhitungan algoritma *Fuzzy C-Means* mencapai 68% dan membutuhkan waktu 5.5416 detik.

KESIMPULAN

Penelitian ini menghasilkan rata-rata tingkat akurasi dari proses perhitungan algoritma *K-Means* mencapai 91% dan membutuhkan waktu 4.4105 detik. Sedangkan rata-rata tingkat akurasi dari proses perhitungan algoritma *Fuzzy C-Means* mencapai 68% dan membutuhkan waktu 5.5416 detik. Sehingga dapat disimpulkan bahwa implementasi dari algoritma *K-Means* lebih baik daripada implementasi *Fuzzy C-Means*.

REFERENSI

- Alimaun, I. (2015). Pengaruh Kedisiplinan Terhadap Hasil Belajar Siswa Kelas V Sekolah Dasar se-Daerah Binaan R.A. Kartini Kecamatan Kutoarjo Kabupaten Purworejo.
- Bancin, M. S. (2014). Implementasi Metode K-Means Clustering dan Metode Analytical Hierarchy Process (AHP) Dalam Penilaian Kedisiplinan Siswa.
- Baskara, D. A. (2016). Implementasi Metode K-Means Clustering Dalam Penilaian Kedisiplinan Siswa (Studi Kasus : Smk Sunan Kalijogo Lumajang).
- Butarbutar, N. (2016). Komparasi Kinerja Algoritma Fuzzy C-Means dan K-Means Dalam Pengelompokkan Data Siswa Berdasarkan Prestasi Nilai Akademik Siswa. 1.

- Febrianti, F. (2016). Perbandingan Pengklusteran Data Iris Menggunakan Metode K-Means Dan Fuzzy C-Means. *02*.
- Yuanita, N. (2016). Implementasi K-Means Clustering untuk Pembagian Kelas Siswa.