

Perbandingan Algoritma *K-Means* Dan *Fuzzy C-Means* Untuk Pengelompokkan Daerah Produktivitas Tanaman Padi Kalimantan Barat

Aryo Nugroho¹, Putri Yuli Utami², Rizki Surtiyan Surya³

^{1,2,3}Program Studi Sistem Informasi, Fakultas Teknik Dan Ilmu Komputer, Universitas Muhammadiyah Pontianak

Corresponding Author: aryo.nugroho0505@gmail.com

Published: Agustus, 2025

ABSTRAK

Indonesia adalah negara agraris, di mana sektor pertanian berperan penting dalam pembangunan nasional dengan menyediakan pangan untuk seluruh penduduk. Luasnya wilayah di Kalimantan Barat dan variasi produktivitas padi di berbagai daerah menunjukkan perlunya pengelompokan wilayah berdasarkan hasil produksinya. Teknik clustering dalam data mining, seperti K-Means dan Fuzzy C-Means, dapat digunakan untuk mengelompokkan daerah produktivitas padi berdasarkan data, sehingga kebijakan dan intervensi bisa dilakukan dengan lebih terfokus. Fuzzy C-Means 2020 Cluster 1 tinggi 4 daerah, Cluster 0 sedang 7 daerah, Cluster 2 rendah 3 daerah. 2021 Cluster 1 tinggi 6 daerah, Cluster 2 sedang 5 daerah, Cluster 0 rendah 3 daerah. 2022 Cluster 1 tinggi 2 daerah, Cluster 0 sedang 5 daerah, Cluster 2 rendah 7 daerah. 2023 Cluster 1 tinggi 5 daerah, Cluster 2 sedang 6 daerah, Cluster 0 rendah 3 daerah. K-Means 2020 Cluster 0 tinggi 4 daerah, Cluster 1 sedang 3 daerah, Cluster 2 rendah 7 daerah. 2021 Cluster 0 tinggi 6 daerah, Cluster 1 sedang 5 daerah, Cluster 2 rendah 3 daerah. 2022 Cluster 2 tinggi 4 daerah, Cluster 0 rendah 5 daerah, Cluster 1 sedang 5 daerah. 2023: Cluster 1 tinggi 5 daerah, Cluster 0 sedang 6 daerah, Cluster 2 rendah 3 daerah. Secara keseluruhan, Fuzzy C-Means unggul pada 8 duster, sementara K-Means hanya unggul pada 4 duster. Hal ini mengidentifikasi bahwa Fuzzy C-Means lebih mampu membentuk kluster.

Kata kunci: Algoritma K-Means, Algoritma Fuzzy C-Means, DBI, Kalimantan Barat, Produktivitas

ABSTRACT

Indonesia is an agrarian country where the agricultural sector plays a crucial role in national development by providing food for the entire population. The vast area of West Kalimantan and the variations in rice productivity across different regions indicate the need for regional grouping based on production outcomes. Clustering techniques in data mining, such as K-Means and Fuzzy C-Means, can be used to group regions based on rice productivity data, allowing for more focused policies and interventions. Fuzzy C-Means in 2020 had Cluster 1 as high with 4 regions, Cluster 0 as medium with 7 regions, and Cluster 2 as low with 3 regions. In 2021, Cluster 1 (high) comprised 6 regions, Cluster 2 (medium) had 5 regions, and Cluster 0 (low) contained 3 regions. For 2022, Cluster 1 (high) covered 2 regions, Cluster 0 (medium) encompassed 5 regions, and Cluster 2 (low) included 7 regions. By 2023, Cluster 1 (high) had 5 regions, Cluster 2 (medium) had 6 regions, and Cluster 0 (low) had 3 regions. K-Means in 2020 placed Cluster 0 (high) with 4 regions, Cluster 1 (medium) with 3 regions, and Cluster 2 (low) with 7 regions. In 2021, Cluster 0 (high) comprised 6 regions, Cluster 1 (medium) had 5 regions, and Cluster 2 (low) included 3 regions. In 2022, Cluster 2 (high) contained 4 regions, Cluster 0 (low) had 5 regions, and Cluster 1 (medium) consisted of 5 regions. For 2023, Cluster 1 (high) contained 5 regions, Cluster 0 (medium) had 6 regions, and Cluster 2 (low) contained 3 regions. Overall, Fuzzy C-Means outperformed in 8 out of 12 cases, while K-Means only excelled in 4. This demonstrates that Fuzzy C-Means is more consistent in producing optimal DBI (Dunn's Validity Index) values.

Keywords: K-Means Algorithm, Fuzzy C-Means Algorithm, Davies-Bouldin Index (DBI), West Kalimantan, Productivity

© The Author(s). 2021 Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0>

PENDAHULUAN

Pembangunan pertanian merupakan bagian integral dari pembangunan ekonomi dan masyarakat secara keseluruhan. Pertanian berkontribusi signifikan kepada masyarakat dan memastikan pembangunan yang inklusif bagi banyak orang yang bergantung pada sektor ini untuk

kehidupan mereka, baik saat ini maupun di masa mendatang. Sebagian besar penduduk Indonesia bergantung pada sektor pertanian, termasuk sebagai petani padi. Dengan kondisi lingkungan mendukung serta iklim cocok untuk budidaya padi. Pada tahun 2018, Indonesia berhasil menjadi salah satu negara produksi beras terbesar di seluruh dunia (Adi et al., n.d.2018).

Luasnya wilayah di Kalimantan Barat dan variasi produktivitas padi di berbagai daerah menunjukkan perlunya pengelompokan wilayah berdasarkan hasil produksinya. Pengelompokan ini akan mempermudah perencanaan kebijakan yang lebih tepat sasaran, serta membantu pemerintah dan pemangku kepentingan dalam merancang program peningkatan produktivitas yang lebih efektif dan efisien. Teknik clustering dalam data mining, seperti K-Means dan Fuzzy C-Means, dapat digunakan untuk mengelompokkan daerah produktivitas padi berdasarkan data, sehingga kebijakan dan intervensi bisa dilakukan dengan lebih terfokus.

Dengan menggunakan analisis Cluster, pengelompokan dilakukan berdasarkan karakteristik masing – masing. Ada tiga kategori yang diidentifikasi dalam kelompok produksi yaitu tinggi, rendah, sedang (Azizah, 2023). Menggunakan metode K-Means untuk menganalisis pengelompokan daerah produktivitas tanaman padi di Kalimantan Barat, di samping itu. Sebagai alternatif, algoritma Fuzzy C-means (FCM) memperkenalkan konsep keanggotaan fuzzy, tidak seperti K-means yang menugaskan setiap data secara eksklusif ke sebuah cluster. FCM memberikan tingkat keanggotaan pada setiap data di setiap cluster, sehingga memberikan fleksibilitas yang lebih besar. Keanggotaan yang fleksibel ini memungkinkan data menjadi anggota beberapa cluster dengan tingkat keanggotaan yang berbeda-beda, sehingga memungkinkan penanganan data yang lebih baik tanpa batasan tetap antar cluster. Selain itu, FCM seringkali dapat memberikan hasil yang lebih stabil dan akurat dibandingkan K-means pada kumpulan data yang kompleks. Dalam konteks produktivitas padi Cluster regional, FCM dapat menangkap perubahan dan ketidakpastian dengan lebih baik, sehingga menghasilkan pemahaman yang lebih mendalam tentang pola produktivitas (Butarbutar et al., 2016).

