

PENGLASTERAN LAHAN SAWAH DI INDONESIA SEBAGAI EVALUASI KETERSEDIAAN PRODUKSI PANGAN MENGUNAKAN *FUZZY C-MEANS*

Nur Affiah¹, Dian C. Rini², Ahmad Lubab³

Matematika, Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Ampel Surabaya

E-mail: afifahaziec22@gmail.com¹, diancrini@uinsby.ac.id², ahmadlubab@uinsby.ac.id³

Abstrak

Luas lahan sawah di Indonesia semakin sempit dengan maraknya pembangunan perumahan dan gedung-gedung. Hal ini berakibat pada ketersediaan produksi pangan yang semakin rendah dan harus mengimpor beras dari negara lain. Dengan mengklasterkan lahan sawah dapat digunakan sebagai evaluasi untuk meningkatkan produksi pangan di Indonesia sehingga kegiatan impor beras dapat terminimalisasi. Metode yang digunakan untuk mengelompokkan lahan sawah adalah metode *Fuzzy C-Means*. Implementasi program pada matlab dengan data *training* dan data *testing*. Pada program *Fuzzy C-Means* tersebut menghasilkan tiga kelompok/*cluster* data, yaitu luas lahan sawah luas, sedang, dan sempit. Hasil pengklusteran, wilayah yang paling berpotensi dalam produksi pangan dari lahan sawah adalah Jawa Timur, Jawa Tengah dan Jawa Barat.

Kata kunci: Lahan sawah, Evaluasi, Fuzzy C-Means

Abstract

The number of rice field in Indonesia is decreasing due to development of residential areas and buildings. Consequently, it reduces foodstuff availability and government should import it from other. Increasing food production and minimizing imported food can be started by clustering fields as an evaluation. This clustering is approached by *Fuzzy C-Means*. Training and Testing data are implemented on Matlab and yield three categories, wide, medium and narrow field. Moreover, the most potential field is East Java, Central Java, and West Java.

Keywords: Field, Evaluation, Fuzzy C-Means

1. Pendahuluan

Dahulu pada masa orde baru kepemimpinan Soeharto, Indonesia mendapat sebutan sebagai lumbung padi, mengubah statusnya dan mencapai swasembada beras pada tahun 1980-an [1]. Selain itu, Indonesia menyandang sebagai Negara produsen beras terbesar ke tiga di dunia [2] dalam kegiatan ekspor beras. Hal ini dikarenakan luas lahan sawah di Indonesia masih sangat luas untuk produksi tanaman padi. Namun semakin bertambahnya tahun jumlah produksi beras menurun. Salah satu penyebabnya adalah perkembangan luas lahan sawah dari tahun

ketahun yang semakin menurun. Data pada Badan Pusat Statistik (2001), luas lahan sawah Indonesia pada tahun 1993 \pm 8.500.000 ha, selanjutnya pada tahun 2000 (7 tahun) telah menyusut serius hingga tinggal 7.790.000 ha atau susutnya lahan 710.000 ha atau setiap tahunnya tanah sawah di Indonesia menyusut 59,167 ha[3]. Penurunan ini disebabkan oleh pengalihan fungsi dari lahan sawah untuk pembangunan pemukiman dan industri. Sehingga penurunan tersebut berdampak pada berkurangnya ketersediaan produksi pangan. Dan untuk menanggulangi ketersediaan tersebut, pemerintah mengencar impor beras

untuk memenuhi kebutuhan pangan dalam negeri.

Pengklasteran lahan sawah pada setiap provinsi dapat digunakan sebagai bahan evaluasi dalam meningkatkan ketersediaan produksi pangan pada suatu daerah dengan mengetahui luas lahan sawah yang ada di Indonesia. Dengan diimbangi teknologi, hal tersebut dapat meminimalisasi kegiatan impor beras dalam pemenuhan kebutuhan pangan dalam negeri.

