

IMPLEMENTASI DATA MINING PENINGKATAN PRODUKSI BERAS MENGUNAKAN METODE K-MEANS CLUSTERING

¹Intan Febriani✉, ¹M. Safii, ²Ommi Alfina

¹STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

²Universitas Potensi Utama, Medan, Indonesia

Email: intanfebriani0208@gmail.com

DOI: <https://doi.org/10.46880/methoda.Vol12No3.pp258-268>

ABSTRACT

Increasingly high growth rates in the world, especially in developing countries like Indonesia, have reduced the area of agricultural land, thereby reducing food production. Rice is one of the staple food ingredients which is an important commodity, especially in Indonesia because rice is the main consumption ingredient for people to obtain carbohydrate intake. This study discusses the implementation of data mining in increasing rice production in North Sumatra province using the K-Means Clustering algorithm as a solution to solving cases. The source of data in this study was obtained from BPS North Sumatra with 32 processed data. Data analysis in this study used 2 (two) cluster levels, namely the high cluster (C1) and the low cluster (C2). The research results obtained are that there are 3 cities/regencies that are included in the high cluster (C1) and there are 29 other cities/regencies that are included in the low cluster (C2). It is hoped that the research results can become input, suggestions and efforts for the North Sumatra provincial government to pay more attention to increasing rice production in each region so that it can meet the basic needs of the community so that it can increase food security more optimally.

Keyword: Data Mining, K-Means, Clustering, Production, Rice.

ABSTRAK

Tingkat laju pertumbuhan yang semakin tinggi di dunia terutama di negara-negara berkembang seperti Indonesia membuat luas lahan pertanian semakin menyempit sehingga mengurangi produksi bahan pangan. Beras merupakan salah satu bahan makanan pokok yang menjadi komoditas penting terutama di negara Indonesia dikarenakan beras menjadi bahan konsumsi utama bagi masyarakat untuk memperoleh asupan karbohidrat. Penelitian ini membahas tentang implementasi data mining dalam meningkatkan produksi beras di provinsi Sumatera Utara menggunakan algoritma K-Means Clustering sebagai solusi penyelesaian kasus. Sumber data pada penelitian ini diperoleh dari BPS Sumatera Utara dengan data olah sebanyak 32 data. Analisa data pada penelitian ini menggunakan 2 (dua) tingkat cluster yaitu cluster tinggi (C1) dan cluster rendah (C2). Hasil penelitian yang diperoleh yaitu terdapat 3 daerah kota/kabupaten yang masuk ke dalam cluster tinggi (C1) dan terdapat 29 kota/kabupaten lainnya yang termasuk ke dalam cluster

rendah (C2). Diharapkan hasil penelitian dapat menjadi masukan, saran serta upaya pada pemerintah provinsi Sumatera Utara untuk lebih memperhatikan peningkatan produksi beras di setiap daerah agar dapat memenuhi kebutuhan pokok masyarakat sehingga dapat meningkatkan ketahanan pangan yang lebih optimal.

Kata Kunci: Data Mining, K-Means, Clustering, Produksi, Beras.

PENDAHULUAN

Tingkat laju pertumbuhan yang semakin tinggi di dunia terutama di negara-negara berkembang seperti Indonesia membuat luas lahan pertanian semakin menyempit sehingga mengurangi produksi bahan pangan. Ketahanan pangan sangat bergantung pada tingkat panen maupun produksi bahan makanan terutama bahan pokok. Kekurangan ketersediaan lahan pertanian tidak hanya membuat produksi pangan menurun tetapi juga akan menyebabkan harga-harga pangan semakin melonjak naik sehingga dapat mempersulit perekonomian masyarakat (Saputra & Purnomo, 2022). Tolak ukur lainnya untuk mengetahui berhasil atau tidaknya suatu wilayah dalam mencapai tingkat ketahanan pangan yaitu dengan melihat akses pangan dan penyerapan pangan pada wilayah tersebut. Wilayah yang memiliki akses pangan yang baik dapat menunjukkan bahwa setiap orang akan memiliki akses fisik dan ekonomi yang layak untuk kebutuhan konsumsi sehari-hari untuk memperoleh kondisi hidup yang sehat dan produktif (Destiarsono & Darwanto, 2022).

Beras merupakan salah satu bahan makanan pokok yang menjadi komoditas penting terutama di negara Indonesia dikarenakan beras menjadi bahan konsumsi utama bagi masyarakat untuk memperoleh asupan karbohidrat. Kestabilan produksi beras sangat penting menjadi perhatian pemerintah dikarenakan selain sebagai konsumsi utama, beras juga merupakan salah satu sumber devisa juga menyerap

banyak tenaga kerja, khususnya untuk jutaan rumah tangga petani di wilayah pedesaan (Mardiyah, Fajar, & Badruzzaman, 2022).

Semakin banyaknya impor beras ke Indonesia juga menunjukkan bahwa perekonomian Indonesia menurun terhadap ketersediaan bahan pangan. Hal tersebut dapat disimpulkan bahwa ketahanan bahan pangan di Indonesia semakin melemah dikarenakan luas lahan pertanian yang semakin sempit dan produksi bahan makanan yang semakin sedikit tidak dapat memenuhi kebutuhan penduduk yang semakin padat (Zaeroni & Rustariyuni, 2016).

Berdasarkan data dari Badan Pusat Statistik (BPS) provinsi Sumatera Utara dijelaskan bahwa produksi beras di provinsi Sumatera Utara semakin menurun di setiap tahunnya sesuai dengan wilayah produksinya. Permintaan kebutuhan beras lebih banyak di bandingkan dengan jumlah produksinya sehingga masyarakat masih banyak mengirim permintaan beras dari provinsi lain seperti provinsi Jawa Tengah.