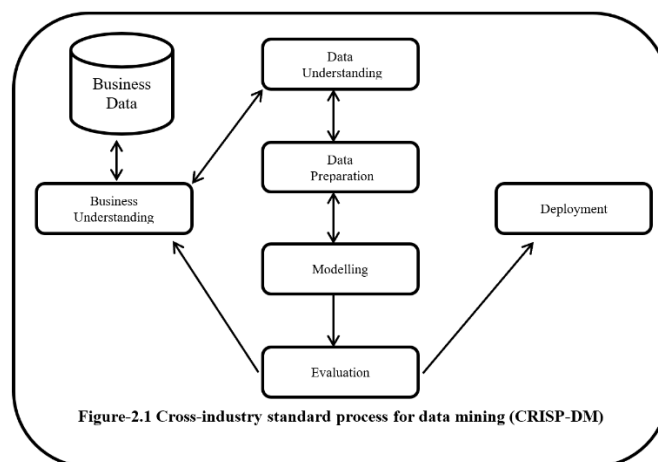
Terdapat penelitian sebelumnya pernah dilakukan diantaranya oleh (Klau et al., 2023) tentang perbandingan metode K-Means dan Fuzzy c-means terhadap indikator kemiskinan Nusa Tenggara Timur (NTT) 2021 hasilnya adalah dapat diketahui metode Davies Bouldin Index (DBI) untuk penentuan metode terbaik adalah metode K-Means dengan nilai sebesar 1,615 dan R-Squared sebesar 27.2% dengan jumlah cluster sebanyak 2, sehingga dapat disimpulkan bahwa metode terbaik dalam pengelompokan kabupaten/kota di provinsi NTT berdasarkan indikator variabel tingkat kemiskinan adalah metode K-Means clustering dengan banyaknya cluster 2. Selanjutnya penelitian yang dilakukan oleh (Dame Tampubolon et al., 2021) menggunakan K-Means dan K-Medoid. Penelitian ini bertujuan mengelompokkan provinsi berdasarkan tingkat tindak kriminalitas yang terjadi di Indonesia menggunakan K-Means dan K-Medoid. Hasil algoritma K-Means menunjukkan 6 provinsi termasuk dalam kelompok tingkat kriminalitas tinggi, sementara 28 provinsi dalam kelompok tingkat

kecriminalitas rendah. Sedangkan hasil algoritma K-Means menunjukkan 7 provinsi dalam kelompok tingkat kriminalitas tinggi dan 27 provinsi dalam kelompok tingkat kriminalitas rendah.

Berdasarkan Penelitian ini bertujuan untuk mengduster wilayah-wilayah di Kalimantan Barat berdasarkan tingkat produktivitas padi menggunakan pendekatan metode K-Means serta Fuzzy C-means. Dengan membandingkan kedua algoritma ini, penelitian diharapkan dapat menentukan metode mana yang lebih efektif dalam mengduster wilayah-wilayah di Kalimantan Barat berdasarkan tingkat produktivitas padi.

METODE PENELITIAN

Penelitian ini mengadopsi model Cross-Industry Standard Process for Data Mining (CRISPDM). Alasan mengapa peneliti menggunakan model CRISP-DM dibandingkan model lainnya adalah karena model CRISPDM merupakan metode standar untuk mengelola data mining, ditunjukkan pada gambar 3.1.



Gambar 1. Tahapan CRIPS-DM

Tahap ini dimulai dengan memahami masalah dan tujuan yang ingin diselesaikan atau tujuan yang ingin dicapai dengan proyek data mining. Tujuan penelitian ini untuk membandingkan dua algoritma pengelompokan, yaitu K-Means dan Fuzzy C-Means, dalam mengidentifikasi kelompok daerah dengan produktivitas padi di Kalimantan Barat. Pada tahap ini, semua data yang tersedia diperiksa dan dipahami secara menyeluruh. Ini termasuk memeriksa dan memahami struktur, kualitas, dan kemungkinan masalah yang ada. Data yang digunakan yaitu (Kabupaten/Kota, tahun, bulan, luas panen padi, produksi padi, produktivitas padi) data yang diperoleh dari Badan Pusat Statistik Kalimantan Barat dan Dinas Pangan tahun 2016-2023.

HASIL DAN PEMBAHASAN

Hasil Analisis *Cluster* Algoritma *K-means*

Berikut langkah-langkah perhitungan algoritma K-Means

1. Menentukan jumlah duster

Pada penelitian ini jumlah duster yang ditentukan adalah 3 duster yaitu C1, C2, dan C3.

2. Tentukan titik centroid atau pusat duster.

Titik centroid ditentukan secara acak dari data tanaman padi. Pada tabel 4.1 merupakan nilai titik centroid awal yang akan digunakan dalam proses perhitungan algoritma K-Means dimana C1 merupakan data ke-24, C2 data ke-33 dan C3 adalah data ke-81.

3. Hitung jarak data terhadap masing-masing centroid dengan rumus Eudidean.

Hasil iterasi pertama dengan menggunakan data sebanyak 336 data menghasilkan anggota duster C1 sebanyak 56 anggota C2 sebanyak 203 anggota dan C3 sebanyak 77 anggota.

4. Memperbarui Nilai Titik Centroid

Proses mencari nilai centroid baru adalah dengan membagi hasil total jumlah atribut dengan total jumlah anggota setiap masing-masing duster, selanjutnya hasil tersebut adalah angka yang akan digunakan dalam perhitungan selanjutnya, dengan kata lain hasil tersebut adalah nilai centroid baru..

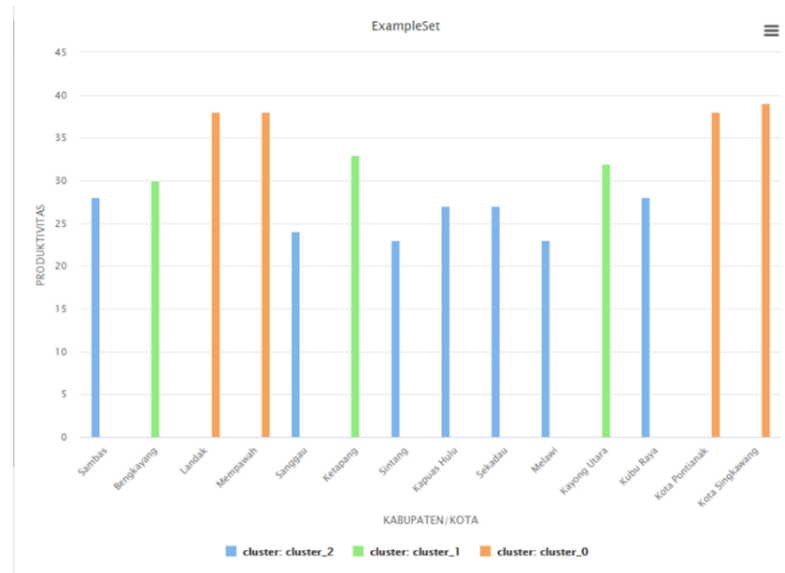
5. Mengulang Iterasi

Ulangi langkah 2 dan seterusnya dengan memakai pusat Cluster yang baru. Pada iterasi ke-2 dihasilkan C1= 208, C2= 41, C3= 87, iterasi ke-3 C1= 45, C2= 92, C3= 198, iterasi ke-4 C1= 76, C2 = 24, C3= 236, iterasi ke-5 C1= 85, C2= 24, C3= 227. Karena pada iterasi ke-6 dan ke-7 anggota Cluster tidak ada yang berubah maka proses dihentikan karena sudah konvergen dengan hasil Cluster "1 sebanyak 87 anggota," Cluster "2 sebanyak 22 anggota dan Cluster 3 sebanyak 227 anggota.

Untuk melihat hasil Clusterisasi daerah berdasarkan produktivitas padi, silakan merujuk pada tabel iterasi ke-7. rincian mengenai pengelompokan daerah sesuai dengan tingkat produktivitas menyajikan hasil centroid iterasi ke-6 dan ke-7 dari analisis Cluster produktivitas padi yang telah dikelompokkan berdasarkan tingkat produktivitasnya, yaitu rendah, sedang, dan tinggi. Centroid dalam konteks ini merepresentasikan nilai rata-rata dari masing-masing variabel yang digunakan dalam proses clustering untuk setiap kelompok produktivitas tersebut.

