

Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Daerah Produksi Kakao

<http://dx.doi.org/10.28932/jutisi.v8i2.4897>

Riwayat Artikel

Received: 07 Juni 2022 | Final Revision: 10 Agustus 2022 | Accepted: 12 Agustus 2022

Creative Commons License 4.0 (CC BY – NC)



Nur Awalia Syukri Zainal Abidin[✉]#1, Rizna Dwi Avila^{#2}, Arudji Hermatyar^{#3}, Rismayani^{#4}

Program Studi Teknik Informatika, Universitas Dipa Makassar
Jl. Perintis Kemerdekaan Km. 9 Makassar, 90245, Indonesia

¹awalabidin@gmail.com

²riznadwiavila30@gmail.com

³lostarmy1211@gmail.com

⁴rismayani@undipa.ac.id

✉Corresponding author: awalabidin@gmail.com

Abstrak — Kakao merupakan salah satu komoditas unggulan dari sektor perkebunan, produksi kakao mampu meningkatkan devisa negara. Di Indonesia khususnya Provinsi Sulawesi Selatan memiliki produksi kakao yang besar, hampir seluruh kabupaten/kota di Sulawesi Selatan dapat memproduksi kakao. Penelitian ini bertujuan untuk mengelompokkan daerah produksi kakao di Provinsi Sulawesi Selatan. Algoritma yang digunakan adalah K-Means dan K-Medoids, algoritma K-Means mengelompokkan data dengan membaginya menjadi beberapa cluster berdasarkan karakteristik yang sama. Sedangkan algoritma K-Medoids memilih objek nyata untuk mewakili cluster. Pada penelitian ini, kedua algoritma tersebut dibandingkan dengan menggunakan satu dataset. Perbandingan dilakukan dengan melihat nilai Davies-Bouldin Index (DBI) pada RapidMiner. Kemudian hasil yang diperoleh berdasarkan penelitian ini adalah pengelompokan menggunakan algoritma K-Means lebih efektif dibandingkan dengan menggunakan K-Medoids dalam pengelompokan daerah produksi kakao di Provinsi Sulawesi Selatan. Dengan nilai DBI yang diperoleh, K-Means dan K-Medoids memiliki nilai DBI masing-masing sebesar 0,292 dan 0,365.

Kata Kunci— Davies-Bouldin Index; Kakao; K-Means; K-Medoids.

Comparison of K-Means and K-Medoids Algorithms for Grouping Cocoa Production Areas

Abstract — Cocoa is one of the leading commodities from the plantation sector, cocoa production is able to increase the country's foreign exchange. In Indonesia, especially South Sulawesi Province has a large cocoa production, almost all districts/cities in South Sulawesi can produce cocoa. This study aims to classify cocoa production areas in South Sulawesi Province. The algorithm used is K-Means and K-Medoids, the K-Means algorithm groups data by dividing it into several clusters based on the same characteristics. While the K-Medoids algorithm chooses real objects to represent the cluster. In this study, the two algorithms were compared using one dataset. The comparison is made by looking at the Davies-Bouldin Index (DBI) value on RapidMiner. The results obtained based on this study are that grouping using the K-Means algorithm is more effective than using K-Medoids in grouping cocoa production areas in South Sulawesi Province. With the DBI values obtained, K-Means and K-Medoids have DBI values of 0.292 and 0.365, respectively.

Keywords— Cocoa; Davies-Bouldin Index; K-Means; K-Medoids.