Untuk mengetahui daerah-daerah di provinsi Sumatera Utara yang termasuk ke dalam kategori produksi beras yang rendah maupun tinggi perlu dilakukan analisis data berdasarkan data yang telah ada sebelumnya dengan menggunakan algoritma *K-Means Clustering* sebagai solusi dalam menyelesaikan kasus secara kompleks (Reyhan). Hal tersebut bertujuan untuk meminimalisir kekurangan kebutuhan beras pada daerah tertentu

dengan mempersiapkan ketersediaan sesuai dengan perkiraan yang cukup (Antari, Harini, & Tastrawati, 2021).

Algoritma *K-Means Clustering* telah banyak digunakan secara praktis dan terstruktur di berbagai penelitian terdahulu dan diperoleh hasil penelitian yang efektif dan juga efisien (Sudirman, Windarto, & Wanto, 2018)(Dewi, Purnama, & Utami, 2022).

Berdasarkan penelitian terdahulu yang dilakukan oleh (agung) diperoleh hasil penelitian bahwa algoritma *K-Means Clustering* dapat menjadi rekomendasi yang baik dalam analisa kasus penjualan sehingga diperoleh hasil *performance* data yang lebih akurat. Penelitian terdahulu tentang kasus penentuan judul skripsi juga menggunakan algoritma *K-Means Clustering* sebagai alternatif penyelesaian kasus dengan hasil penelitian bahwa mahasiswa lebih mudah untuk mengetahui ide judul skripsi dengan cepat dan tepat sesuai dengan kemampuan analisa masing-masing (Sembiring, Hanum, & Tamba, 2022).

Menurut penelitian yang dilakukan oleh (Normah, Nurajizah, & Salbinda, 2021) juga menggunakan algoritma *K-Means Clustering* dalam menyelesaikan kasus

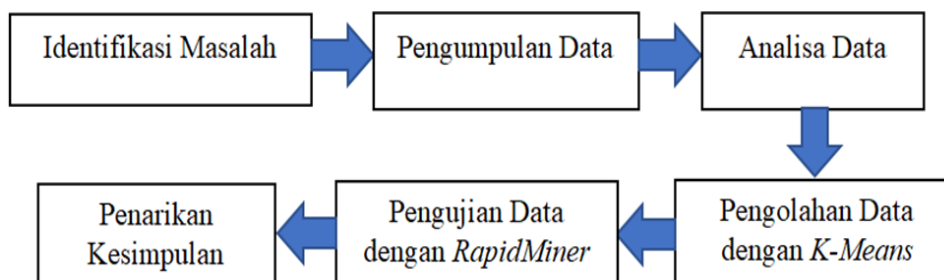
analisa penjualan hijab. Dari hasil penelitian diperoleh bahwa dengan penggunaan algoritma *K-Means Clustering* dapat mempermudah distributor dalam memilah barang laris dan tidak laris sehingga dapat menyediakan stok dengan tepat sasaran sehingga proses penjualan dapat berjalan dengan optimal.

Berdasarkan pemaparan latar belakang masalah, diharapkan penelitian ini dapat mengimplementasikan data mining algoritma *K-Means Clustering* dalam meningkatkan produksi beras di provinsi Sumatera Utara. Dari hasil penelitian yang diperoleh, diharapkan dapat menjadi masukan, saran, dan upaya bagi pemerintah khususnya di provinsi Sumatera Utara untuk selalu memperhatikan daerah-daerah yang kekurangan akan produksi beras serta dapat mengoptimalkan akses lahan pertanian sehingga dapat memenuhi kebutuhan konsumsi masyarakat dalam meningkatkan ketahanan pangan yang realistis.

METODOLOGI PENELITIAN

Tahapan Penelitian

Kerangka kerja pada penelitian ini dapat diuraikan seperti berikut:



Gambar 1. Tahapan Penelitian

Berdasarkan pada Gambar 1., dapat dijelaskan bahwa proses penelitian dilakukan dengan berbagai tahap yaitu dimulai dengan melakukan identifikasi data dengan mempelajari mengenai pengaruh produksi beras terhadap ketahanan pangan dan penggunaan algoritma *K-Means Clustering* dari berbagai penelitian terdahulu seperti jurnal, buku, dan karya ilmiah lainnya. Selanjutnya melakukan proses pengumpulan data yang dalam hal ini dilakukan secara online dengan mengakses website BPS Sumatera Utara. Berikutnya data yang dikumpulkan kemudian dianalisis sesuai dengan kebutuhan dan referensi yang telah ditentukan. Kemudian data yang telah diperoleh diolah dengan menggunakan algoritma *K-Means Clustering* sesuai dengan alur kerjanya di aplikasi *Microsoft Excel*. Selanjutnya hasil pengolahan data dengan algoritma *K-Means* di uji kembali pada aplikasi *RapidMiner 5.3* untuk mengetahui akurasi kebenaran pengolahan data. Terakhir, hasil pengujian data yang diperoleh di evaluasi kembali sehingga dapat ditarik suatu kesimpulan pada penelitian.

Data Mining

Data mining merupakan salah satu ilmu bidang komputer yang telag banyak digunakan untuk menyelesaikan berbagai kasus penelitian. Terdapat berbagai macam algoritma dalam data mining seperti *K-Means*, *K-Medoids*, *C4.5*, *Naïve Bayes*, *Regresi Linier* dan lainnya. Adapun istilah yang ada pada padanan kata data mining yaitu seperti *pattern recognition* yang maksudnya yaitu data mining dapat melakukan pengenalan pola data yang bersifat tersembunyi pada suatu bongkahan data. Adapula istilah *knowledge discovery* yang maksudnya data mining dapat dilakukan penemuan pengetahuan pada data tersembunyi yang terdapat di dalam

bongkahan data dikarenakan pada realitanya data mining digunakan untuk memperoleh pengetahuan secara terperinci (Zulfa, Auliya, & Zaenal, 2021).