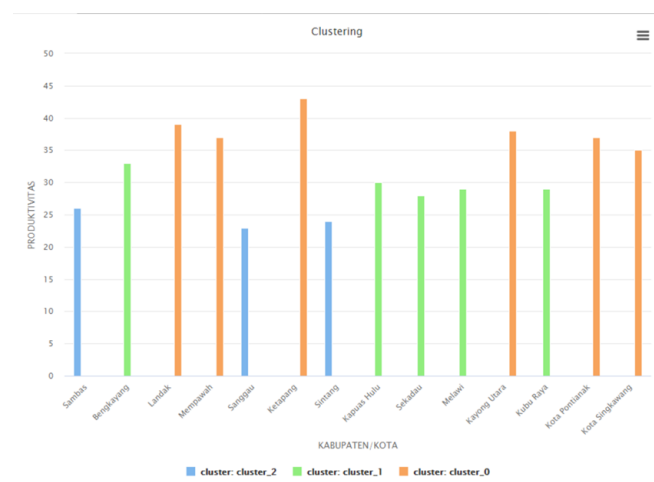
Implementasi *K-Means* menggunakan *Rapidminer*

Berdasarkan dustering menggunakan metode K-Means ditentukan dengan menetapkan parameter $k = 3$ sebagai kelompok dustering Terdapat 3 duster dalam penelitian yaitu duster produktivitas tinggi, produktivitas sedang, produktivitas rendah dari tahun 2020-2023.



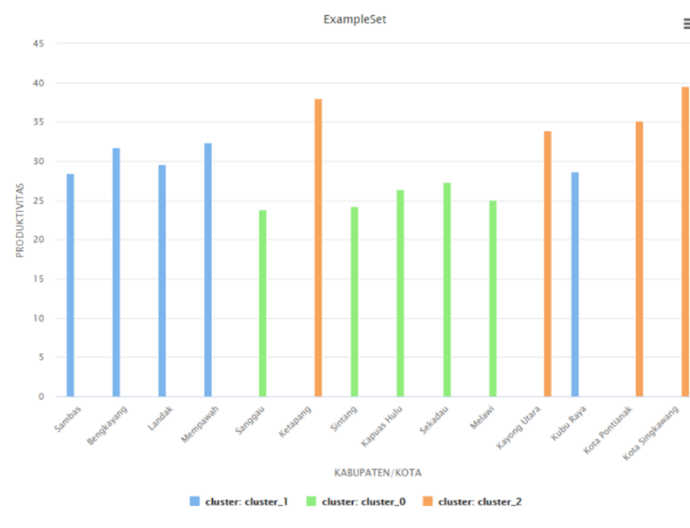
Gambar 2. Grafik Cluster K-Means Daerah Produktivitas Padi 2020

Pada gambar 2 pengelompokan daerah produktivitas padi tahun 2020, grafik menunjukkan distribusi produktivitas di beberapa kabupaten/kota yang dikelompokkan ke dalam tiga duster. Cluster 0 berwarna orange beranggotakan kabupaten/kota dengan produktivitas tertinggi dengan nilai produktivitas padi di atas 35 ton pertahun, seperti Mempawah, Landak, Kota Pontianak, dan Kota Singkawang. Cluster 1 berwarna hijau beranggotakan kabupaten/kota dengan produktivitas sedang, dengan rentang 30-35 ton pertahun seperti Bengkayang, Ketapang dan Kayong Utara. Cluster 2 berwarna biru rendah beranggotakan kabupaten/kota dengan produktivitas rendah yang memiliki nilai produktivitas di bawah 30 ton pertahun seperti Sambas, Sanggau, Sintang, Kapuas Hulu, Sekadau, Melawi dan Kubu Raya.



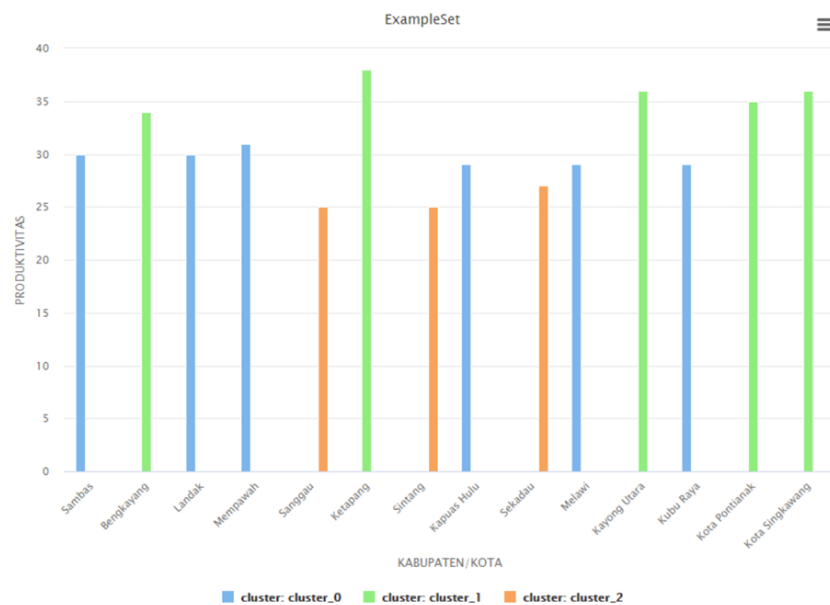
Gambar 3. Grafik Cluster K-Means Daerah Produktivitas Padi 2021

Pada gambar 3. pengelompokan daerah produktivitas padi tahun 2021, grafik menunjukkan distribusi produktivitas di beberapa kabupaten/kota yang dikelompokkan ke dalam tiga Cluster. Cluster 0 berwarna orange mencakup kabupaten/kota dengan produktivitas tertinggi dengan nilai produktivitas 40 ton pertahun, seperti Landak, Mempawah, Ketapang, Kayong Utara, Kota Pontianak dan Singkawang. Daerah-daerah ini menunjukkan produktivitas yang konsisten tinggi. Cluster 1 berwarna hijau meliputi kabupaten/kota dengan produktivitas sedang dengan nilai produktivitas 30 ton pertahun, seperti Bengkayang, Kapuas Hulu, Sekadau, Melawi dan Kubu Raya. Cluster 2 berwarna biru dengan produktivitas rendah, mencakup kabupaten/kota seperti Sambas, Sanggau, Sintang, dengan nilai produktivitas berkisar antara 25 ton pertahun.



Gambar 4. Grafik Cluster K-Means Daerah Produktivitas Padi 2022

Pada gambar pengelompokan daerah produktivitas padi tahun 2022, grafik menunjukkan distribusi produktivitas di beberapa kabupaten/kota yang dikelompokkan ke dalam tiga Cluster. Cluster 2 berwarna orange mencakup kabupaten/kota dengan produktivitas tertinggi dengan nilai produktivitas sekitar 35 ton pertahun, seperti Ketapang, Kota Pontianak dan Kota Singkawang. Daerah-daerah ini menunjukkan produktivitas yang konsisten tinggi. Cluster 0 berwarna hijau meliputi kabupaten/kota dengan produktivitas rendah dengan nilai produktivitas 25 ton pertahun, seperti Sanggau, Sintang, Kapuas Hulu, Sekadau, Melawi. Cluster 1 berwarna biru produktivitas sedang mencakup kabupaten/kota seperti Sambas, Bengkayang, Landak, Mempawah, Kubu Raya, dengan produktivitas 30 ton pertahun.