I. PENDAHULUAN

Kakao adalah salah satu dari beberapa komoditas unggulan dari sektor perkebunan yang memberikan kontribusi dalam hal meningkatkan devisa negara juga sebagai sumber pendapatan petani, mendorong agribisnis maupun agroindustri dan berperan dalam pengembangan wilayah [1]. Di Indonesia saat ini salah satu provinsi yang menjadi produsen kakao terbesar yaitu Sulawesi Selatan. Sulawesi Selatan terbagi menjadi 24 Kabupaten/Kota, hampir semua Kabupaten/Kota yang terdapat di Sulawesi Selatan memproduksi kakao. Pada tahun 2021, luas areal produksi kakao di tiap-tiap Kabupaten/Kota di Sulawesi Selatan sebesar 187.422 hektar, sedangkan untuk jumlah produksi kakao mencapai 106.380 ton. Berdasarkan luas areal dan tingginya angka produksi tersebut maka diperlukan adanya pengelompokan daerah produksi kakao di Sulawesi Selatan. Hal tersebut bertujuan agar dapat mengetahui Kabupaten/Kota mana saja yang memiliki hasil produksi kakao tertinggi dan yang masih kurang, dengan begitu dapat membantu mengoptimalkan produksi komoditas kakao serta meningkatkan intensifikasi, terlebih beberapa daerah dianggap memiliki potensi yang besar untuk memproduksi kakao.

Adapun algoritma yang digunakan untuk melakukan pengelompokan ialah K-Means beserta K-Medoids. Algoritma K-means adalah metode data mining yang termasuk unsupervised dengan kata lain tanpa supervisi ketika melakukan proses pemodelan, algoritma ini juga merupakan algoritma dengan sistem partisi untuk melakukan pengelompokan data. Pada proses pengelompokan data, terdapat dua jenis metode, yakni Hierarchical dan Non-Hierarchical, untuk algoritma K-Means tergolong dalam metode Non-Hierarchical atau dapat juga disebut dengan istilah Partitional Clustering. Algoritma k-Means melakukan pembagian data dengan membaginya ke dalam kelompok atau cluster, yang mana data di dalam satu cluster memiliki ciri-ciri yang sama pada tiap datanya namun memiliki ciri-ciri yang berbeda terhadap data yang terdapat pada cluster lainnya [2].

Algoritma kedua yang diaplikasikan yaitu K-Medoids, algoritma ini dikenal juga dengan sebutan Algoritma Partitioning Around Medoid yang disingkat PAM. Pada algoritma K-Medoids pengelompokan partisi memakai objek nyata yang dipilih untuk mewakili klaster [3]. Bertujuan untuk menghasilkan partisi yang tidak terlalu rentan berhubung karena nilai-nilai pada dataset yang ekstrim [4].

