

ABSTRAK

Produksi padi merupakan indikator utama dalam ketahanan pangan di Indonesia. Ketidakstabilan produksi padi dari tahun ke tahun dapat berdampak pada ketergantungan terhadap impor dan menurunnya ketahanan pangan nasional. Pulau Madura, dengan empat kabupatennya yaitu Bangkalan, Sampang, Pamekasan, dan Sumenep, merupakan salah satu wilayah penghasil padi yang memiliki fluktuasi produksi cukup signifikan. Oleh karena itu, dibutuhkan suatu metode untuk mengelompokkan daerah berdasarkan produktivitas padi guna mendukung pengambilan kebijakan yang lebih tepat sasaran. Penelitian ini mengusulkan perbandingan metode clustering untuk menganalisis produktivitas padi di Pulau Madura. Pada penelitian ini akan mengusulkan tiga metode *clustering*, yaitu *K-Means*, *K-Medoids*, dan *Fuzzy C-Means*. Dataset yang digunakan terdiri dari 417 data produksi padi dari tahun 2018 hingga 2023, dengan empat parameter utama yaitu luas tanam, luas panen, produktivitas, dan produksi. Proses penelitian meliputi tahapan preprocessing data, normalisasi menggunakan *Min-Max*, dan penerapan metode *clustering*. Evaluasi hasil *clustering* dilakukan dengan *Davies Bouldin Index* (DBI), dan dibandingkan dari segi waktu eksekusi dan jumlah iterasi. Dari hasil penelitian diperoleh algoritma *K-Means* memberikan performa terbaik pada jumlah cluster optimal $k=2$ dengan nilai DBI sebesar 0.7057. Sementara itu, metode *Fuzzy C-Means* menghasilkan nilai DBI sebesar 0.7198, sedangkan *K-Medoids* memperoleh nilai DBI sebesar 0.7735. Dengan nilai evaluasi yang lebih baik dan waktu eksekusi yang lebih efisien, *K-Means* menjadi metode paling optimal dalam mengelompokkan wilayah produktivitas padi di Pulau Madura.

Kata kunci : Produksi Padi, Clustering, K-Means, K-Medoids, Fuzzy C-Means, Davies Bouldin Index

ABSTRACT

Rice production is a key indicator of food security in Indonesia. The instability of rice production from year to year can impact dependence on imports and reduce national food security. Madura Island, with its four regencies: Bangkalan, Sampang, Pamekasan, and Sumenep, is one of the rice-producing regions with significant production fluctuations. Therefore, a method is needed to group regions based on rice productivity to support more targeted policymaking. This study proposes a comparison of clustering methods to analyze rice productivity on Madura Island. This study will propose three clustering methods: K-Means, K-Medoids, and Fuzzy C-Means. The dataset used consists of 417 rice production data from 2018 to 2023, with four main parameters: planted area, harvested area, productivity, and production. The research process includes data preprocessing, normalization using Min-Max, and application of the clustering method. Evaluation of clustering results was carried out using the Davies Bouldin Index (DBI), and compared in terms of execution time and number of iterations. The results showed that the K-Means algorithm provided the best performance at the optimal number of clusters $k = 2$ with a DBI value of 0.7057. Meanwhile, the Fuzzy C-Means method produced a DBI value of 0.7198, while K-Medoids obtained a DBI value of 0.7735. With better evaluation values and more efficient execution time, K-Means is the most optimal method in grouping rice productivity areas on Madura Island.

Keywords: Rice Production, Clustering, K-Means, K-Medoids, Fuzzy C-Means, Davies Bouldin Index

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERNYATAAN ORISINALITAS	iii
MOTTO & PERSEMBAHAN.....	iv
ABSTRAK	v
<i>ABSTRACT</i>	vi
KATA PENGANTAR.....	vii
DAFTAR ISI	ix
DAFTAR GAMBAR	xii
DAFTAR TABEL.....	xiv
DAFTAR PROGRAM	xv
DAFTAR LAMPIRAN	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.2.1 Permasalahan.....	5
1.2.2 Metode Usulan	6
1.2.3 Pertanyaan penelitian	6
1.3 Tujuan Dan Manfaat.....	6
1.3.1 Tujuan Penelitian.....	6
1.3.2 Manfaat Penelitian	6
1.4 Batasan Masalah.....	7
1.5 Sistematika Penulisan	7
BAB II TINJAUAN PUSTAKA	9