Fuzzy C-Means adalah metode yang banyak digunakan untuk menyelesaikan permasalahan yang berhubungan dengan klaster/klasifikasi. Diantaranya adalah digunakan untuk mengklasterkan varietas padi[4], diagnose penyakit jantung[5], analisa kepercayaan data berdasarkan metode klasifikasi[6]. Sehingga untuk menyelesaikan permasalahan klaster lahan sawah pada penelitian ini dapat digunakan metode *Fuzzy C-Means*. Metode *Fuzzy C-Means* adalah suatu teknik pengklusteran data yang mana keberadaan tiap-tiap titik data dalam suatu *cluster* ditentukan oleh derajat keanggotaan[7]. sehingga dapat ditentukan *cluster* lahan sawah.

2. Tinjauan Pustaka

Teori himpunan fuzzy akan memberikan jawaban terhadap suatu permasalahan yang mengandung ketidakpastian[8]. Salah satu bab dalam fuzzy adalah *fuzzy clustering*, fuzzy ini digunakan untuk mengklusterkan data. Dalam *clustering* ini dibagi menjadi empat metode, yaitu metode *Fuzzy Subtractive Clustering*, *Mountain*, *K-means* dan *C-Means*.

Fuzzy C-Means (FCM) adalah suatu teknik pengklasteran data yang mana keberadaan tiap-tiap data dalam suatu *cluster* ditentukan oleh nilai keanggotaan[9]. Pada kondisi awal, pusat *cluster* masih belum akurat sehingga dibutuhkan perbaikan pusat *cluster* secara berulang hingga berada pada titik yang tepat. Setiap data akan memiliki derajat keanggotaan untuk setiap *clusternya*.

Algoritma dari *Fuzzy C-Means* adalah sebagai berikut[10] :

1. Input data yang akan di *cluster* x, berupa matriks berukuran n x m (n= jumlah sampel data, m= atribut setiap data). X_{ij} = data ke-i(i=1,2,...,n) , atribut ke-j(j=1,2,...,m)
2. Tentukan :
 - Jumlah *cluster* (c)
 - Pangkat (w)
 - Maksimum iterasi
 - Error terkecil yang diharapkan (ξ)
 - Fungsi objektif awal ($P_0=0$)
 - Iterasi awal (t=1)
3. Bangkitkan bilangan random μ_{ik} sebagai elemen matriks partisi awal U. Hitung umlah setiap kolom :

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

Kemudian hitung:

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

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

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

5. Hitung fungsi objektif pada iterasi ke-t, P_t (yan, 1994):

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

6. Hitung perubahan matriks partisi :

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

7. Cek kondisi berhenti :
 - Jika $t > \max$ iter maka berhenti
 - Jika tidak, $t=t+1$, ulangi langkah ke-4

3. Metode Penelitian

3.1. Studi Literatur

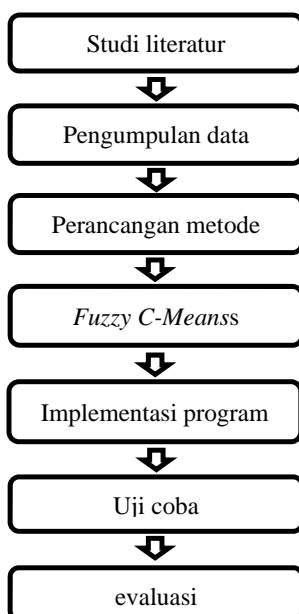
Lahan yang diambil sebagai obyek penelitian adalah lahan sawah dari seluruh provinsi di Indonesia. Data luas lahan sawah ini didapatkan dari Badan Pusat Statistik (BPS). Data yang digunakan untuk penelitian ini adalah data luas lahan sawah dari tahun 2004 sampai 2013.