Terdapat penelitian terkait yang sebelumnya pernah dilakukan diantaranya oleh Hotma Dame Tampubolon et.al. [5] untuk clustering tindak kriminalitas berdasarkan Provinsi. Penelitian tersebut bertujuan mengelompokkan Provinsi yang rawan tindak kriminalitas agar dapat menentukan Provinsi mana saja yang membutuhkan ditingkatkan pengawasannya. Hasil penelitian yang dapat ditarik yaitu melalui perhitungan K-Means, didapatkan kelompok tindak kejahatan tinggi (cluster 1) terdapat 6 anggota, dan untuk kelompok tindak kejahatan rendah (cluster 2) ada 28 anggota. Sedangkan untuk hasil K-Medoids didapatkan untuk tindak kejahatan tinggi (cluster 1) memiliki 7 anggota dan 27 anggota tergabung dalam kelompok tindak kejahatan rendah (cluster 2). Kemudian penelitian lainnya yang melakukan perbandingan atau dengan kata lain komparasi K-Means dan K-Medoids adalah riset Novita Lestari Anggreini dan Shandy Tresnawati [6]. Riset tersebut menjelaskan tentang komparasi dari kedua algoritma untuk pembagian dan hasil dari riset ini dijadikan rekomendasi supaya dapat menetapkan strategi promosi yang tepat dalam pelaksanaan penerimaan mahasiswa baru di Politeknik TEDC Bandung. Hasil yang diperoleh dari eksperimen ini adalah K-Means memiliki kinerja yang lebih ideal bila dipadankan dengan metode K-Medoids dalam kasus penentuan program promosi yang sesuai untuk Politeknik yang berlokasi di Bandung tersebut. Penelitian yang dilakukan oleh Reza Gustrinda dan Dadang Iskandar Mulyana [7] membahas bahwa terdapat 2 macam jenis analisis cluster pada metode clustering yang bertautan, yakni K-Means dan K-Medoids. Menurut ulasan tersebut, penggunaan dua algoritma tersebut untuk melakukan analisa terhadap produk apa saja yang sangat disukai dan kurang disukai oleh konsumen PT. XYZ Indonesia. Adapun data yang digunakan diambil dari data transaksi penjualan pada bulan Mei hingga bulan Oktober 2020. Data tersebut berupa kode dan nama produk, tanggal dilakukannya transaksi serta total penjualan dan stok barang dari perusahaan tersebut. Berdasarkan implementasi kedua algoritma tersebut diperoleh hasil bahwa algoritma K-Means dinilai sebagai metode yang cocok digunakan untuk permasalahan dalam penentuan produk terbaik dengan nilai Davies Bouldin sebesar -0,430. Penelitian yang dilakukan oleh Mufidah Herviany, et.al. [8] bertujuan untuk melakukan pengelompokan kawasan rawan bencana tanah longsor di Provinsi Jawa Barat, hal tersebut penting mengingat tingginya kasus bencana tanah longsor yang ada di Jawa Barat, sehingga pemerintah dapat menangani bencana secara terkendali dan tepat. Pada penelitian ini juga melakukan komparasi metode K-Means juga K-Medoids sebagai teknik klasterisasinya. Pengelompokan dengan algoritma K-Means disebutkan lebih diatas daripada K-Medoids yang mana merupakan kesimpulan eksperimen. Pada penelitian tersebut didapatkan nilai $k = 6$, yang merupakan nilai k paling ideal. Selanjutnya ada penelitian terkait implementasi dari K-Means dan K-Medoids yang digunakan sebagai teknik segmentasi pelanggan pada data transaksi e-commerce. Tujuan dilakukannya penelitian ini adalah mengimplementasikan metode clustering pada data transaksi e-commerce dengan memakai metode K-Means dan k-Medoids. Percobaan menampilkan nilai performa K-Medoids sebesar 0,337575 yang mana hal tersebut membuktikan bahwa pemetaan pelanggan dengan memakai algoritma K-Medoids lebih unggul dari K-Means [9]. Sena Wijayanto dan M Yoka Fathoni dalam penelitian [10], menggunakan algoritma K-Means sebagai metode untuk pembagian Kabupaten/Kota. Penerapan metode K-Means ini menghasilkan 3 cluster yaitu C0 yang mencerminkan 12 daerah yang memiliki produktivitas padi sedang. C1 mencerminkan 18 daerah yang memiliki produktivitas padi rendah, dan C2 yang mencerminkan 12 daerah yang memiliki produktivitas padi tinggi. Penelitian yang dilakukan pada

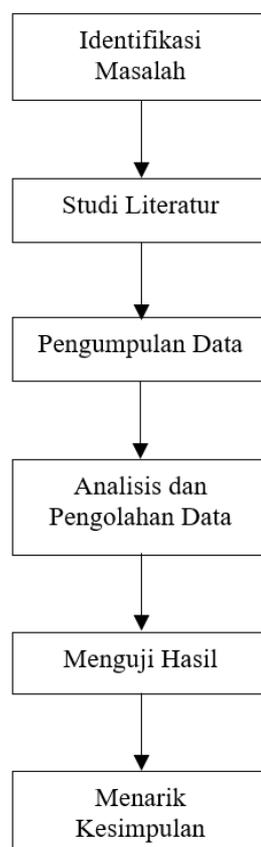
tahun 2021 yang lalu [11], penelitian ini tujuannya untuk mengaplikasikan metode K-Medoids pada proses pengelompokan tingkat masa studi tepat waktu mahasiswa. Informasi yang dihasilkan dari pengaplikasian algoritma K-Medoids pada Metode clustering ini dapat bermanfaat bagi pihak pendidikan khususnya pihak kampus dalam melakukan kegiatan-kegiatan untuk mengurangi jumlah mahasiswa yang diperkirakan tidak mampu menuntaskan masa studi tepat waktu. Hasil yang diperoleh yakni sebanyak 35 mahasiswa terhitung sebagai kelompok mahasiswa yang bisa menuntaskan masa studinya tepat waktu dari total mahasiswa 50 orang, lalu 15 orang mahasiswa lainnya tercatat ke dalam kelompok mahasiswa yang tidak bisa menuntaskan masa studi secara tepat waktu.