Secara luas, pengertian data mining adalah suatu proses yang digunakan untuk melakukan identifikasi atau pengenalan data yang ber-potensi tinggi dan dapat digunakan dengan baik dalam proses pengelolaan dataset yang besar (Miharja & Suhendri, 2021).

K-Means Clustering

Algoritma *K-Means* merupakan salah satu algoritma *clustering* yang sangat umum digunakan dalam pengelompokan data yang sesuai dengan karakteristik yang serupa. Adapun kelompok data yang dihasilkan disebut sebagai *cluster*/klaster yang dirancang untuk mengklasifikasikan objek berdasarkan karakteristik yang sama. *K-Means* juga merupakan algoritma *clustering* yang hanya dapat bekerja pada tiap atribut yang bersifat angka atau *numeric* dikarenakan memiliki basis jarak yang dapat membagi data ke dalam sejumlah *cluster* tertentu (Atros, Robi Padri, Nurdiawan, Faqih, & Anwar, 2021).

Langkah-langkah melakukan klasterisasi data dengan menggunakan algoritma *K-Means* yaitu sebagai berikut:

- a. Menentukan nilai jumlah *cluster* (K)
- b. Melakukan perhitungan pada iterasi pertama dengan menetapkan nilai *centroid* awal yang ditentukan secara acak atau *random*. Selanjutnya pada iterasi selanjutnya, penentuan nilai *centroid* yang baru dilakukan sesuai dengan rumus dibawah ini:

$$V_{ij} = \frac{1}{n} \sum_{k=1}^n x_{ik} \quad (1)$$

Ket.:

v_{ij} = data *cluster* ke-i kolom j

W_{kj} = data ke-k kolom ke- j

X_i = Banyaknya anggota *cluster* ke- i

c. Mencari jarak antara titik pada pusat *cluster* dengan menghitung *centroid* dengan titik setiap objeknya. Hasil yang diperoleh adalah nilai jarak terpendek dari setiap titik *cluster* dengan menggunakan rumus *Euclidean Distance* sebagai berikut :

$$P1_{(x_2-x_1)} = j = 1m(x_2 - x_1)^2 \quad (2)$$

Ket.:

$P_1 (X_1-X_2)$ = jarak objek antara objek i dan k

M = dimensi data

X_{2j} = koordinat dari obyek i pada dimensi k

X_{ij} = koordinat obyek j pada dimensi k

d. Melakukan proses pengelompokan data dari setiap objek dengan cara mencari jarak minimum pada objek yang serupa untuk menentukan bagian tiap *cluster*.

e. Lakukan proses perhitungan secara berulang sampai diperoleh nilai *centroid* awal dengan nilai *centroid* sebelumnya bernilai sama. Atau dapat dilihat dari

hasil pengelompokan *cluster* pada iterasi pertama dengan iterasi selanjutnya. Apabila nilai yang disebutkan bernilai sama maka proses perhitungan berhenti dan langkah kerja *K-Means Clustering* selesai.

HASIL DAN PEMBAHASAN

Berdasarkan pemaparan sebelumnya, proses penelitian akan melalui beberapa tahapan sesuai dengan alur penelitian dan juga langkah kerja algoritma *K-Means Clustering*.

Analisa Data

Pada tahap proses analisa data dilakukan dengan cara mencari dan menetapkan data yang akan diolah yang sesuai dengan referensi dan kasus penelitian. Adapun data yang digunakan pada penelitian ini yaitu data jumlah produksi beras di provinsi Sumatera Utara dari tahun 2019-2022. Data mentah yang akan diolah pada penelitian ini dapat dilihat pada Tabel 1.

Tabel 1. Data Jumlah Produksi Beras Provinsi Sumatera Utara

| No. | Kabupaten/ Kota | Produksi Beras (ton) Provinsi Sumatera Utara (Ton) | | | |
|-----|--------------------|--|-----------|-----------|-----------|
| | | 2019 | 2020 | 2021 | 2022 |
| 1 | Nias | 18075.18 | 19472.43 | 20194.5 | 20858.52 |
| 2 | Mandailing Natal | 47170.28 | 43272.28 | 42437.74 | 43943.88 |
| 3 | Tapanuli Selatan | 55987.25 | 54977.77 | 51848.99 | 54511.87 |
| 4 | Tapanuli Tengah | 29719.7 | 24374.33 | 23928.4 | 29904.97 |
| 5 | Tapanuli Utara | 63795.2 | 66494.84 | 62913.44 | 74252.63 |
| 6 | Toba Samosir | 73737.34 | 62677.78 | 60586.14 | 62946.74 |
| 7 | Labuhan Batu | 43155.3 | 34770.60 | 33208.82 | 33654.58 |
| 8 | Asahan | 35103.33 | 36854.41 | 35010.19 | 31926.01 |
| 9 | Simalungun | 95610.69 | 97418.15 | 99754.03 | 103516.35 |
| 10 | Dairi | 15975.76 | 22115.61 | 20150.9 | 19950.98 |
| 11 | Karo | 33308.76 | 33493.45 | 33007.88 | 41099.53 |
| 12 | Deli Serdang | 177352.74 | 177547.89 | 179847.64 | 186953.05 |
| 13 | Langkat | 73212.2 | 77618.70 | 79795.29 | 72478.85 |
| 14 | Nias Selatan | 33964.81 | 25123.52 | 26365.98 | 35187.7 |
| 15 | Humbahas | 38969.39 | 33736.79 | 32179.44 | 31365.43 |
| 16 | Pakpak Bharat | 3073.55 | 2558.74 | 2125.35 | 2873.86 |
| 17 | Samosir | 27174.65 | 22401.22 | 21173.46 | 22971.31 |
| 18 | Serdang Bedagai | 159617.25 | 174556.12 | 169684.38 | 153282 |
| 19 | Batu Bara | 37558.74 | 43164.91 | 42194.07 | 41644.3 |