Provinsi	Luas Lahan Sawah (Hektar)										
	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	
ACEH	346305	356649	315277	312803	323010	359751	313649	307556	308973	300808	
SUMATERA UTARA	502839	462767	460486	453372	478521	464256	468724	467138	448722	438346	
SUMATERA BARAT	231939	228176	229469	227355	225623	228176	229693	231463	230775	224182	
RIAU	125966	118955	124985	128242	122255	122738	115961	115897	109585	93338	
JAMBI	122126	117482	119242	117543	116212	117336	112434	113757	113379	113546	
SUMATERA SELATAN	474429	484207	523922	530204	577821	611072	611386	629355	610314	612424	
BENGKULU	85641	84164	83885	93779	89244	89614	92976	90217	88877	93382	
LAMPUNG	316017	313621	317413	342507	348732	349144	345437	350949	364111	360237	
KEP. BANGKA BELITUNG	3773	4111	4048	4176	3506	5017	4056	5932	6133	5358	
KEP. RIAU	-	76	82	124	133	238	442	393	559	487	
DKI JAKARTA	2563	1866	1466	1200	1200	1215	1312	1098	1001	895	
JAWA BARAT	932337	925900	926782	934845	945544	937373	930268	930507	923575	925042	
JAWA TENGAH	996197	964102	963401	962942	963984	960768	962471	960970	962289	952525	
DI YOGYAKARTA	56982	57188	56218	55540	55332	55325	55291	55023	55336		

Gambar 1 Sampel Data

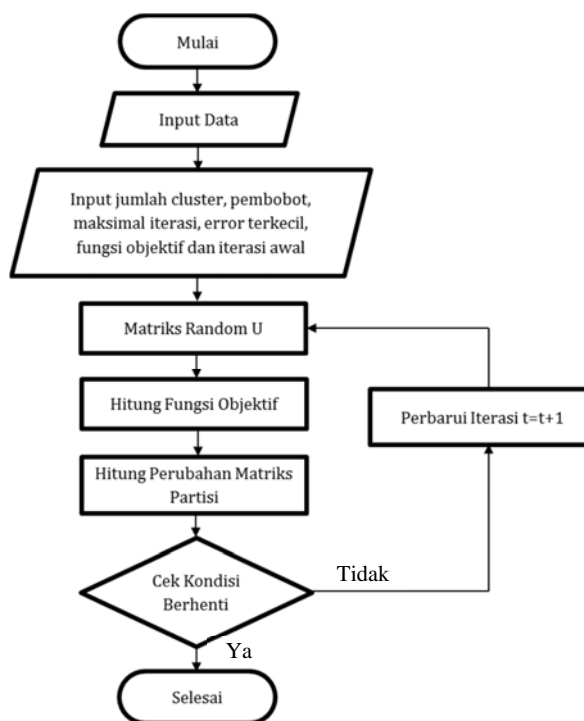
3.2. Pengolahan data

Dari data lahan sawah tiap provinsi yang diperoleh sebanyak 34 data akan diklusterkan menjadi 3 tingkat luasan, yaitu tinggi, sedang, dan rendah. Proses pengolahan data menjadi beberapa tahap. Berikut adalah tahapan alur pengolahan data seperti pada Gambar 2 berikut.



Gambar 2. Alur olah data

Pengklasteran ini dilakukan menggunakan algoritma *Fuzzy C-Means* dan aplikasi matlab dengan fungsi *c-means*.



Gambar 3. Diagram alur proses clustering

Untuk perhitungan *cluster* menggunakan penentuan jumlah kluster 3, yaitu sempit, sedang, dan luas dengan pangkat berbobot 2 dan maksimal iterasi 100.

Setelah penentuan tersebut, dilakukan perhitungan matriks random dengan persamaan (1) yang dilanjutkan dengan menghitung fungsi objektif hingga algoritma selesai. Namun pada implementasi program, data dirubah dalam bentuk matriks dan disimpan dengan format “.dat”. kemudian untuk mencari *cluster* fungsi yang digunakan pada program adalah :

```

X=load('data.dat');
[Center,U,ObjFcn]=fcm(X,3)

```

Fungsi center untuk mencari pusat *cluster* pada matriks random U dan kemudian dicari fungsi objektif.