Gambar 5. Grafik Cluster K-Means Daerah Produktivitas Padi 2023

Pada gambar 5 pengelompokan daerah produktivitas padi tahun 2023, grafik menunjukkan distribusi produktivitas di beberapa kabupaten/kota yang dikelompokkan ke dalam tiga Cluster Cluster 1 berwarna hijau mencakup kabupaten/kota dengan produktivitas tertinggi dengan nilai produktivitas 35 ton pertahun, seperti Bengkayang, Ketapang, Kayong Utara, Kota Pontianak, Kota Singkawang. Daerah-daerah ini menunjukkan produktivitas yang konsisten tinggi. Cluster 0 berwarna biru meliputi kabupaten/kota dengan produktivitas sedang 30 ton pertahun, seperti Sambas, Landak, Mempawah, Kapuas Hulu, Melawi, Kuburaya. Cluster 2 orange rendah mencakup kabupaten/kota seperti Sanggau, Sintang, Sekadau, dengan produktivitas berkisar antara 25 ton pertahun. bawah menunjukkan hasil rata-rata produktivitas dan nilai DBI setiap duster. Pada data padi di Kalimantan Barat dari tahun 2020 hingga 2023.

Tabel 1. Tabel Nilai DBI Algoritma K-Means

Tahun	Cluster	Rata-rata Produktivitas	Nilai DBI
2020	0	38	-0.938
	1	31	-2.222
	2	25	-5.347
2021	0	38	-6.972
	1	29	-3.760
	2	24	-2.222
2022	0	25	-2.485
	1	30	-1.124
	2	36	-5.024

2023	0	29	-1.389
	1	35	-2.560
	2	25	-1.556

Pada Tabel 1 tabel nilai DBI algoritma K-Means dari tahun 2020 hingga 2023, produktivitas dan kualitas pemisahan antar duster menunjukkan variasi yang signifikan. Pada tahun 2020, duster 0 memiliki produktivitas tertinggi (38) dan DBI -0,938 yang mendekati 0, menandakan kualitas pemisahan yang baik. Cluster 1 mencatat produktivitas 31 dengan DBI -2,222, sementara duster 2 dengan produktivitas terendah (25) dan DBI -5,347 menunjukkan kualitas pemisahan yang buruk. Pada 2021, produktivitas duster 0 tetap tertinggi (38), namun DBI-nya memburuk menjadi -6,972, sedangkan duster 1 turun produktivitasnya ke 29 dengan DBI -3,760, dan duster 2 menjadi 24 dengan DBI -2,222, menunjukkan sedikit peningkatan pemisahan. Pada 2022, produktivitas duster 0 turun ke 25 dengan DBI -2,485, sementara duster 1 meningkat produktivitasnya ke 30 dengan DBI -1,124 yang paling mendekati 0, menunjukkan pemisahan terbaik. Cluster 2 mencapai produktivitas tertinggi (36) namun memiliki DBI -5,024 yang menunjukkan pemisahan sangat buruk. Di tahun 2023, duster 0 mencatat produktivitas 29 dengan DBI -1,389, menunjukkan kualitas pemisahan terbaik, sedangkan duster 1 memiliki produktivitas tertinggi (35) tetapi DBI -2,560 yang lebih rendah dalam kualitas pemisahan dibandingkan duster 0. Cluster 2 tetap terendah dalam produktivitas (25) dengan DBI -1,556, menunjukkan sedikit peningkatan dari tahun sebelumnya, namun masih di bawah duster 0.

Hasil Analisis Algoritma *Fuzzy C-Means*

Proses perhitungan manual algoritma Fuzzy C-Means :

1. Pada Proses awal dalam menghitung Fuzzy C-Means dengan memasukan data sebanyak 336 yang akan di duster. Dengan atribut data sebagai berikut :
 X_{i3} = Luas Panen
 X_{i4} = Produksi
 X_{i5} = Produktivitas
Jumlah duster = 3
2. Tahapan dalam perhitungan menggunakan Fuzzy C-Means pada iterasi 1 adalah pertama kali dilakukan penentuan tiga nilai partisi awal secara random dengan rentang nilai 0-1. Membentuk matriks partisi awal U (derajat keanggotaan dalam duster). Menghitung Pusat Cluster

Pada proses menghitung pusat duster. hasilnya diperoleh dari hasil pangkat dari matriks U pada setiap kolomnya dengan $w = 2$ dan akan dipakai pada perkalian dengan matriks X . Nilai dari total adalah hasil penjumlahan dari setiap kolomnya dan akan dipakai

untuk menghitung hasil pusat duster 1,2,3. merupakan hasil dari proses perkalian antara kolom μ_i^w1 dengan atribut pada matriks X dan nilai total adalah hasil penjumlahan untuk setiap kolomnya dan nilai-nilai dari total tersebut akan dipakai dalam menghitung hasil pusat duster 1.

Selanjutnya pada Tabel 4.13 adalah proses hasil perkalian antara kolom μ_i^w2 pada setiap kolom pada matriks X dan total adalah hasil penjumlahan untuk setiap kolomnya dan nilai-nilai dari total tersebut akan dipakai untuk menghitung hasil pusat duster 2. merupakan proses hasil perkalian antara kolom μ_i^w3 pada setiap kolom pada matriks X dan total adalah hasil penjumlahan untuk setiap kolomnya dan nilai dari total tersebut akan dipakai untuk menghitung hasil pusat duster 3.

proses hasil perhitungan pusat duster 1, duster 2 dan pusat duster 3. Hasil untuk duster 1 pada baris pertama diperoleh dari baris pada hasil penjumlahan setiap total pada tabel 4.12 dibagi dengan nilai total pada kolom μ_i^w1 dan untuk duster 2 pada baris kedua hasilnya diperoleh dari baris pada hasil penjumlahan setiap total pada tabel 4.13 dibagi dengan nilai total pada kolom μ_i^w2 , untuk duster 3 pada baris ketiga hasilnya diperoleh dari baris pada hasil penjumlahan setiap total pada tabel 4.14 dibagi dengan nilai total pada kolom μ_i^w3 .

3. Menghitung Fungsi Objektif

Proses menghitung fungsi objektif dimana kolom $C1$ hasilnya diperoleh dari setiap baris pada matriks X dikurangi dengan duster 1 pada Tabel 4.15 kemudian dipangkatkan 2. Selanjutnya hasil penjumlahan setiap baris pada kolom $C1$ dapat dilihat pada kolom $Tc1$. Pada kolom $Tc1 * \mu_i^w1$ hasilnya diperoleh dari perkalian antara kolom $Tc1$ dengan kolom μ_i^w1 pada Tabel 4.11. Hasil pada kolom $Tc1$ akan dipakai untuk mencari matriks U baru dan hasil pada kolom $Tc1 * \mu_i^w1$ akan dipakai untuk menghitung nilai fungsi objektif ($P1$).