| | | | | | |
|----|---------------------|----------|----------|----------|----------|
| 20 | Padang Lawas Utara | 16697.06 | 20439.04 | 19683.16 | 17109.81 |
| 21 | Padang Lawas | 13595.73 | 16522.38 | 16672.67 | 15240.53 |
| 22 | Labuhanbatu Selatan | 147.50 | 377.69 | 356.32 | 237.36 |
| 23 | Labuanbatu Utara | 27267.70 | 31754.59 | 32028.68 | 21942.67 |
| 24 | Nias Utara | 27877.67 | 22325.73 | 18105.25 | 23530.26 |
| 25 | Nias Barat | 4951.03 | 6384.66 | 6020.31 | 8105.13 |
| 26 | Tanjungbalai | 244.69 | 217.13 | 243.61 | 231.76 |
| 27 | Pematangsiantar | 6456.21 | 7397.25 | 7221.72 | 6847.53 |
| 28 | Tebing Tinggi | 1782.37 | 1505.61 | 1526.94 | 1338.96 |
| 29 | Medan | 3440.34 | 2807.37 | 2868.31 | 3102.44 |
| 30 | Binjai | 3412.98 | 3970.10 | 4490.89 | 4248.81 |
| 31 | Padangsidempuan | 11089.27 | 11452.45 | 11190.64 | 10229.92 |
| 32 | Gunungsitoli | 6824.29 | 7069.45 | 7619.72 | 8552.52 |

Berdasarkan pada Tabel 1., dapat dijelaskan bahwa data yang digunakan pada penelitian ini yaitu data jumlah produksi beras di 32 daerah di provinsi Sumatera Utara dari tahun 2019-2022. Apabila data mentah penelitian telah selesai di analisa, selanjutnya melakukan proses akumulasi data ke bagian data yang lebih singkat yang bertujuan untuk mempermudah proses perhitungan dan pengujian data selanjutnya. Proses perhitungan akumulasi data ditentukan dengan menghitung nilai rata-rata (*mean*) dari setiap data per daerah di setiap tahunnya. Adapun hasil perhitungan akumulasi data mentah pada penelitian ini dapat dilihat pada Tabel 2.

| | | |
|----|---------------------|-----------|
| 14 | Nias Selatan | 31839.50 |
| 15 | Humbang Hasundutan | 31772.44 |
| 16 | Pakpak Bharat | 2499.61 |
| 17 | Samosir | 22072.39 |
| 18 | Serdang Bedagai | 161483.19 |
| 19 | Batu Bara | 41919.19 |
| 20 | Padang Lawas Utara | 18396.49 |
| 21 | Padang Lawas | 15956.60 |
| 22 | Labuhanbatu Selatan | 296.84 |
| 23 | Labuanbatu Utara | 26985.68 |
| 24 | Nias Utara | 20817.76 |
| 25 | Nias Barat | 7062.72 |
| 26 | Tanjungbalai | 237.69 |
| 27 | Pematangsiantar | 7034.63 |
| 28 | Tebing Tinggi | 1432.95 |
| 29 | Medan | 2985.38 |
| 30 | Binjai | 4369.85 |
| 31 | Padangsidempuan | 10710.28 |
| 32 | Gunungsitoli | 8086.12 |

Tabel 2. Hasil Akumulasi Data Mentah Penelitian

| No. | Kabupaten/ Kota | Rata-Rata (ton) |
|-----|------------------|-----------------|
| 1 | Nias | 19709.40 |
| 2 | Mandailing Natal | 44517.30 |
| 3 | Tapanuli Selatan | 54116.04 |
| 4 | Tapanuli Tengah | 27851.02 |
| 5 | Tapanuli Utara | 66987.09 |
| 6 | Toba Samosir | 65756.74 |
| 7 | Labuhan Batu | 36672.90 |
| 8 | Asahan | 34013.18 |
| 9 | Simalungun | 99627.02 |
| 10 | Dairi | 18692.55 |
| 11 | Karo | 35805.39 |
| 12 | Deli Serdang | 181384.48 |
| 13 | Langkat | 75162.11 |

Proses Algoritma K-Means Clustering Perhitungan Iterasi I (Pertama)

Sesuai pada pemaparan sebelumnya mengenai langkah kerja algoritma *K-Means Clustering*, langkah awal yang dilakukan yaitu menentukan nilai *centroid* awal yang dapat ditentukan secara acak. Pada penelitian ini, nilai *centroid* awal di ambil dari nilai akumulasi data yaitu nilai rata-rata tertinggi sebagai *cluster* 1 (C1) dan nilai terendah sebagai *cluster* 2 (C2). Nilai *centroid* awal dapat dilihat pada Tabel 3.

Tabel 3. Nilai *Centriod* Awal

| <i>Cluster</i> | Nilai |
|----------------------|--------|
| C1 = Nilai Tertinggi | 181384 |
| C2 = Nilai Terendah | 238 |

perhitungan untuk mencari nilai titik pusat cluster sesuai dengan rumus pada persamaan 2 (dua) yang telah dijelaskan sebelumnya. Hasil nilai jarak titik pusat pada setiap cluster dapat dilihat pada Tabel 4.