Dalam pengelolaan data dalam matlab ada 2 macam data yang digunakan yaitu data

input dan data output. Untuk mendapatkan data output dilakukan cluster terlebih dahulu sehingga didapatkan derajat keanggotaan setiap cluster-nya. Data hasil cluster tersebut akan menjadi data output. Dari kedua data tersebut digunakan untuk menghitung tingkat akurasi atau error dengan fungsi RMSE. Dari data input dan output juga diambil data sebagai training dan testing untuk pengujian. Sehingga dapat diketahui error keseluruhan.

Data hasil clustering dengan error terkecil akan digunakan sebagai evaluasi dalam meningkatkan ketersediaan produksi pangan di Indonesia.

4. Hasil dan Pembahasan

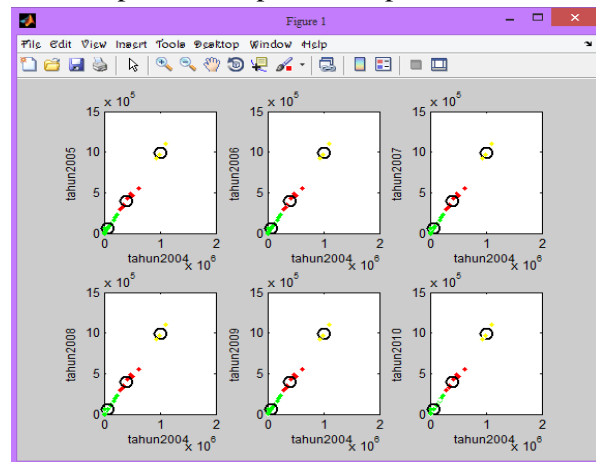
Implementasi program untuk penelitian ini terbagi menjadi dua, yaitu proses clustering dan proses pengujian. Dalam proses clustering menggunakan data awal dan dihasilkan tiga cluster dengan derajat keanggotaannya. Warna hijau menunjukkan cluster 1 yaitu sempit, warna merah menunjukkan cluster 2 yaitu sedang, dan warna kuning menunjukkan cluster 3 yaitu luas.

Proses clustering menghasilkan data output. Dari data tersebut di ambil data training dan data testing yang kemudian dari kedua data tersebut digunakan untuk menghitung error.

Nilai pada cluster di atas tidak akan sama jika dilakukan cluster ulang. Hal ini disebabkan oleh nilai matriks partisi U awal yang dibangkitkan secara random. Namun perubahan dari nilai-nilai tersebut tidak terlalu signifikan karena tidak mempengaruhi keanggotaan cluster.

Hasil cluster terlihat pada gambar 4 dengan perbedaan warna untuk setiap cluster-nya. Cluster yang berada paling atas merupakan pusat cluster terbesar yaitu lahan yang memiliki wilayah sawah yang luas. begitu pula untuk cluster yang berada di bawah merupakan wilayah dengan luas lahan sawah

sempit. Dari proses clustering dihasilkan data untuk tiap cluster dapat dilihat pada Tabel 1.