2.1	Padi.....	9
2.2	Clustering.....	9
2.3	K–Means	10
2.4	K-Medoids	12
2.5	Fuzzy C-Means	13
2.6	Normalisasi Data.....	15
2.7	Evaluasi.....	16
2.9.1	Davies Bouldin Index.....	16
2.9.2	Waktu Eksekusi (<i>Execution Time</i>)	17
2.8	Penelitian Terdahulu.....	17
BAB III METODE PENELITIAN.....		21
3.1	Pengumpulan Data	22
3.1.1	Sumber Data.....	22
3.1.2	Variabel Penelitian	23
3.2	Preprocessing	24
3.2.1	Data Cleaning.....	24
3.2.2	Normalisasi Data.....	24
3.3	<i>Clustering</i> Data Produksi Padi.....	24
3.3.1	K-Means.....	24
3.3.2	K-Medoids	26
3.3.3	Fuzzy C-Means	27
3.4	Evaluasi Model <i>Clustering</i>	29
3.5	Skenario Uji Coba.....	30
BAB IV HASIL DAN PEMBAHASAN.....		31
4.1	Lingkungan Uji Coba.....	31
4.2	Deskripsi Data.....	33

4.3	Preprocessing Data	35
4.3.1	Deteksi Missing Value.....	35
4.3.2	Normalisasi Data	36
4.4	K-Means.....	39
4.4.1	Proses K-Means	39
4.4.2	Hasil Clustering K-Means.....	41
4.5	K-Medoids	45
4.5.1	Proses K-Medoids	45
4.5.2	Hasil Clustering K-Medoids	48
4.6	Fuzzy C-Means	52
4.6.1	Proses Fuzzy C-Means.....	52
4.6.2	Hasil Clustering Fuzzy C-Means	54
4.7	Hasil Skenario Uji Coba.....	59
4.7.1	Skenario Pertama	59
4.7.2	Skenario Kedua	61
4.7.3	Skenario Ketiga.....	63
4.8	Hasil Analisa Pengujian	65
4.9	Hasil Analisis Clustering.....	69
4.10	Implementasi.....	71
BAB V PENUTUP.....		81
5.1	Kesimpulan	81
5.2	Saran.....	81
DAFTAR PUSTAKA		82
LAMPIRAN.....		86
DAFTAR RIWAYAT HIDUP		96

DAFTAR GAMBAR

Gambar 1. 1 Data Produksi Tanaman Padi Madura Tahun 2018-2023	2
Gambar 2. 1 Contoh Clustering	10
Gambar 3. 1 Arsitektur Penelitian	21
Gambar 3. 2 Algoritma Proses Clustering K-Means.....	25
Gambar 3. 3 Algoritma Proses Clustering K-Medoids	26
Gambar 3. 4 Algoritma Proses Clustering Fuzzy C-Means	28
Gambar 4. 1 Hasil Perhitungan Jumlah Missing Value.....	36
Gambar 4. 2 Hasil Nilai Davies-Bouldin Index Menggunakan Metode K-Means	59
Gambar 4. 3 Hasil Nilai Davies-Bouldin Index Menggunakan Metode K-Medoids	61
Gambar 4. 4 Hasil Nilai Davies-Bouldin Index Menggunakan Metode Fuzzy C- Means	63
Gambar 4. 5 Hasil Perbandingan Clustering berdasarkan Davies-Bouldin Index	68
Gambar 4. 6 Hasil Perbandingan Clustering berdasarkan Waktu eksekusi dan Jumlah Iterasi	68
Gambar 4. 7 Fitur Ringkasan Penelitian Bagian Latar Belakang	71
Gambar 4. 8 Fitur Ringkasan Penelitian Bagian Pertanyaan, Tujuan, dan Alur Penelitian.....	72
Gambar 4. 9 Fitur Proses Clustering Data Bagian Input Data	72
Gambar 4. 10 Fitur Proses Clustering Data Bagian Cek Missing Value Pada Tahap Preprocessing.....	73
Gambar 4. 11 Fitur Proses Clustering Data Bagian Normalisasi Data Pada Tahap Preprocessing.....	73
Gambar 4. 12 Fitur Proses Clustering Data Bagian Evaluasi Davies-Bouldin Index, Waktu eksekusi, dan Jumlah Iterasi Pada K-Means.....	74
Gambar 4. 13 Fitur Proses Clustering Data Bagian Output Evaluasi Davies-Bouldin Index, Waktu eksekusi, dan Jumlah Iterasi Pada K-Means.....	74
Gambar 4. 14 Fitur Proses Clustering Data Bagian Hasil Clustering Pada K-Means	75
Gambar 4. 15 Fitur Proses Clustering Data Bagian Evaluasi Davies-Bouldin Index, Waktu eksekusi, dan Jumlah Iterasi Pada K-Medoids	76