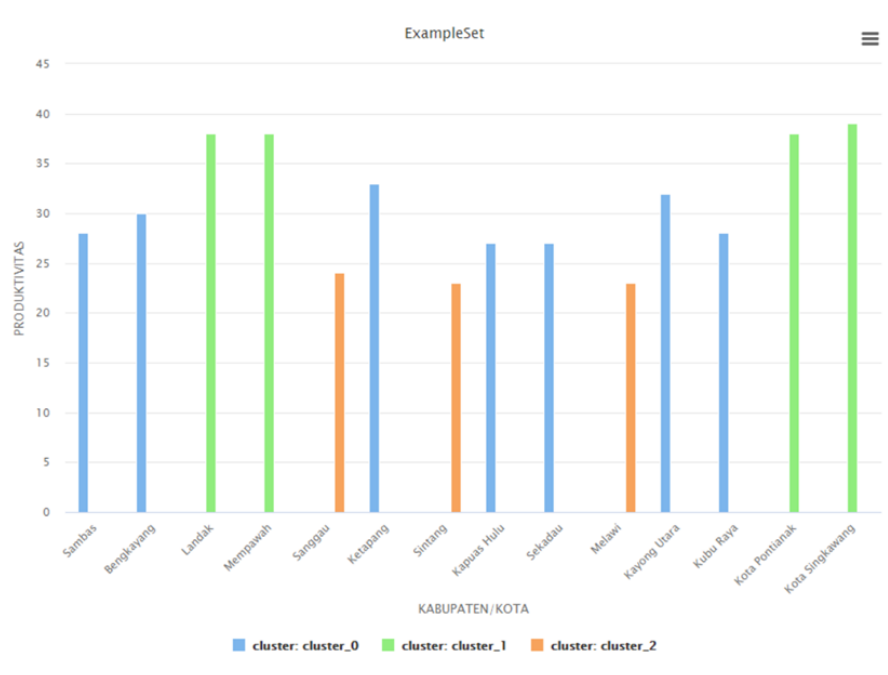
Proses menghitung fungsi objektif dimana kolom $C2$ hasilnya diperoleh dari setiap baris pada matriks X dikurangi dengan duster 2 pada tabel 4.15. Kemudian dipangkatkan 2. Selanjutnya hasil penjumlahan setiap baris pada kolom $C2$ dapat dilihat pada kolom $Tc2$ kemudian pada kolom $Tc2 * \mu_i^w2$ hasilnya diperoleh dari perkalian antara kolom $Tc2$ dengan kolom μ_i^w2 pada tabel 4.11. Hasil pada kolom $Tc2$ akan dipakai untuk mencari matriks U baru dan hasil pada kolom $Tc2 * \mu_i^w2$ akan dipakai untuk menghitung nilai fungsi objektif ($P1$).

Proses menghitung fungsi objektif dimana kolom $C3$ hasilnya diperoleh dari setiap baris pada matriks X dikurangi dengan duster 3 pada Tabel 4.15. Kemudian dipangkatkan 2. Selanjutnya hasil penjumlahan setiap baris pada kolom $C3$ dapat dilihat pada kolom $Tc3$

kemudian pada kolom $Tc3 * \mu^{w3}$ hasilnya diperoleh dari perkalian antara kolom $Tc3$ dengan kolom μ^{w3} pada tabel 4.11. Hasil pada kolom $Tc3$ akan dipakai untuk mencari matriks U baru dan hasil pada kolom $Tc3 * \mu^{w3}$ akan dipakai untuk menghitung nilai fungsi objektif ($P1$).

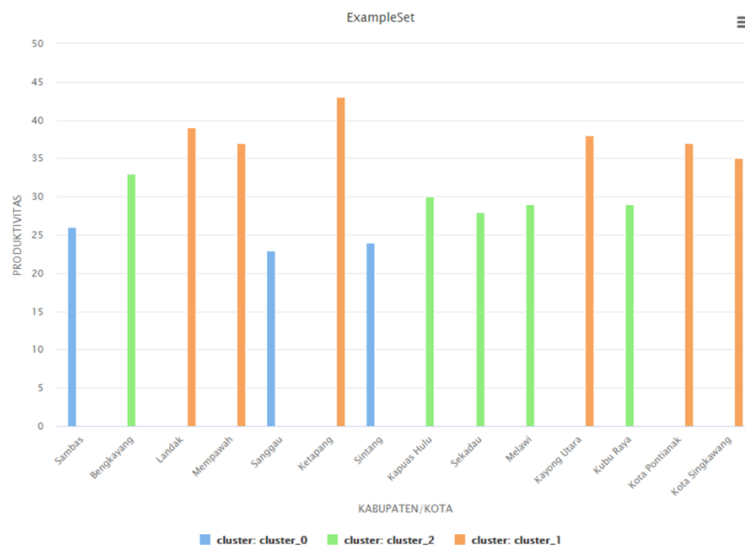
Implementasi Algoritma *Fuzzy C-Means* Menggunakan *RapidMiner*

Berdasarkan *dustering* menggunakan metode *Fuzzy C-Means* ditentukan dengan menetapkan parameter $k = 3$ sebagai kelompok *dustering* Terdapat 3 *duster* dalam penelitian yaitu *duster* produktivitas tinggi, produktivitas sedang, produktivitas rendah dari tahun 2020-2023.



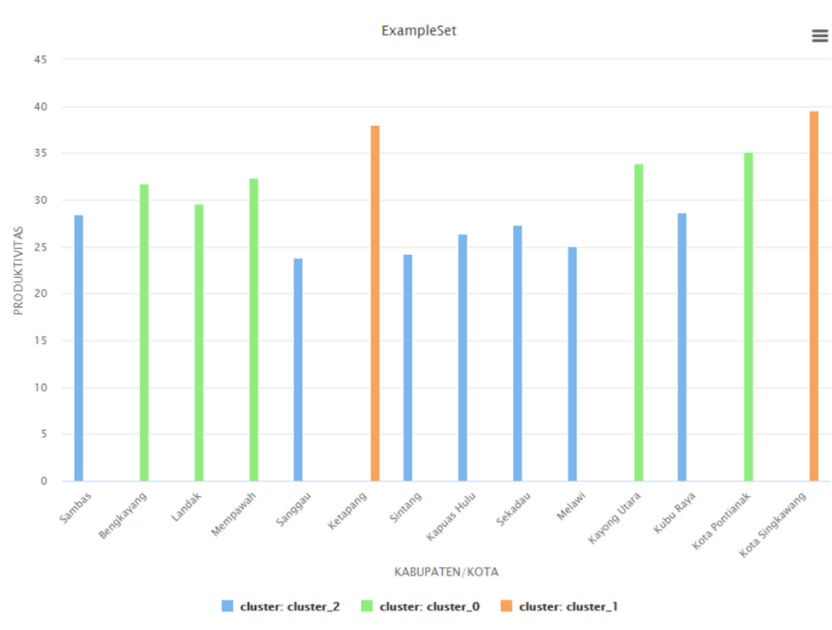
Gambar 6. Grafik Cluster FuzzyC-Means Daerah Produktivitas Padi 2020

Pada gambar 6 grafik produktifitas tahun 2020, grafik menunjukkan distribusi produktivitas di beberapa kabupaten/kota yang dikelompokkan ke dalam tiga Cluster. Cluster 1 berwarna hijau mencakup kabupaten/kota dengan nilai produktivitas tertinggi sekitar 35-40 ton pertahun, seperti Mempawah, Landak, Kota Pontianak, dan Kota Singkawang. Daerah-daerah ini menunjukkan produktivitas yang konsisten tinggi. Cluster 0 berwarna biru meliputi kabupaten/kota dengan nilai produktivitas 30 ton pertahun, seperti Bengkayang, Sambas, Ketapang, Kayong Utara, Kapuas Hulu, Kubu Raya, Sekadau. Cluster 2 berwarna orange mencakup kabupaten/kota seperti Sanggau, Sintang, Melawi, dengan nilai produktivitas berkisar antara 20 sampai 25 ton pertahun.