Tujuan penelitian ini adalah untuk mengelompokkan daerah produksi kakao yang ada di Sulawesi Selatan dengan menggunakan K-Means serta K-Medoids, agar dapat diketahui daerah mana saja yang memiliki produksi kakao yang kurang, lumayan banyak, berlimpah, dan sangat berlimpah.

II. METODE PENELITIAN

A. Prosedur Penelitian

Gambar 1 memvisualisasikan prosedur atau tahapan – tahapan yang dilakukan dalam pengelompokan daerah produksi kakao di Sulawesi Selatan.



Gambar 1. Prosedur Penelitian.

Tahapan identifikasi masalah adalah suatu proses untuk merumuskan masalah dalam penelitian. Agar dalam proses analisis data dapat lebih terstruktur, sehingga dapat memahami lebih rinci mengenai ruang lingkup masalah penelitian. Tahapan berikutnya ialah studi literatur. Studi literatur disini merujuk pada proses mencari, mengumpulkan serta mempelajari berbagai referensi yang dijadikan sebagai bahan penunjang dalam penelitian yang diangkat. Dalam hal ini literatur yang dimaksudkan ialah berupa jurnal-jurnal penelitian maupun buku yang membahas tentang K-Means dan juga algoritma K-Medoids.

Selanjutnya adalah mengumpulkan data yang akan diproses. Setelah data terkumpul masuk ke tahapan analisis dan pengolahan atau pemrosesan data. Pada fase ini, data yang dikumpulkan akan melalui proses penyortiran terlebih dahulu sehingga hanya mengambil data yang diperlukan saja. Setelah itu data diolah dan dievaluasi menggunakan RapidMiner,

dengan menentukan jumlah cluster yang terbaik melalui Davies Bouldin Index. Kemudian tahap terakhir ialah menarik kesimpulan berdasarkan hasil yang diperoleh dari tahap sebelumnya.

B. K-Means

Algoritma K-Means memecah data berdasarkan jarak tiap data terhadap cluster yang sudah ditentukan. K-Means bertumpu pada fungsi untuk menghitung data yang memiliki karakteristik yang sama. Berikut ini langkah-langkah untuk pengelompokan data dengan K-Means [12] :

1. Menetapkan jumlah klaster.
2. Inisialisasi awal dan centroid (pusat klaster) ditentukan secara acak.
3. Menaksir jarak dari tiap centroid dengan menggunakan formula Euclidean Distance.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

4. Menghitung centroid baru dengan kedudukan baru pula dengan cara mencari rata-rata objek pada cluster
5. Mengulangi langkah sebelumnya sampai hasil pengelompokan tidak mengalami perubahan.

C. K-Medoids

Algoritma K-Medoids menggunakan metode partisi clustering dalam melakukan klusterisasi sejumlah n objek menjadi suatu kelompok [11]. Tahapan proses algoritma K-Medoids adalah [7]:

1. Menetapkan jumlah klaster.
2. Menentukan centroid dengan random sebanyak jumlah klaster.
3. Mengukur jarak dari setiap data dengan formula berikut:

$$d_1 = \sqrt{(x_1 - x_{c0})^2 + (y_1 - y_{c0})^2} \quad (2)$$

Adapun nilai dari x_1 mewakili luas areal dan nilai dari y_1 mewakili jumlah produksi berdasarkan dataset pada tabel 1. Sehingga akan dicari satu per satu berdasarkan data yang ada untuk mengetahui jarak terdekatnya. Dalam hal ini, nilai pada dataset untuk x_1 dan y_1 akan dikurangkan dengan data yang menjadi centroid kemudian dikuadratkan lalu dijumlah, setelah itu diakarkan.

