

Penerapan Fuzzy C-Means Pada Curah Hujan di Kalimantan Timur

Ahmad Wahbi Fadillah^{*1}, Andi Tejawati², Novianti Puspitasari³

^{1,2,3}Jurusan Teknologi Informasi dan Komunikasi, Universitas Mulawarman, Samarinda
e-mail: ahmadwahby.cisco2@gmail.com, anditejawati@yahoo.com, miechan.novianti@gmail.com

Abstrak

Saat ini teknologi informasi berkembang sangat cepat. Begitu juga dengan perkembangan ilmu pengetahuan. Namun seiring dengan perkembangan ilmu pengetahuan dan teknologi, sistem informasi memiliki fungsi yang beragam, baik sebagai alat bantu perhitungan matematis, simulasi, pemodelan grafis, juga sebagai alat bantu untuk melakukan prediksi berdasarkan data yang termasuk di dalamnya yaitu melakukan pengelompokan curah hujan. Pengelompokan curah hujan mengacu pada rangkaian aktivitas yang dilakukan untuk menghasilkan sekumpulan informasi mengenai kondisi iklim dan unsur - unsurnya. Informasi ini berguna dalam banyak hal salah satunya adalah untuk prakiraan curah hujan yang memiliki pengaruh yang besar terhadap segala macam aktifitas kehidupan. Contohnya, para petani sangat membutuhkan informasi mengenai prakiraan musim hujan agar mereka dapat mempersiapkan masa tanam dengan lebih baik. Bidang lain yang membutuhkan informasi mengenai curah hujan antara lain pariwisata, perikanan, pelayaran, perkebunan, kehutanan, pembangunan gedung, penataan wilayah, dan kesehatan. Dalam hal ini masyarakat membutuhkan sistem informasi untuk mengetahui kategori curah hujan di Kalimantan Timur. Maka dari itu dibuatlah suatu sistem informasi untuk mengelompokkan curah hujan di Kalimantan Timur menggunakan metode Fuzzy C-Means. Dari penelitian ini menghasilkan suatu sistem informasi guna membantu masyarakat untuk mengetahui tingkat curah hujan di daerahnya yang ada di Kalimantan Timur. Hasil dari sistem ini akan memberikan informasi tingkat curah hujan dalam tiga kategori yaitu kategori tertinggi, sedang, dan rendah untuk daerah Kalimantan Timur.

Kata kunci : Fuzzy, C-Means, Curah Hujan.

1. PENDAHULUAN

Saat ini teknologi informasi berkembang sangat cepat. Begitu juga dengan perkembangan ilmu pengetahuan. Namun seiring dengan perkembangan ilmu pengetahuan dan teknologi, sistem informasi memiliki fungsi yang beragam, baik sebagai alat bantu perhitungan matematis, simulasi, pemodelan grafis, juga sebagai alat bantu untuk melakukan prediksi berdasarkan data yang termasuk di dalamnya yaitu melakukan pengelompokan curah hujan.

Pengelompokan curah hujan mengacu pada rangkaian aktivitas yang dilakukan untuk menghasilkan sekumpulan informasi mengenai kondisi iklim dan unsur-unsurnya. Informasi ini berguna dalam banyak hal salah satunya adalah untuk prakiraan curah hujan yang memiliki pengaruh yang besar terhadap segala macam aktifitas kehidupan. Contohnya, para petani sangat membutuhkan informasi mengenai prakiraan musim hujan agar mereka dapat mempersiapkan masa tanam dengan lebih baik. Bidang lain yang membutuhkan informasi mengenai curah hujan antara lain pariwisata, perikanan, pelayaran, perkebunan, kehutanan, pembangunan gedung, penataan wilayah, dan kesehatan.

Informasi yang berkaitan dengan kondisi dimasa yang akan datang tidak dapat ditentukan dengan pasti tetapi hanya bisa diprediksi atau diramalkan. Informasi tentang curah hujan, apapun bentuknya asalkan dapat dipercaya dan valid, sungguh sangat diperlukan oleh masyarakat. Salah satunya adalah penentuan curah hujan suatu wilayah, namun karena sulitnya untuk menentukan perulangan masa atau periode curah hujan secara manual, sehingga diperlukan sebuah perangkat lunak yang dapat mengelompokkan data curah hujan menurut

kemiripan datanya. Maka dari itu dibutuhkan suatu sistem informasi yang dapat digunakan untuk menentukan curah hujan.