Gambar 7 Grafik Cluster Fuzzy C-Means Daerah Produktivitas Padi 2021

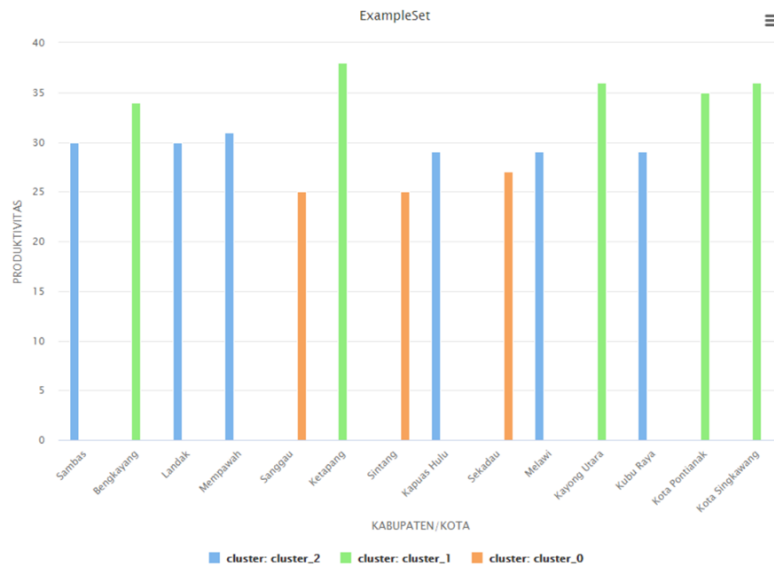
Pada gambar 7 grafik produktifitas tahun 2021, grafik menunjukkan distribusi produktivitas di beberapa kabupaten/kota yang dikelompokkan ke dalam tiga Cluster. Cluster 1 berwarna orange mencakup kabupaten/kota dengan nilai produktivitas tertinggi sekitar 40 ton pertahun, seperti Mempawah, Landak, Kota Pontianak, Kota Singkawang, Kayong Utara, Ketapang. Daerah-daerah ini menunjukkan produktivitas yang konsisten tinggi. Cluster 2 berwarna hijau meliputi kabupaten/kota dengan nilai produktivitas sedang 30 ton pertahun, seperti Bengkayang, Kapuas Hulu, Sekadau, Melawi, Kubu Raya. Cluster 0 berwarna biru rendah mencakup kabupaten/kota seperti Sambas, Sanggau, Sintang, dengan dengan nilai produktivitas berkisar 25 ton pertahun.



Gambar 8. Grafik Cluster Fuzzy C-Means Daerah Produktivitas Padi 2022

Pada gambar 8 grafik produktivitas tahun 2022, grafik menunjukkan distribusi produktivitas di beberapa kabupaten/kota yang dikelompokkan ke dalam tiga Cluster. Cluster 1 berwarna orange

mencakup kabupaten/kota dengan produktivitas tertinggi dengan nilai 40 ton pertahun, seperti Kota Singkawang, Ketapang. Daerah-daerah ini menunjukkan produktivitas yang konsisten tinggi. Cluster 0 berwarna hijau meliputi kabupaten/kota dengan nilai produktivitas sedang berkisaran nilai 35 ton pertahun, seperti Bengkayang, Landak, Mempawah, Kayong Utara, Kota Pontianak. Cluster 2 berwarna biru rendah mencakup kabupaten/kota seperti Sambas, Sanggau, Sintang, Kapuas Hulu, Sekadau, Melawi dan Kubu Raya, dengan nilai produktivitas 30 ton pertahun.



Gambar 9. Grafik Cluster Fuzzy C-Means Daerah Produktivitas Padi 2023

Pada gambar 9 grafik produktivitas tahun 2023, grafik menunjukkan distribusi produktivitas di beberapa kabupaten/kota yang dikelompokkan ke dalam tiga Cluster. Cluster 1 berwarna hijau mencakup kabupaten/kota dengan produktivitas tertinggi dengan nilai 35 ton pertahun, seperti Bengkayang, Ketapang, Kayong Utara, Kota Pontianak, dan Kota Singkawang. Daerah-daerah ini menunjukkan produktivitas yang konsisten tinggi. Cluster 2 berwarna biru meliputi kabupaten/kota dengan produktivitas sedang dengan nilai 30 ton pertahun, seperti Sambas, Landak, Mempawah, Kapuas Hulu, Melawi, Kubu Raya. Cluster 0 berwarna orange rendah mencakup kabupaten/kota seperti Sanggau, Sintang, Sekadau, dengan nilai produktivitas 25 ton pertahun. Tabel 2 di bawah menunjukkan hasil pengelompokan produktivitas padi. Pada data padi di Kalimantan Barat dari tahun 2016 hingga 2023.

Tabel 2 Nilai DBI *Fuzzy C-Means*

Tahun	Cluster	Rata-Rata Produktivitas	DBI
2020	0	29	- 5.061

	1	38	- 0.188
	2	23	- 0.222
2021	0	24	- 1.556
	1	38	- 6.139
	2	29	- 2.960
2022	0	32	- 3.493
	1	38	- 0.624
	2	26	- 3.329
2023	0	25	- 0.889
	1	35	- 1.760
	2	29	- 0.556

Berdasarkan tabel 2 nilai DBI Fuzzy C-Means terbaru ini, berikut adalah analisis singkat mengenai produktivitas rata-rata dan kualitas pemisahan (DBI) dari setiap Cluster selama tahun 2020 hingga 2023:

Pada tahun 2020, Cluster 1 memiliki produktivitas tertinggi (38) dengan DBI -0,188, yang mendekati 0, menunjukkan kualitas pemisahan yang baik. Cluster 0 mencatat produktivitas 29 dan DBI -5,061, menandakan kualitas pemisahan yang sangat buruk. Cluster 2 memiliki produktivitas terendah (23) dengan DBI -0,222, menunjukkan pemisahan yang cukup baik. Pada tahun 2021,

produktivitas Cluster 1 tetap tinggi di 38, tetapi kualitas pemisahan memburuk (DBI -6,139). Cluster 0 mengalami penurunan produktivitas menjadi 24 dan DBI -1,556, menunjukkan kualitas pemisahan yang lebih baik dari tahun sebelumnya. Cluster 2 mencatat produktivitas 29 dengan DBI -2,960. Pada 2022, Cluster 1 tetap dengan produktivitas tertinggi (38) dan DBI -0,624 yang mendekati 0, menunjukkan kualitas pemisahan yang baik. Cluster 0 mencatat peningkatan produktivitas menjadi 32, namun kualitas pemisahan tetap buruk (DBI -3,493). Cluster 2 memiliki produktivitas 26 dengan DBI -3,329. Pada 2023, Cluster 1 kembali mencatat produktivitas tertinggi (35) namun DBI-nya menurun ke -1,760. Cluster 2 memiliki produktivitas 29 dengan DBI -0,556 yang mendekati 0, menunjukkan kualitas pemisahan terbaik. Cluster 0 mencatat produktivitas terendah (25) dengan DBI -0,889.

Pembahasan

Dari hasil perhitungan Algoritma Fuzzy C-Means dan K-means dengan menggunakan $K=3$ diperoleh anggota duster yang berbeda dengan data produktivitas padi dari tahun 2020-2023.

1. Algoritma K-Means

Pada 2020, Cluster 0 orange dengan nilai produktivitas di atas 35 ton pertahun mencakup Mempawah, Landak, Kota Pontianak, dan Kota Singkawang dengan produktivitas tertinggi. Cluster 1 hijau dengan nilai produktivitas 30-35 ton pertahun berisi Bengkayang, Ketapang, dan Kayong Utara untuk produktivitas sedang. Cluster 2 biru dengan nilai produktivitas dibawah 30 ton pertahun terdiri dari Sambas, Sanggau, Sintang, Kapuas Hulu, Sekadau, Melawi, dan Kubu Raya dengan produktivitas terendah. Pada 2021, Cluster 0 orange dengan nilai produktivitas tinggi sekitar 40 ton pertahun terdiri dari Landak, Mempawah, Ketapang, Kayong Utara, Kota Pontianak, dan Singkawang. Cluster 1 hijau sedang dengan nilai produktivitas 30 ton pertahun meliputi Bengkayang, Kapuas Hulu, Sekadau, Melawi, dan Kubu Raya, sementara Cluster 2 biru rendah dengan nilai produktivitas 25 ton pertahun berisi Sambas, Sanggau, dan Sintang. Pada 2022, Cluster 2 orange menunjukkan produktivitas tertinggi dengan nilai 35 ton pertahun beranggotakan Singkawang, Kota Pontianak, Ketapang, Kayong utara. Cluster 0 hijau rendah dengan nilai produktivitas 25 ton pertahun berisi Sanggau, Sintang, Kapuas Hulu, Sekadau, dan Melawi, sedangkan Cluster 1 biru sedang dengan nilai produktivitas 30 ton pertahun mencakup Sambas, Bengkayang, Landak, Mempawah, dan Kubu Raya. Pada 2023, Cluster 1 hijau tinggi dengan nilai produktivitas 35 ton pertahun mencakup Bengkayang, Ketapang, Kayong Utara, Kota Pontianak, dan Kota Singkawang. Cluster 0 biru sedang dengan nilai produktivitas 30 ton pertahun terdiri dari Sambas, Landak, Mempawah, Kapuas Hulu, Melawi, dan Kubu Raya, sementara Cluster 2 orange rendah dengan nilai produktivitas sekitar 25 ton pertahun berisi Sanggau, Sintang, dan Sekadau dengan produktivitas terendah.