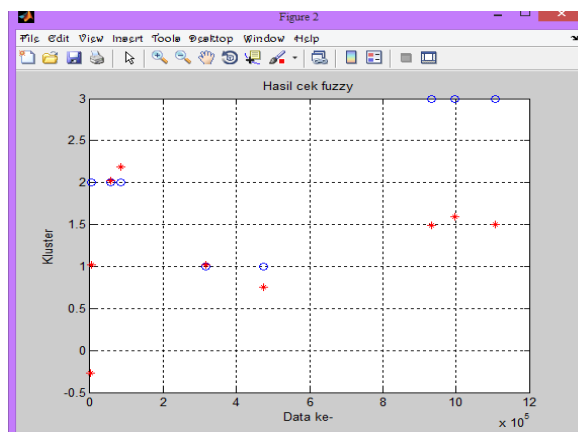


Gambar 4. Hasil clustering

Tabel 1. Data hasil clustering

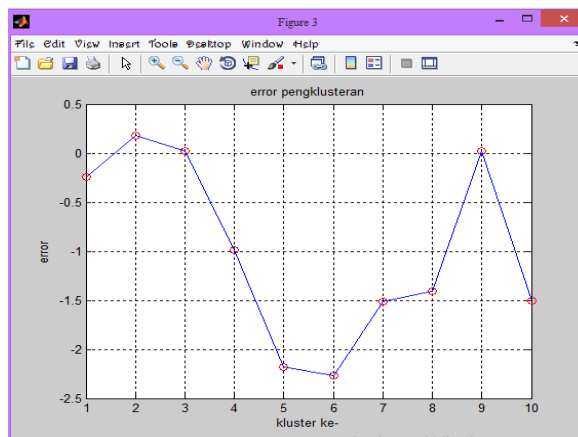
CLUSTER	DATA KE	PROVINSI
1	3	SUMATRA BARAT
	4	SUMATRA UTARA
	5	JAMBI
	7	BENGKULU
	9	KEP. BANGKA BELITUNG
	10	KEP. RIAU
	11	DKI JAKARTA
	14	DI YOGYAKARTA
	16	BANTEN
	17	BALI
	18	NUSA TENGGARA BARAT
	19	NUSA TENGGARA TIMUR
	21	KALIMANTAN TENGAH
	23	KALIMANTAN TIMUR
	24	KALIMANTAN UTARA
	25	SULAWESI UTARA
26	SULAWESI TENGAH	
28	SULAWESI TENGGARA	
29	GORONTALO	
30	SULAWESI BARAT	
31	MALUKU	
32	MALUKU UTARA	

	33	PAPUA BARAT
	34	PAPUA
2	1	ACEH
	2	SUMATERA UTARA
	6	SUMATERA SELATAN
	8	LAMPUNG
	20	KALIMANTAN BARAT
	22	KALIMANTAN SELATAN
	27	SULAWESI SELATAN
3	12	JAWA BARAT
	13	JAWA TENGAH
	15	JAWA TIMUR

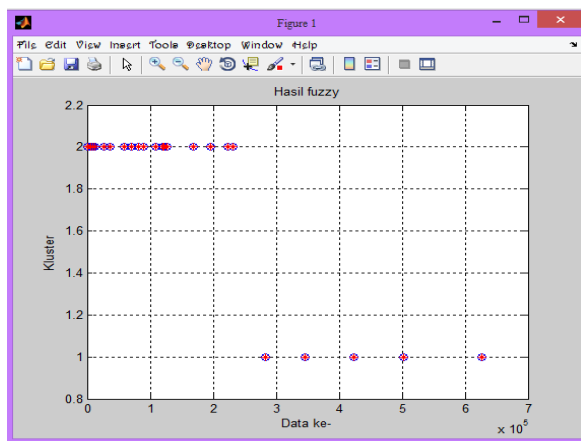


Gambar 6. Data testing

Proses pengujian data *training* dan data *testing* dilakukan dengan beberapa kali percobaan untuk menghasilkan nilai dengan tingkat error terkecil. Sehingga dari nilai ini bisa di terima untuk keakurasian data.



Gambar 7. Checking Error



Gambar 5. Data training

Dari data yang di *testing* dapat diketahui nilai error untuk setiap datanya. Dan untuk mengetahui tingkat keakuratan digunakan fungsi RMSE (*Root Mean Square Error*), didapatkan nilai error keseluruhan dari 5 percobaan seperti pada Tabel 2.