Berdasarkan uraian diatas maka penulis mengambil tema “Penerapan Metode Fuzzy C-Means Pada Curah Hujan Untuk Perwilayahan Provinsi Kalimantan Timur”.

2. METODE PENELITIAN

Logika *fuzzy* merupakan suatu teori himpunan logika yang dikembangkan untuk mengatasi konsep nilai yang terdapat diantara kebenaran (*true*) dan kesalahan (*false*). Logika *fuzzy* meniru cara berpikir manusia dengan menggunakan konsep kesamaan suatu nilai. Dalam teori logika *fuzzy* suatu nilai bisa bernilai benar atau salah secara bersama. Namun berapa besar kebenaran dan kesalahan suatu tergantung pada bobot keanggotaan yang dimilikinya.

2.1 Fuzzy Clustering Means (FCM)

Fuzzy Clustering adalah salah satu teknik untuk menentukan cluster optimal dalam suatu ruang vector yang didasarkan pada bentuk normal euclidian untuk jarak antara vektor. Fuzzy Clustering sangat berguna bagi pemodelan *fuzzy* terutama dalam mengidentifikasi aturan – aturan *fuzzy*

2.1.1 Algoritma Fuzzy Clustering Means (FCM)

Algoritma *Fuzzy C-Means* adalah sebagai berikut.

1. Input data yang akan dikelompokkan, yaitu X , berupa matrix berukuran $n \times m$ (n =jumlah sampel data, m = atribut setiap data). X_{ij} data sampel ke- i ($i=1,2,\dots,n$), atribut ke- j ($j=1,2,\dots,m$).
2. Tentukan :
 - a) Jumlah *cluster* = (c).
 - b) Pangkat untuk matriks partisi = (w).
 - c) Maksimum iterasi = (MaxIter).
 - d) *Error* terkecil yang diharapkan = (ξ).
 - e) Fungsi objektif awal = (Po=0).
 - f) Iterasi awal = (t=1).
3. Bangkitkan bilangan random μ_{ik} , $i=1,2,\dots,n$; $k=1,2,\dots,c$ sebagai elemen matrik partisi awal U .

$$\text{Menghitung tiap baris : } Q_i = \sum_{k=1}^c \mu_{ik} \quad (1)$$

Dengan $j = 1,2,\dots,n$

$$\text{Menghitung : } \mu_{ik} = \frac{\mu_{ik}}{Q_i} \quad (2)$$

4. Hitung pusat *cluster* ke- k : V_{kj} , dengan $k=1,2,\dots,c$; dan $j=1,2,\dots,n$;

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

5. Hitung fungsi objektif pada iterasi ke- t , P_t ;

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

6. Hitung perubahan matriks partisi :

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

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

7. Cek Kondisi berhenti :

- a) Jika : $(|P_t - P_{t-1}| < \xi)$ atau $t > \text{MaxIter}$ maka berhenti;

b) Jika tidak : $t = t + 1$, ulangi langkah ke-4.

Keterangan :

μ = Himpunan

Q = Jumlah setiap kolom

V = Pusat *cluster*

P = Fungsi objektif

2.2.2 Sum of Square Error (SSE)

Metode ini berusaha untuk meminimalkan variasi antar data yang ada di dalam suatu *cluster* dan memaksimalkan variasi dengan data yang ada di *cluster* lainnya. Digunakan fungsi kriteria: jumlah dari kesalahan kuadrat (*sum of squared-error*, SSE) yang dapat mengukur kualitas klastering yang dibuat dengan persamaan (6).