2. Algoritma Fuzzy C-Means

Pada tahun 2020, Cluster 1 hijau tinggi dengan nilai produktivitas 35-40 ton pertahun mencakup Mempawah, Landak, Kota Pontianak, dan Kota Singkawang. Cluster 0 biru sedang dengan nilai produktivitas 30 ton pertahun terdiri dari Bengkayang, Sambas, Ketapang, Kayong Utara, Kapuas Hulu, dan Kubu Raya, sedangkan Cluster 2 orange rendah dengan nilai produktivitas 20 sampai 25 ton pertahun mencakup Sanggau, Sintang, dan Melawi. Pada tahun 2021, Cluster 1 orange tinggi dengan nilai produktivitas 40 ton pertahun berisi Mempawah, Landak, Kota Pontianak, Kota Singkawang, Kayong Utara, dan Ketapang. Cluster 2 hijau sedang dengan nilai produktivitas 30 ton pertahun mencakup Bengkayang, Kapuas Hulu, Sekadau, Melawi, dan Kubu Raya, sementara Cluster 0 biru rendah dengan nilai produktivitas 20 ton pertahun mencakup Sambas, Sanggau, Sintang. Pada tahun 2022, Cluster 1 orange tinggi dengan nilai produktivitas 40 ton pertahun terdiri dari Kota Singkawang dan Ketapang. Cluster 0 hijau sedang dengan nilai produktivitas 35~30 ton pertahun meliputi Bengkayang, Landak, Mempawah, Kayong Utara, dan Kota Pontianak, sedangkan Cluster 2 biru rendah dengan nilai produktivitas di bawah 30 ton pertahun mencakup Sambas, Sanggau, Sintang, Kapuas Hulu, Sekadau, Melawi, dan Kubu Raya. Pada tahun 2023, Cluster 1 hijau tinggi dengan nilai produktivitas diatas 35 ton pertahun mencakup Bengkayang, Ketapang, Kayong Utara, Kota Pontianak, dan Kota Singkawang. Cluster 2 biru sedang dengan nilai produktivitas di bawah 35 ton pertahun berisi Sambas, Landak, Mempawah, Kapuas Hulu, Melawi, dan Kubu Raya, sedangkan Cluster 0 orange dengan nilai produktivitas 25~20 ton pertahun mencakup Sanggau, Sintang, dan Sekadau.

Tabel 1. Perbandingan Nilai DBI

Tahun	Cluster	DBI K-Means	Cluster	DBI Fuzzy C-Means
2020	0	-0.938	0	-5.061
	1	-2.222	1	-0.188
	2	-5.347	2	-0.222
2021	0	-6.972	0	-1.556
	1	-3.760	1	-6.139
	2	-2.222	2	-2.960
2022	0	-2.485	0	-3.493
	1	-1.124	1	-0.624
	2	-5.024	2	-3.329

2023	0	-1.389	0	-0.889
	1	-2.560	1	-1.760
	2	-1.556	2	-0.556

Dilihat pada tabel 4.25 berikut. Davies-Bouldin Index (DBI) adalah metrik yang digunakan untuk mengevaluasi kualitas duster yang dihasilkan oleh algoritma duster. Nilai DBI yang lebih kecil menunjukkan duster yang lebih baik menurut (Prasetyaningrum & Susanti, 2023) dan (Harahap, 2021). Berdasarkan analisis nilai DBI untuk algoritma K-Means dan Fuzzy C-Means, didapati bahwa kriteria penilaian yang digunakan adalah nilai DBI yang mendekati 0 sebagai yang terbaik. Berdasarkan data tahun 2020 hingga 2023, terlihat bahwa metode Fuzzy C-Means cenderung memiliki nilai yang lebih mendekati nol dibandingkan K-Means di beberapa kluster. Pada tahun 2020, di Cluster 0, K-Means (-0.938) lebih mendekati nol dibandingkan Fuzzy C-Means (-5.061), namun pada Cluster 1 dan Cluster 2, Fuzzy C-Means dengan nilai (-0.188 dan -0.222) lebih unggul dibandingkan K-Means (-2.222 dan -5.347). Tahun 2021 menunjukkan keunggulan Fuzzy C-Means di Cluster 0 (-1.556) dibandingkan K-Means (-6.972). Cluster 2 (-2.960 K-Means unggul dengan nilai -2.222), sedangkan Cluster 1 menunjukkan K-Means lebih baik dengan nilai -3.760 dibandingkan Fuzzy C-Means -6.139. Pada tahun 2022, K-Means unggul pada Cluster 0 (-2.485, -3.493) dan Fuzzy C-Means unggul pada Cluster 1 (-0.624, -1.124), serta unggul pada Cluster 2 dengan nilai -3.329 dibandingkan -5.024. Di tahun 2023, Fuzzy C-Means tetap menunjukkan keunggulannya dengan DBI yang lebih mendekati nol pada Cluster 0 (-0.889 dibandingkan K-Means -1.389), Cluster 1 (-1.760 dibandingkan K-Means -2.560), serta Cluster 2 (-0.556 dibandingkan K-Means -1.556). Secara umum, Fuzzy C-Means menghasilkan nilai DBI yang lebih baik dalam hal mendekati nol, menunjukkan kualitas klusterisasi yang lebih baik dibandingkan K-Means di sebagian besar kluster yang dianalisis. Fuzzy C-Means unggul 8 duster dari 12 kasus duster sedangkan K-Means hanya unggul 4 duster dari kasus 12 duster.

KE SIMPULAN

Berdasarkan analisis menggunakan algoritma K-Means dan Fuzzy C-Means dalam mengelompokkan produktivitas kabupaten/kota dari tahun 2020 hingga 2023 adalah sebagai berikut. Berdasarkan hasil K-Means tahun 2020, warna orange melambangkan duster 0 yang merepresentasikan produktivitas padi tinggi, dimana terdapat 4 daerah, duster 1 dilambangkan dengan warna hijau dengan produktivitas padi sedang, dimana terdapat 3 daerah, duster 2 biru rendah terdapat 7 daerah. Pada K-Means tahun 2021, duster 0 orange nilai produktivitas tinggi terdapat 6 daerah, duster 1 hijau produktivitas sedang terdapat 5 daerah, duster 2 biru rendah terdapat 3 daerah. Pada K-Means tahun 2022, duster 2 orange dengan nilai produktivitas tinggi terdapat 4 daerah, duster 0 hijau produktivitas

rendah terdapat 5 daerah, duster 1 biru sedang terdapat 5 daerah. Pada K-Means 2023, duster 2 orange dengan nilai produktivitas rendah terdapat 3 daerah, duster 1 hijau produktivitas tinggi terdapat 5 daerah, duster 0 biru sedang terdapat 6 daerah. Pada Fuzzy C-Means 2020, duster 2 orange dengan nilai produktivitas rendah terdapat 3 daerah, duster 1 hijau produktivitas tinggi terdapat 4 daerah, duster 0 biru sedang terdapat 7 daerah. Pada Fuzzy C-Means 2021, duster 1 orange dengan nilai produktivitas tinggi terdapat 6 daerah, duster 2 hijau produktivitas sedang terdapat 5 daerah, duster 0 biru rendah terdapat 3 daerah. Pada Fuzzy C-Means 2022, duster 1 orange dengan nilai produktivitas tinggi terdapat 2 daerah, duster 0 hijau produktivitas sedang terdapat 5 daerah, duster 2 biru rendah terdapat 7 daerah. Pada Fuzzy C-Means 2023, duster 0 orange dengan nilai produktivitas rendah terdapat 3 daerah, duster 1 hijau produktivitas tinggi terdapat 5 daerah, duster 2 biru sedang terdapat 6 daerah