Tabel 2. Data Hasil Percobaan *Training* dan *Testing*

No	Data Training	Data Testing	Hasil Error
1	24	10	2.72×10^{-15}
2	23	11	1.40×10^{-14}
3	27	7	7.44×10^{-15}
4	28	6	2.60×10^{-14}
5	25	9	1.93×10^{-14}

Dari hasil percobaan diatas didapatkan nilai error terkecil yaitu pengujian yang dilakukan dengan 24 data *training* dan 10 data *testing* yang menghasilkan nilai error 0.00000000000000272.

5. KESIMPULAN

Metode *Fuzzy C-Means (FCM)* dapat diimplementasikan dalam pengklasteran lahan sawah di Indonesia. *Cluster* yang dihasilkan dari proses *clustering* ada 3, yaitu lahan sawah sempit, sedang, dan luas. Pengklusteran ini menggunakan 34 data. Dari proses *clustering* dihasilkan data output sehingga dapat dilakukan proses *training* dan *testing*. Dari 34 data *input* dan *output* dengan beberapa percobaan *training* dan *testing* didapatkan percobaan dengan nilai error terkecil yaitu 24 data *training* dan 10 data *testing* serta menghasilkan nilai error yang sangat kecil yaitu 0.00000000000000272. Sehingga dari pengklasteran ini dapat menjadi sebagai bahan untuk evaluasi untuk pemerintah dalam meningkatkan ketersediaan produksi pangan sebagai minimasi impor beras, khususnya pemaksimalan pada lahan sawah yang luas yaitu daerah Jawa, yaitu Jawa Barat, Jawa Tengah dan Jawa Timur.

Berdasarkan luas lahan, pada daerah lahan dengan luas sedang, pemerintah dan masyarakat harus bekerja sama menekan faktor yang membuat semakin menyempitnya lahan seperti maraknya pembangunan perumahan, pembangunan mall dan gedung bertingkat dan seharusnya pemerintah dan masyarakat bekerjasama dalam menambahkan dan menggunakan metode-metode baru dalam peningkatan produksi pertanian.

REFERENSI

- [1] Khairunnisa. Kompas, 12 april 2011. Diakses pada 20 Juni 2016. <http://kompas.com>
- [2] Sawit, Husein. 2006. *Indonesia Dalam Tatanan Perubahan Perdagangan Beras Dunia*. Jurnal. Bogor.Pusat Analisis Social Ekonomi Dan Kebijakan Pertanian

- [3] Priyono.2011. *Alih Fungsi Lahan Pertanian Merupakan Suatu Kebutuhan Atau Tantangan*. Jurnal. Bengkulu
- [4] Nurjanah. Farmadi, Andi. 2014. *Implementasi Metode Fuzzy C-Means Pada Sistem Clustering Data Varietas Padi*. Jurnal ISJD
- [5] Hammouda, Khaled. *A Comparative Studi Of Data Clustering Techniques*. Journal Universitas Of Waterloo. Ontario. Canada
- [6] Asyali, Musa. Alci, Musa. *Reliability Analysis Of Microarray Data Using Fuzzy C-Means And Normal Mixture Modeling Based Classification Methods*. Jurnal Bioinformatics Vol.21 No.5 2005, Pages 644-649
- [7] Kusumadewi, Sri. Purnomo, Hadi. 2010. *Aplikasi Logika Fuzzy Untuk Pendukung Keputusan Eds.2*. Yogyakarta. Graha Ilmu
- [8] Kusumadewi, Sri. Purnomo, Hadi. 2010. *Aplikasi Logika Fuzzy Untuk Pendukung Keputusan Eds.2*. Yogyakarta. Graha Ilmu
- [9] Kusumadewi, Sri hartati. 2006. *Neuro fuzzy: integrasi sistem fuzzy dan jaringan syaraf*. Yogyakarta. Graha ilmu
- [10] Kusumadewi, Sri. Purnomo, Hadi. 2010. *Aplikasi Logika Fuzzy Untuk Pendukung Keputusan Eds.2*. Yogyakarta. Graha Ilmu