Setelah diperoleh nilai centroid awal, maka langkah selanjutnya melakukan

Tabel 4. Hasil Perhitungan Jarak *Cluster* Pusat Pada Iterasi 1

| No. | Kabupaten/Kota | Rata-Rata | Clustering | | Jarak Terpendek | Hasil |
|-----|---------------------|-----------|------------|--------|-----------------|-------|
| | | | C1 | C2 | | |
| 1 | Nias | 19709.40 | 161675 | 19472 | 19472 | C2 |
| 2 | Mandailing Natal | 44517.30 | 136867 | 44280 | 44280 | C2 |
| 3 | Tapanuli Selatan | 54116.04 | 127268 | 53878 | 53878 | C2 |
| 4 | Tapanuli Tengah | 27851.02 | 153533 | 27613 | 27613 | C2 |
| 5 | Tapanuli Utara | 66987.09 | 114397 | 66749 | 66749 | C2 |
| 6 | Toba Samosir | 65756.74 | 115628 | 65519 | 65519 | C2 |
| 7 | Labuhan Batu | 36672.90 | 144712 | 36435 | 36435 | C2 |
| 8 | Asahan | 34013.18 | 147371 | 33775 | 33775 | C2 |
| 9 | Simalungun | 99627.02 | 81757 | 99389 | 81757 | C1 |
| 10 | Dairi | 18692.55 | 162692 | 18455 | 18455 | C2 |
| 11 | Karo | 35805.39 | 145579 | 35568 | 35568 | C2 |
| 12 | Deli Serdang | 181384.48 | 0 | 181147 | 0 | C1 |
| 13 | Langkat | 75162.11 | 106222 | 74924 | 74924 | C2 |
| 14 | Nias Selatan | 31839.50 | 149545 | 31602 | 31602 | C2 |
| 15 | Humbang Hasundutan | 31772.44 | 149612 | 31535 | 31535 | C2 |
| 16 | Pakpak Bharat | 2499.61 | 178885 | 2262 | 2262 | C2 |
| 17 | Samosir | 22072.39 | 159312 | 21835 | 21835 | C2 |
| 18 | Serdang Bedagai | 161483.19 | 19901 | 161246 | 19901 | C1 |
| 19 | Batu Bara | 41919.19 | 139465 | 41682 | 41682 | C2 |
| 20 | Padang Lawas Utara | 18396.49 | 162988 | 18159 | 18159 | C2 |
| 21 | Padang Lawas | 15956.60 | 165428 | 15719 | 15719 | C2 |
| 22 | Labuhanbatu Selatan | 296.84 | 181088 | 59 | 59 | C2 |
| 23 | Labuanbatu Utara | 26985.68 | 154399 | 26748 | 26748 | C2 |
| 24 | Nias Utara | 20817.76 | 160567 | 20580 | 20580 | C2 |
| 25 | Nias Barat | 7062.72 | 174322 | 6825 | 6825 | C2 |
| 26 | Tanjungbalai | 237.69 | 181147 | 0 | 0 | C2 |
| 27 | Pematangsiantar | 7034.63 | 174350 | 6797 | 6797 | C2 |
| 28 | Tebing Tinggi | 1432.95 | 179952 | 1195 | 1195 | C2 |
| 29 | Medan | 2985.38 | 178399 | 2748 | 2748 | C2 |
| 30 | Binjai | 4369.85 | 177015 | 4132 | 4132 | C2 |
| 31 | Padangsidempuan | 10710.28 | 170674 | 10473 | 10473 | C2 |
| 32 | Gunungsitoli | 8086.12 | 173298 | 7848 | 7848 | C2 |

Selanjutnya dilakukan proses pengelompokan data untuk mempermudah dalam mengetahui obyek data yang tergolong pada kelompok atau *cluster* yang sama. Hasil pemilahan data tersebut dapat dilihat pada Tabel 5.

Tabel 5. Hasil Pengelompokan Data Pada Iterasi 1

| No. | Kabupaten/Kota | Clustering | |
|-----|------------------|------------|----|
| | | C1 | C2 |
| 1 | Nias | | 1 |
| 2 | Mandailing Natal | | 1 |
| 3 | Tapanuli Selatan | | 1 |
| 4 | Tapanuli Tengah | | 1 |
| 5 | Tapanuli Utara | | 1 |
| 6 | Toba Samosir | | 1 |
| 7 | Labuhan Batu | | 1 |
| 8 | Asahan | | 1 |
| 9 | Simalungun | 1 | |

| | | | |
|---------------|---------------------|----------|-----------|
| 10 | Dairi | | 1 |
| 11 | Karo | | 1 |
| 12 | Deli Serdang | 1 | |
| 13 | Langkat | | 1 |
| 14 | Nias Selatan | | 1 |
| 15 | Humbang Hasundutan | | 1 |
| 16 | Pakpak Bharat | | 1 |
| 17 | Samosir | | 1 |
| 18 | Serdang Bedagai | 1 | |
| 19 | Batu Bara | | 1 |
| 20 | Padang Lawas Utara | | 1 |
| 21 | Padang Lawas | | 1 |
| 22 | Labuhanbatu Selatan | | 1 |
| 23 | Labuhanbatu Utara | | 1 |
| 24 | Nias Utara | | 1 |
| 25 | Nias Barat | | 1 |
| 26 | Tanjungbalai | | 1 |
| 27 | Pematangsiantar | | 1 |
| 28 | Tebing Tinggi | | 1 |
| 29 | Medan | | 1 |
| 30 | Binjai | | 1 |
| 31 | Padangsidempuan | | 1 |
| 32 | Gunungsitoli | | 1 |
| Jumlah | | 3 | 29 |

Berdasarkan pada Tabel 5., dapat diketahui bahwa terdapat 3 kota/kab di provinsi Sumatera Utara yang termasuk pada kelompok *cluster* tinggi (C1) sedangkan

kelompok pada *cluster* rendah (C2) terdapat pada 29 kota/kab lainnya.