Gambar 4. 16 Fitur Proses Clustering Data Bagian Output Evaluasi Davies-Bouldin Index, Waktu eksekusi, dan Jumlah Iterasi Pada K-Medoids	76
Gambar 4. 17 Fitur Proses Clustering Data Bagian Hasil Clustering Pada K-Medoids.....	77
Gambar 4. 18 Fitur Proses Clustering Data Bagian Evaluasi Davies-Bouldin Index, Waktu eksekusi, dan Jumlah Iterasi Pada Fuzzy C-Means	77
Gambar 4. 19 Fitur Proses Clustering Data Bagian Output Evaluasi Davies-Bouldin Index, Waktu eksekusi, dan Jumlah Iterasi Pada Fuzzy C-Means	78
Gambar 4. 20 Fitur Proses Clustering Data Bagian Hasil Clustering Pada Fuzzy C-Means	78
Gambar 4. 21 Fitur Perbandingan Metode Clustering	79
Gambar 4. 22 Fitur Prediksi Hasil Clustering.....	80

DAFTAR TABEL

Tabel 2. 1 Kriteria Silhoutte Coefficient	20
Tabel 2. 2 Penelitian Terdahulu.....	17
Tabel 3. 1 Dataset Produksi Padi.....	22
Tabel 3. 2 Atribut Data	23
Tabel 3. 3 Skenario Pengujian Perbandingan Metode Clustering.....	30
Tabel 4. 1 Spesifikasi Perangkat Keras	31
Tabel 4. 2 Spesifikasi Perangkat Lunak	31
Tabel 4. 3 Daftar Library.....	32
Tabel 4. 4 Tabel DataFrame 5 baris.....	35
Tabel 4. 5 Data Hasil Normalisasi Menggunakan MinMax Scaler.....	37
Tabel 4. 6 Data Hasil Normalisasi dengan Pengembalian Kabupaten, Kecamatan, dan Tahun	38
Tabel 4. 7 Hasil Clustering K-Means Untuk Data Tahun 2018.....	42
Tabel 4. 8 Hasil Clustering K-Medoids Untuk Data Tahun 2018	49
Tabel 4. 9 Hasil Clustering Fuzzy C-Means Untuk Data Tahun 2018.....	55
Tabel 4. 10 Hasil Perbandingan Evaluasi DBI, Waktu eksekusi dan Jumlah Iterasi Pada Metode K-Means.....	60
Tabel 4. 11 Hasil Perbandingan Evaluasi DBI, Waktu eksekusi dan Jumlah Iterasi Pada Metode K-Medoids.....	62
Tabel 4. 12 Hasil Perbandingan Evaluasi DBI, Waktu eksekusi dan Jumlah Iterasi Pada Metode Fuzzy C-Means	64
Tabel 4. 13 Hasil Nilai Uji Coba DBI, Waktu eksekusi dan Jumlah Iterasi Menggunakan Metode K-Means.....	65
Tabel 4. 14 Hasil Nilai Uji Coba DBI, Waktu eksekusi dan Jumlah Iterasi Menggunakan Metode K-Medoids	66
Tabel 4. 15 Hasil Nilai Uji Coba DBI, Waktu eksekusi dan Jumlah Iterasi Menggunakan Metode Fuzzy C-Means	67
Tabel 4. 16 Titik Pusat Cluster Pada Dataset Menggunakan Metode K-Means ...	70

DAFTAR PROGRAM

Program 4. 1 Import Library	33
Program 4. 2 Menampilkan Dataset.....	34
Program 4. 3 Menampilkan Jumlah Missing Value	35
Program 4. 4 Normalisasi Data	36
Program 4. 5 Penggabungan Kolom Dataset Setelah di Normalisasi	38
Program 4. 6 Evaluasi Hasil Clustering K-Means dengan Uji Cluster 2 hingga 10	39
Program 4. 7 Proses Clustering Algoritma K-Means.....	41
Program 4. 8 Evaluasi Hasil Clustering K-Medoids dengan Uji Cluster 2 hingga 10	46
Program 4. 9 Proses Clustering Algoritma K-Medoids	48
Program 4. 10 Evaluasi Hasil Clustering Fuzzy C-Means dengan Uji Cluster 2 hingga 10.....	52
Program 4. 11 Proses Clustering Algoritma Fuzzy C-Means	54

DAFTAR LAMPIRAN

Lampiran 1 Data Penelitian.....	86
---------------------------------	----