$$SSE = \sum_{i=1}^k \sum_{p \in C_i} d(p, m_i)^2 \quad (6)$$

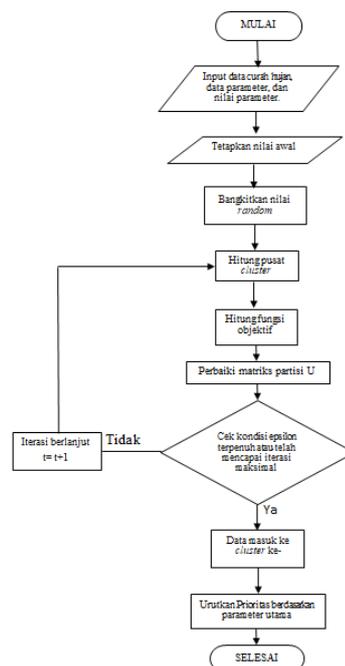
$p \in C_i$ = tiap data poin pada cluster i ,

m_i = centroid dari cluster i ,

d = jarak/ distances/ variance terdekat pada masing-masing cluster i .

Mengoptimalkan nilai fungsi kriteria tersebut, nilai SSE tergantung pada jumlah klaster dan bagaimana data dikelompokkan ke dalam klaster-klaster. Semakin kecil nilai SSE semakin bagus hasil klastering yang dibuat.

Dalam penelitian ini, metode yang digunakan adalah *Fuzzy Clustering Means* (FCM). Metode *Fuzzy C-Means* mengelompokkan objek sehingga objek yang paling dekat kesamaannya dengan objek lain berada dalam cluster yang sama. Berikut adalah *Flowchart* Algoritma *Fuzzy C-Means* yang dapat dilihat pada Gambar 1.



Gambar 1. *Flowchart* Algoritma *FuzzyC-Means*

3. HASIL DAN PEMBAHASAN

Penelitian ini menguji berupa data curah hujan 1 tahun (12 Bulan) pada tahun 2016, dan berasal dari 10 stasiun pengamatan BMKG Propinsi Kalimantan Timur dengan parameter yang sudah ditentukan.

Tabel 1. Data Curah Hujan

Lokasi	Jumlah Curah Hujan (mm)		
	Tertinggi	Terendah	Rata – Rata
Long Iram	1.287	98	682
Kota Bangun	461	60	184
MA. Kaman	413	71	207
Tenggarong Seberang	313	60	166
Tenggarong	274	69	159
MA. Ancalong	448	33	212
Temindung	367	98	209
Samboja	438	97	190
Balikpapan	538	45	189
Berau	528	123	252

Untuk memperoleh hasil akhir fungsi objektif, lakukan penjumlahan kolom dari hasil penjumlahan kolom dari hasil perhitungan diatas. Sehingga hasil dari fungsi objektif hasil *clustering* dapat dilihat pada tabel 2.

Tabel 2. Hasil Clustering

No	Lokasi	Derajat keanggotan data pada <i>cluster</i> ke-			Data cenderung masuk ke <i>cluster</i> ke-		
		1	2	3	1	2	3
1	Long Iram	1,0000	0,0000	0,0000	*		
2	Kota Bangun	0,0006	0,0282	0,9712			*
3	Ma. Kaman	0,0026	0,2412	0,7562			*
4	Tenggarong Seb.	0,0002	0,9921	0,0077		*	
5	Tenggarong	0,0015	0,9523	0,0463		*	
6	Ma. Ancalong	0,0020	0,0888	0,9092			*
7	Temindung	0,0030	0,7166	0,2804		*	
8	Samboja	0,0018	0,1108	0,8874			*
9	Balikpapan	0,0060	0,0962	0,8978			*
10	Berau	0,0093	0,1319	0,8587			*

Dari perhitungan iterasi yang dilakukan maka diperoleh hasil klasifikasi data curah hujan tinggi, sedang, dan rendah. Hasil Pengelompokan *C-Means* dapat dilihat pada tabel 3.

Tabel 3. Hasil Pengelompokan

No	Lokasi	Kelompok	Curah Hujan
1	Long Iram	C1	Tertinggi
2	Kota Bangun	C3	Terendah
3	Ma. Kaman	C3	Terendah
4	Tenggarong Seb.	C2	Sedang
5	Tenggarong	C2	Sedang
6	Ma. Ancalong	C3	Terendah
7	Temindung	C2	Sedang
8	Samboja	C3	Terendah
9	Balikpapan	C3	Terendah
10	Berau	C3	Terendah

Pengujian Hasil perhitungan dilakukan untuk melihat seberapa besar perbedaan antara hasil pengelompokan curah hujan yang ditentukan dengan sistem yang merepakan metode *Fuzzy K-Means* dengan *Fuzzy C-Means*. Hasil perbandingan dapat dilihat pada tabel 4.