Perhitungan Iterasi 2 (Kedua)

Sesuai dengan langkah-langkah kerja metode *K-Means Clustering* yang telah dijelaskan pada bagian sebelumnya, maka proses perhitungan diulang kembali pada iterasi selanjutnya yaitu iterasi kedua yang dimulai kembali dari perhitungan nilai *centroid* awal. Adapun hasil *centroid* kedua pada iterasi kedua dapat dilihat pada Tabel 6.

Tabel 6. Hasil Nilai *Centroid* Kedua

| <i>Cluster</i> | Nilai |
|----------------------|--------|
| C1 = Nilai Tertinggi | 147498 |
| C2 = Nilai Terendah | 25302 |

Hasil perhitungan jarak titik antar *cluster* pusat pada iterasi kedua dapat dilihat pada Tabel 7.

Tabel 7. Hasil Perhitungan Jarak *Cluster* Pusat Pada Iterasi 2

| No. | Kabupaten/Kota | Rata-Rata | Clustering | | Jarak Terpendek | Hasil |
|-----|---------------------|-----------|------------|--------|-----------------|-------|
| | | | C1 | C2 | | |
| 1 | Nias | 19709.40 | 127789 | 5593 | 5593 | C2 |
| 2 | Mandailing Natal | 44517.30 | 102981 | 19215 | 19215 | C2 |
| 3 | Tapanuli Selatan | 54116.04 | 93382 | 28814 | 28814 | C2 |
| 4 | Tapanuli Tengah | 27851.02 | 119647 | 2549 | 2549 | C2 |
| 5 | Tapanuli Utara | 66987.09 | 80511 | 41685 | 41685 | C2 |
| 6 | Toba Samosir | 65756.74 | 81741 | 40455 | 40455 | C2 |
| 7 | Labuhan Batu | 36672.90 | 110825 | 11371 | 11371 | C2 |
| 8 | Asahan | 34013.18 | 113485 | 8711 | 8711 | C2 |
| 9 | Simalungun | 99627.02 | 47871 | 74325 | 47871 | C1 |
| 10 | Dairi | 18692.55 | 128806 | 6610 | 6610 | C2 |
| 11 | Karo | 35805.39 | 111693 | 10503 | 10503 | C2 |
| 12 | Deli Serdang | 181384.48 | 33886 | 156082 | 33886 | C1 |
| 13 | Langkat | 75162.11 | 72336 | 49860 | 49860 | C2 |
| 14 | Nias Selatan | 31839.50 | 115659 | 6537 | 6537 | C2 |
| 15 | Humbang Hasundutan | 31772.44 | 115726 | 6470 | 6470 | C2 |
| 16 | Pakpak Bharat | 2499.61 | 144999 | 22802 | 22802 | C2 |
| 17 | Samosir | 22072.39 | 125426 | 3230 | 3230 | C2 |
| 18 | Serdang Bedagai | 161483.19 | 13985 | 136181 | 13985 | C1 |
| 19 | Batu Bara | 41919.19 | 105579 | 16617 | 16617 | C2 |
| 20 | Padang Lawas Utara | 18396.49 | 129102 | 6906 | 6906 | C2 |
| 21 | Padang Lawas | 15956.60 | 131542 | 9345 | 9345 | C2 |
| 22 | Labuhanbatu Selatan | 296.84 | 147201 | 25005 | 25005 | C2 |
| 23 | Labuhanbatu Utara | 26985.68 | 120513 | 1684 | 1684 | C2 |
| 24 | Nias Utara | 20817.76 | 126680 | 4484 | 4484 | C2 |
| 25 | Nias Barat | 7062.72 | 140436 | 18239 | 18239 | C2 |

| | | | | | | |
|----|-----------------|----------|--------|-------|-------|----|
| 26 | Tanjungbalai | 237.69 | 147261 | 25064 | 25064 | C2 |
| 27 | Pematangsiantar | 7034.63 | 140464 | 18267 | 18267 | C2 |
| 28 | Tebing Tinggi | 1432.95 | 146065 | 23869 | 23869 | C2 |
| 29 | Medan | 2985.38 | 144513 | 22317 | 22317 | C2 |
| 30 | Binjai | 4369.85 | 143128 | 20932 | 20932 | C2 |
| 31 | Padangsidempuan | 10710.28 | 136788 | 14592 | 14592 | C2 |
| 32 | Gunungsitoli | 8086.12 | 139412 | 17216 | 17216 | C2 |

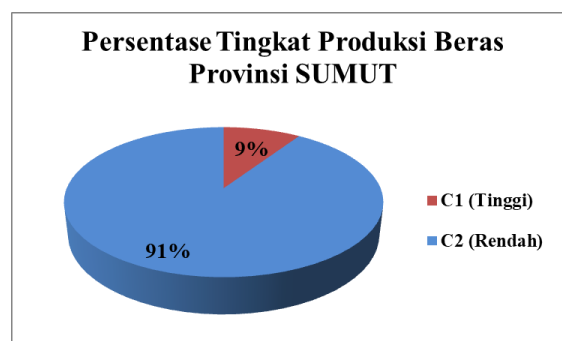
Selanjutnya hasil pengelompokan data *cluster* pertama dan *cluster* kedua pada iterasi kedua dapat dilihat pada Tabel 8.