Perbandingan antara algoritma K-Means dan Fuzzy C-Means. Secara umum, Fuzzy C-Means menghasilkan nilai DBI yang lebih mendekati nol, menunjukkan kualitas kluster yang lebih baik dibandingkan K-Means. Dari 12 kasus duster, Fuzzy C-Means unggul pada 8 duster, sementara K-Means hanya unggul pada 4 duster. Hal ini mengindikasikan bahwa Fuzzy C-Means lebih mampu membentuk kluster dengan batasan yang lebih jelas dan representatif dalam kebanyakan kasus dibandingkan K-Means.

DAFTAR PUSTAKA

- Aberson, A. M., & Sidebang, S. (2024). Pemetaan Kemiskinan Digital Kabupaten/Kota di Sumatera Utara Menggunakan Ward Hierarchical Clustering Digital Poverty Mapping of Regencies/Cities in North Sumatra Using Ward Hierarchical Clustering. In JOURNAL OF ANALYTICAL RESEARCH (Vol. 3, Issue 2).
- Adi, M., Hutabarat, P., Julham, M., Wanto, A., Studi, P., Informatika, T., Tunas, S., & Pematangsiantar, B. (n.d.). PENERAPAN ALGORITMA BACKPROPAGATION DALAM MEMPREDIKSI PRODUKSI TANAMAN PADI SAWAH MENURUT KABUPATEN/KOTA DI SUMATERA UTARA. 4(1), 77–86.
- Ahmadian, I., Yustiati, A., & Andriani, D. Y. (n.d.). PRODUKTIVITAS BUDIDAYA SISTEM MINA PADI UNTUK MENINGKATKAN KETAHANAN PANGAN DI INDONESIA: A REVIEW. In Jurnal Akuatek (Vol. 2, Issue 1).
- Azizah, N. (2023). Pemodelan Spatial Autoregressive (SAR-X) pada Perkawinan Usia Anak di Indonesia. Jurnal Riset Statistika, 1–10. <https://doi.org/10.29313/jrs.v3i1.1643>
- Butarbutar, N., Perdana Windarto, A., Hartama, D., STIKOM Tunas Bangsa Pematangsiantar, M., & STIKOM Tunas Bangsa Pematangsiantar, D. (2016). KOMPARASI KINERJA ALGORITMA FUZZY C-MEANS DAN K-MEANS DALAM PENGELOMPOKAN DATA SISWA BERDASARKAN PRESTASI NILAI AKADEMIK SISWA (Studi Kasus : SMP Negeri 2 Pematangsiantar). 1(1).

- Dame Tampubolon, H., Safii, M., & Suhendro, D. (2021). Penerapan Algoritma K-Means dan K-Medoids Clustering untuk Mengelompokkan Tindak Kriminalitas Berdasarkan Provinsi. 2(2), 6–12. [Http://creativecommons.org/licenses/by/4.0/](http://creativecommons.org/licenses/by/4.0/)
- Farissa, R. A., Mayasari, R., & Umaidah, Y. (2021). Perbandingan Algoritma K-Means dan K-Medoids Untuk Pengelompokkan Data Obat dengan Silhouette Coefficient. In Journal of Applied Informatics and Computing (JAIC) (Vol. 5, Issue 2). [Http://jurnal.polibatam.ac.id/index.php/JAIC](http://jurnal.polibatam.ac.id/index.php/JAIC)
- Handayani, Y., Syafrudin, S., Suherman, S., ini bertujuan untuk membagi Sungai Pepe Atas Hulu, P., Pepe Bawah Hulu, S., Anyar Hilir, S., Jenes Hilir, S., Kedung Jumbleng Hilir, S., Anyar Tengah, S., Pepe Bawah Hilir, S., Anyar Hulu, S., Gajah Putih Hulu, S., Brojo Hilir, S., Premulung Tengah, S., Premulung Hulu, S., & Sungai Premulung Hilir, dan. (n.d.). Analisis Cluster K-Means Terhadap Kualitas Air Sungai di Kota Surakarta, Jawa Tengah (K-Means Cluster Analysis on River Water Quality in Surakarta City, Central Java).
- Hayati, F. N., Silfiani, M., & Nurlaily, D. (2023). PERBANDINGAN PENGELOMPOKKAN PUSAT KESEHATAN MASYARAKAT DI KOTA BALIKPAPAN MENGGUNAKAN METODE K-MEANS DAN FUZZY C-MEANS. VARIANCE: Journal of Statistics and Its Applications, 5(1), 55–66. <https://doi.org/10.30598/variancevol5iss1page55-66>
- Klau, A. R. D., Suryowati, K., & Statistika, J. (2023). PERBANDINGAN METODE K-MEANS DAN FUZZY C-MEANS TERHADAP INDIKATOR KEMISKINAN NTT 2021 (Studi Kasus: Indikator Tingkat Kemiskinan di NTT tahun 2021). Jurnal Statistika Industri Dan Komputasi, 08(1), 32–37. <https://ntt.bps.go.id>
- Monalisa, S. (2018). Clusterisasi Customer Lifetime Value dengan Model LRFM menggunakan Algoritma K-Means. Jurnal Teknologi Informasi Dan Ilmu Komputer, 5(2), 247–252. <https://doi.org/10.25126/jtiik.201852690>
- Nova Lita, T., Soekartomo, S., Guritno Jurusan Budidaya Pertanian, B., & Pertanian, F. (2013). THE EFFECT OF THE DIFFERENT CROPPING SYSTEMS ON GROWTH AND YIELD OF RICE (*Oryza sativa* .L) IN LOWLAND. 1(4).
- Octavia, S. F., & Mustakim, M. (2021). Penerapan K-Means dan Fuzzy C-Means untuk Pengelompokan Data Kasus Covid-19 di Kabupaten Indragiri Hilir. Building of Informatics, Technology and Science (BITS), 3(2), 88–94. <https://doi.org/10.47065/bits.v3i2.1005>
- Ratna Sari, Y., Sudewa, A., Ayu Lestari, D., & Ika Jaya, T. (2020). PENERAPAN ALGORITMA K-MEANS UNTUK CLUSTERING DATA KEMISKINAN PROVINSI BANTEN MENGGUNAKAN RAPIDMINER (Vol. 5, Issue 2 Georgia: Global Text Project).
- Studi, P., Informasi, S., Kisaran, R., & Yamin, J. M. (2018). PEMANFAATAN RAPID MINER STUDIO 8.2 UNTUK PENGELOMPOKAN DATA PENJUALAN AKSESORIS MENGGUNAKAN ALGORITMA K-MEANS *Mardalius*. IV (2), 123–132.

- Wajong, Andre M.R. 2015. Applying Performance Dashboard in Hospitals. Industrial Engineering, Faculty of Engineering. Bina Nusantara University.
- Wijayanto, S., Yoka Fathoni, M., Panjaitan No, J. DI, Purwokerto Selatan, K., Banyumas, K., & Tengah, J. (n.d.). Pengelompokan Produktivitas Tanaman Padi di Jawa Tengah Menggunakan Metode Clustering K-Means (Vol. 13, Issue 2).