Table 4. Hasil Perbandingan

Lokasi	Hasil Pengelompokan KET (T/F)		
	K-MEANS	C-MEANS	
Long Iram	TERTINGGI	TERTINGGI	T
Kota Bangun	SEDANG	SEDANG	T
Ma. Kaman	SEDANG	SEDANG	T
Tenggarong Seb.	TERENDAH	TERENDAH	T
Tenggarong	TERENDAH	TERENDAH	T
Ma. Ancalong	SEDANG	SEDANG	T
Temindung	SEDANG	TERENDAH	F
Samboja	SEDANG	SEDANG	T
Balikpapan	SEDANG	SEDANG	T
Berau	SEDANG	SEDANG	T

Ket : T = True; F = False

Terlihat bahwa terdapat 1 data lokasi yang memiliki nilai *false*. Dimana, nilai *false* ini menunjukkan perbedaan antara hasil keputusan menggunakan metode *K-Means* dan hasil keputusan menggunakan metode *C-Means*. Selanjutnya, untuk mengetahui seberapa besar persentase perbedaan hasil pengelompokan data, dilakukan perhitungan dengan menggunakan persamaan (7).

$$\text{Hasil} = \frac{9}{10} \times 100\% = 90\% \quad (7)$$

Berdasarkan hasil perhitungan diatas, diperoleh persentase sebesar 90%. Hal ini menunjukkan bahwa tidak ada perbedaan yang signifikan antara hasil keputusan menggunakan metode *Fuzzy K-Means* dengan *C-Means*.

Tahap terakhir adalah pengujian menggunakan metode *Sum of Squared Error* (SSE) yang memberikan informasi *error* jarak data ke *centroid*. Dalam penelitian ini, SSE digunakan untuk menguji hasil *clustering* yang dibuat sebanyak tiga *cluster* dan hasil *clustering* sebanyak dua *cluster*. Hasil pengujian menggunakan metode SSE dapat dilihat pada tabel 5.

Table 5. Hasil Pengujian Metode Sum of Squared Error

Cluster	Hasil SSE	Selisih K2 & K3
K=2	9,84664932	1,769402731
K=3	8,077246589	

Dari tabel diatas, diperoleh informasi bahwa nilai SSE tiga cluster menunjukkan nilai yang lebih kecil dibandingkan dengan nilai SSE dua cluster. Hal ini menunjukkan bahwa hasil klastering dengan jumlah tiga cluster merupakan jumlah cluster yang ideal untuk mendapatkan informasi pola curah hujan yang terjadi.

Implementasi sistem merupakan penciptaan perangkat lunak untuk perancangan sistem. Implementasi bertujuan untuk mengkonfirmasi modul – modul perancangan, sehingga pengguna dapat memberikan masukan kepada pembangunan system. Untuk implementasi sistem dari Tampilan halaman informasi *user* dapat dilihat pada Gambar 2.



Gambar 2. Halaman Utama Sistem

Halaman ini merupakan halaman depan untuk memilih menu menu yang disediakan sistem terdiri dari data master, *fuzzyC-Means* dan hasil klasifikasi. Untuk hasil pengelompokan curah hujan dapat dilihat pada Gambar 3.

HASIL PENGELOMPOKAN CURAH HUJAN		
No	WILAYAH	HASIL PENGELOMPOKAN
1	Long Iram	Tinggi
2	Kota Bangun	Sedang
3	Ma. Kaman	Sedang
4	Berau	Sedang
5	Ma. Ancalong	Sedang
6	Samboja	Sedang
7	Balikpapan	Sedang
8	Temindung	Rendah
9	Tenggarong Seberang	Rendah
10	Tenggarong	Rendah