Tabel 8. Hasil Pengelompokan Data Pada Iterasi 2

| No. | Kabupaten/Kota | Clustering | |
|---------------|---------------------|------------|----|
| | | C1 | C2 |
| 1 | Nias | | 1 |
| 2 | Mandailing Natal | | 1 |
| 3 | Tapanuli Selatan | | 1 |
| 4 | Tapanuli Tengah | | 1 |
| 5 | Tapanuli Utara | | 1 |
| 6 | Toba Samosir | | 1 |
| 7 | Labuhan Batu | | 1 |
| 8 | Asahan | | 1 |
| 9 | Simalungun | 1 | |
| 10 | Dairi | | 1 |
| 11 | Karo | | 1 |
| 12 | Deli Serdang | 1 | |
| 13 | Langkat | | 1 |
| 14 | Nias Selatan | | 1 |
| 15 | Humbang Hasundutan | | 1 |
| 16 | Pakpak Bharat | | 1 |
| 17 | Samosir | | 1 |
| 18 | Serdang Bedagai | 1 | |
| 19 | Batu Bara | | 1 |
| 20 | Padang Lawas Utara | | 1 |
| 21 | Padang Lawas | | 1 |
| 22 | Labuhanbatu Selatan | | 1 |
| 23 | Labuanbatu Utara | | 1 |
| 24 | Nias Utara | | 1 |
| 25 | Nias Barat | | 1 |
| 26 | Tanjungbalai | | 1 |
| 27 | Pematangsiantar | | 1 |
| 28 | Tebing Tinggi | | 1 |
| 29 | Medan | | 1 |
| 30 | Binjai | | 1 |
| 31 | Padangsidempuan | | 1 |
| 32 | Gunungsitoli | | 1 |
| Jumlah | | 3 | 29 |

Berdasarkan pada hasil pengelompokan data pada Tabel 8, maka dapat diketahui bahwa hasil pengelompokan data C1 dan C2 pada iterasi pertama dan iterasi kedua adalah

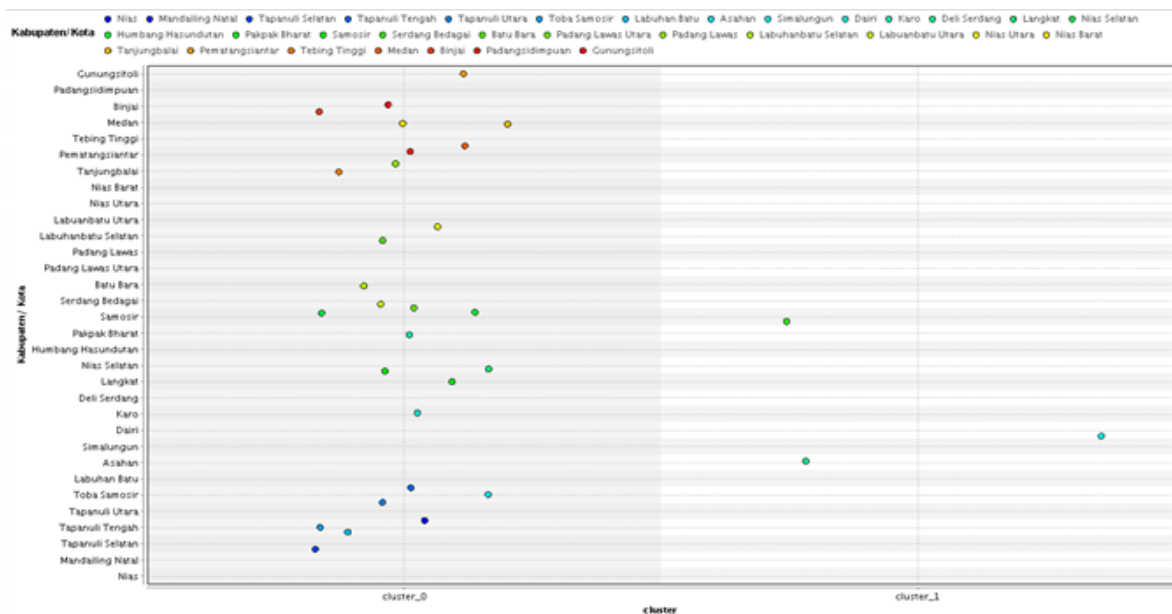
sama atau serupa. Adapun kota/kab yang masuk ke dalam kelompok *cluster* tinggi yaitu sebanyak 3 kota/kab sedangkan 29 kota/kab lainnya termasuk ke dalam *cluster* rendah. Hal tersebut berarti, apabila nilai pengelompokan data pada tiap *cluster* pada iterasi awal dan akhir sama, maka proses perhitungan akan berhenti dan komputasi algoritma *K-Means Clustering* selesai dan berakhir pada iterasi kedua. Untuk mengetahui peningkatan persentase produksi beras di provinsi Sumatera Utara baik tingkat tinggi maupun rendah, maka dapat dilihat pada Gambar 2.



Gambar 2. Persentase *Cluster*

Pengujian Data

Pengujian data pada penelitian ini dilakukan dengan bantuan aplikasi *RapidMiner 5.3*. pengujian data bertujuan untuk mengetahui akurasi ataupun kebenaran pada perhitungan manual. Hasil pengujian data dengan menggunakan *software RapidMiner 5.3* dapat dilihat pada Gambar 3.



Gambar 3. Hasil Pengujian Data Pada *Tool RapidMiner 5.3*

Berdasarkan pada Gambar 3., dapat dijelaskan bahwa pada *tool RapidMiner 5.3* hasil *cluster 0* menunjukkan *cluster* rendah, sedangkan *cluster 1* menunjukkan *cluster* tinggi. Sesuai penjelasan sebelumnya dapat dilihat bahwa hasil perhitungan manual dan pengujian data bernilai sama atau serupa. Adapun kota/kabupaten yang termasuk pada tingkat produksi tinggi hanya terdapat pada 3 daerah yaitu Simalungun, Deli Serdang, dan Serdang Bedagai, sedangkan 29 kota/kab lainnya termasuk pada tingkat produksi rendah. Hal tersebut berarti bahwa tingkat produksi beras di provinsi Sumatera Utara masih tergolong rendah.

KESIMPULAN

Berdasarkan hasil dan pembahasan sebelumnya, maka dapat disimpulkan bahwa metode *K-Means Clustering* dapat diterapkan dalam implementasi peningkatan produksi beras di Provinsi Sumatera Utara dengan menggunakan data produksi beras pada tahun 2019-2022. Hasil perhitungan dengan menggunakan dua tingkat *cluster* diperoleh hasil jumlah *cluster 1* (C1) sebanyak 3 anggota dan *cluster 2* (C2) sebanyak 29 anggota dari 32 total nama

kota/kabupaten. Hal tersebut berarti terdapat 3 kota/kabupaten yang memiliki peningkatan produksi beras yang optimal yang harus lebih diperhatikan dan ditingkatkan lagi proses produksinya pada periode berikutnya, sedangkan pada 29 kota/kabupaten lainnya masih minim produksi berasnya sehingga sangat perlu dilakukan peninjauan lebih lanjut oleh pemerintah setempat di provinsi Sumatera Utara untuk meningkatkan produksi beras dengan menyediakan luas lahan yang cukup dan menyediakan akses pertanian yang produktif sehingga dapat memenuhi ketahanan pangan untuk menunjang pertumbuhan ekonomi yang optimal.

DAFTAR PUSTAKA

- Antari, N. K. D. P., Harini, L. P. I., & Tastrawati, N. K. T. (2021). Analisis Penjadwalan Produksi Menggunakan Metode Campbell Dudek Smith Dan Dannenbring Dalam Meminimumkan Total Waktu Produksi Beras. *E-Jurnal Matematika*, 10(4), 215–221. <https://doi.org/10.24843/mtk.2021.v10.i04.p345>
- Atros, K. S. H. K. Al, Robi Padri, A., Nurdiawan, O., Faqih, A., & Anwar,

- S. (2021). Model Klasifikasi Analisis Kepuasan Pengguna Perpustakaan Online Menggunakan K-Means dan Decision Tree. *JURIKOM (Jurnal Riset Komputer)*, 8(6), 323–329. <https://doi.org/10.30865/jurikom.v8i6.3680>
- Destiariono, M. E., & Darwanto. (2022). Dampak Produksi Beras Terhadap Kemiskinan Moderat: Ardl Bounds-Testing Approach. *Jurnal Ekonomi Pertanian Dan Agribisnis (JEPA)*, 6(4), 1352–1361.
- Dewi, N. L. P. P., Purnama, I. N., & Utami, N. W. (2022). Penerapan Data Mining Untuk Clustering Penilaian Kinerja Dosen Menggunakan Algoritma K-Means (Studi Kasus: STMIK Primakara). *Jurnal Ilmiah Teknologi Informasi Asia*, 16(2), 105. <https://doi.org/10.32815/jitika.v16i2.761>
- Mardiyah, S., Fajar, M. Y., & Badruzzaman, F. H. (2022). Penggunaan Forecasting dan Goal Programming dalam Optimasi Perencanaan Produksi Beras. *Bandung Conference Series: Mathematics*, 2(1), 83–93. <https://doi.org/10.29313/bcsm.v2i1.2033>
- Miharja, J., & Suhendri, S. (2021). Penerapan Data Mining Penerimaan Karyawan Menggunakan Metode Naive Bayes Classifier. *Proceeding SENDIU*, 579–583.
- Normah, Nurajizah, S., & Salbinda, A. (2021). Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Fashion Hijab Banten. *Jurnal Teknik Komputer AMIK BSI*, 7(2), 158–163. <https://doi.org/10.31294/jtk.v4i2>
- Saputra, M. R. T., & Purnomo, D. (2022). Identifikasi Dan Analisis Determinan Ketahanan Pangan Di Indonesia (Studi Empiris Produksi Beras 1980-2020). *Jurnal Ekonomi, Teknologi Dan Bisnis (JETBIS)*, 1(2), 66–77. <https://doi.org/10.57185/jetbis.v1i2.11>
- Sembiring, C. S. D. B., Hanum, L., & Tamba, S. P. (2022). Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Menentukan Judul Skripsi Dan Jurnal Penelitian (Studi Kasus Ftik Unpri). *Jurnal Sistem Informasi Dan Ilmu Komputer Prima (JUSIKOM PRIMA)*, 5(2), 80–85. <https://doi.org/10.34012/journalsisteminformasidanilmukomputer.v5i2.2393>
- Sudirman, Windarto, A. P., & Wanto, A. (2018). Data mining tools | rapidminer: K-means method on clustering of rice crops by province as efforts to stabilize food crops in Indonesia. *IOP Conference Series: Materials Science and Engineering*, 420(1), 1–8. <https://doi.org/10.1088/1757-899X/420/1/012089>
- Zaeroni, R., & Rustariyuni, S. (2016). Pengaruh Produksi Beras, Konsumsi Beras Dan Cadangan Devisa Terhadap Impor Beras Di Indonesia. *E-Jurnal Ekonomi Pembangunan Universitas Udayana*, 5(9), 993–1010.
- Zulfa, N., Auliya, R. I., & Zaenal, A. (2021). Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means. *Jurnal Teknologi Dan Sistem Informasi (JTISI)*, 2(2), 100–108.