Gambar 3. Pengelompokan Curah Hujan

4. KESIMPULAN

Berdasarkan hasil penelitian dan pengujian maka kesimpulan yang di dapat dari penelitian ini adalah metode *FuzzyC-Means* telah diterapkan untuk mengelompokan curah hujan ke dalam tiga kategori (tinggi , rendah, dan sedang) di Provinsi Kalimantan Timur. Hasil penelitian

menunjukkan bahwa metode *C-Meansclustering* dengan tiga *cluster* merupakan jumlah yang ideal dibandingkan dengan hasil *clustering* berjumlah dua *cluster*. Penelitian ini menghasilkan informasi tentang kategori tingkat curah hujan di Kalimantan Timur, dimana daerah Long Iram memiliki tingkat curah hujan yang tinggi dan Tenggarong, Tenggarong Seberang, dan Temindung memiliki kategori curah hujan terendah. Perbandingan penelitian ini dengan penelitian sebelumnya yaitu pada hasil perhitungan pengelompokan curah hujan pada stasiun Temindung. Di dalam penelitian ini curah hujan pada stasiun Temindung menghasilkan pengelompokan kategori rendah, sedangkan pada penelitian sebelumnya menggunakan metode K-Means yang menghasilkan pengelompokan kategori sedang.

5. SARAN

Berdasarkan pembahasan dan kesimpulan dalam penelitian ini, sistem ini dapat dikembangkan dengan menambahkan obyek penelitian berupa menghitung nilai kelembapan suhu minimum, suhu maksimum, kualitas udara dan kecepatan angin. Agar nantinya sistem ini akan memberikan informasi yang lebih lengkap tepat dan akurat kepada masyarakat luas.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada dosen pembimbing ibu Andi Tejawati dan ibu Novianti Puspitasari yang telah memberi bimbingan dan masukan dalam perbaikan penulisan skripsi ini. Dan juga selaku dosen penguji ibu Masna Wati dan bapak Hario Jati Setyadi yang telah memberikan saran dan masukan dalam perbaikan penulisan skripsi.

DAFTAR PUSTAKA

- [1] Adelia, dan Jimmy Setiawan. 2011. Implementasi Customer Relationship Management (CRM) pada Sistem Reservasi Hotel berbasis Website dan Desktop. Bandung: Universitas Kristen Maranatha. Vol. 6, No. 2.
- [2] Aditya Nur Santoso, and Yehuda. 2012 . "Penerapan Algoritma FuzzyC-Means untuk." Program Studi Sistem Informasi FTI-UKSW.
- [3] Agustina, Fajar. 2015. "Penerapan *FuzzyC-Means* Pada Pemetaan Jurusan Sekolah." *Seminar Nasional Ilmu Pengetahuan dan Teknologi* vol. 4, no. 1, pp. 38-43.
- [4] Anhar. 2010. PHP & MySql Secara Otodidak. Jakarta: PT Trans Media Havaluddin. 2011. "Memahami Penggunaan UML (Unified Modelling." *Jurnal Informatika Mulawarman* vol. 6, no. 1, pp 1-2.
- [5] Kusumadewi, Sri, and Hadi Purnomo. 2010 . *Aplikasi Logika Fuzzy untuk Pendukung*. Yogyakarta : Graha Ilmu.
- [6] Luthfi, Emha Taufiq. 2017 . "*Fuzzyc-Means* Untuk *Clustering* Data (Studi Kasus : Data Performance Mengajar Dosen)". *Seminar Nasional Teknologi*.
- [7] Masrur, Mukhamad. 2016. Pemrograman Web Dinamis Menggunakan Java Server Pages dengan Database Relasional MYSQL. Yogyakarta: Andi.
- [8] Pressman, R S. 2010,. *Software Engineering : a practitioner's approach*. New York: McGraw-Hill.

- [9] Puspitasari, Novianti, and Havaluddin. 2016. "Penerapan Metode *K-Means* Dalam Pengelompokan Curah Hujan di Kalimantan Timur." *Jurnal Ilmu Komputer & Teknologi Informasi* 1-6.
- [10] Puspitasari, Novianti, Rosmasari, and Stefanie. 2017. "Penentuan Prioritas Perbaikan Jalan Menggunakan *FuzzyC-Means*: Studi Kasus." *Jurnal Teknologi dan Sistem Komputer* vol. 5, no. 1 7-14.
-