4. Lakukan repetisi medoids dengan mengukur jarak dengan memakai metode taksir jarak, kepada semua data yang bertindak sebagai data proses.
5. Menetapkan objek yang tidak representatif dengan random.

D. Davies-Bouldin Index (DBI)

D. L. Davies dan D. W. Bouldin pertama kali mempublikasikan Davies-Bouldin Index pada tahun 1979. Davies Bouldin berguna untuk evaluasi cluster yang didasarkan pada nilai kohesi dan separasi. Formula yang digunakan untuk mengetahui nilai kohesi menggunakan Sum of square within cluster (SSW) pada cluster ke-i :

$$SSW_i = \frac{1}{m_i} \sum_{j=1}^{m_i} d(x_j, c_i) \quad (3)$$

Pada persamaan SSW m_i adalah jumlah data yang ada pada cluster serta c_i merupakan centroid dari cluster. Kemudian untuk nilai separasi dihitung dengan *Sum of square between cluster* (SSB) yang dirumuskan:

$$SSB_{i,j} = d(c_i, c_j) \quad (4)$$

Jika nilai kohesi dan separasi telah didapatkan selanjutnya mencari pengukuran rasio (R_{ij}), dengan rumus:

$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_{i,j}} \quad (5)$$

Lalu dengan nilai rasio tersebut digunakan untuk mencari nilai DBI dengan formula:

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{j \neq i} (R_{i,j}) \quad (6)$$

Pada formula DBI nilai k mewakili jumlah cluster yang digunakan. Jika nilai DBI minimal artinya suatu kelompok atau clusters memiliki klasterisasi yang optimal[4]. Nilai DBI bersifat absolut atau mutlak. Penelitian ini akan mengaplikasikan metode Davies Bouldin untuk mengelompokkan Kabupaten/Kota berdasarkan luas areal dan jumlah produksi kakao.

E. Metode Elbow

Metode elbow adalah metode yang berguna untuk memberikan informasi dalam menentukan jumlah cluster terbaik dengan cara menghitung nilai *Sum of Square Error* (SSE) dari setiap cluster. Berikut ini adalah rumus SSE pada K-Means:

$$SSE = \sum_{K=1}^K \sum_{x_i \in S_k} \|x_i - \hat{c}_k\|_2^2 \quad (7)$$

F. Data Penelitian

Data yang dipakai untuk riset ini didapatkan dari secara online melalui link dari Badan Pusat Statistik (BPS). Adapun data tersebut adalah luas areal (ha) dan jumlah produksi kakao (ton) Kabupaten/Kota di provinsi Sulawesi Selatan pada 2021. Sebelum data diproses pada aplikasi RapidMiner dilakukan seleksi data, adapun data yang dipakai hanya data yang diperlukan seperti yang tertera dalam Tabel 1.

TABEL 1
DATA LUAS AREAL DAN JUMLAH PRODUKSI KAKAO

Kabupaten/Kota	Luas Areal (hektar)	Produksi (ton)
Kepulauan Selayar	533	181
Bulukumba	7417	4403
Bantaeng	4696	3644
Jeneponto	153	17
Takalar	25	4
Gowa	3510	1147
Sinjai	4315	281
Maros	1585	404
Pangkajene dan Kepulauan	297	63
Barru	759	273
Bone	16163	7279
Soppeng	11430	3774
Wajo	14010	8404
Sidenreng Rappang	6728	4552
Pinrang	19568	11099
Enrekang	6534	2124
Luwu	28009	22000
Tana Toraja	4226	1334
Luwu Utara	39562	30395
Luwu Timur	13730	3707
Toraja Utara	1548	297
Kota Makassar	0	0

Kabupaten/Kota	Luas Areal (hektar)	Produksi (ton)
Kota Parepare	0	0
Kota Palopo	2604	998

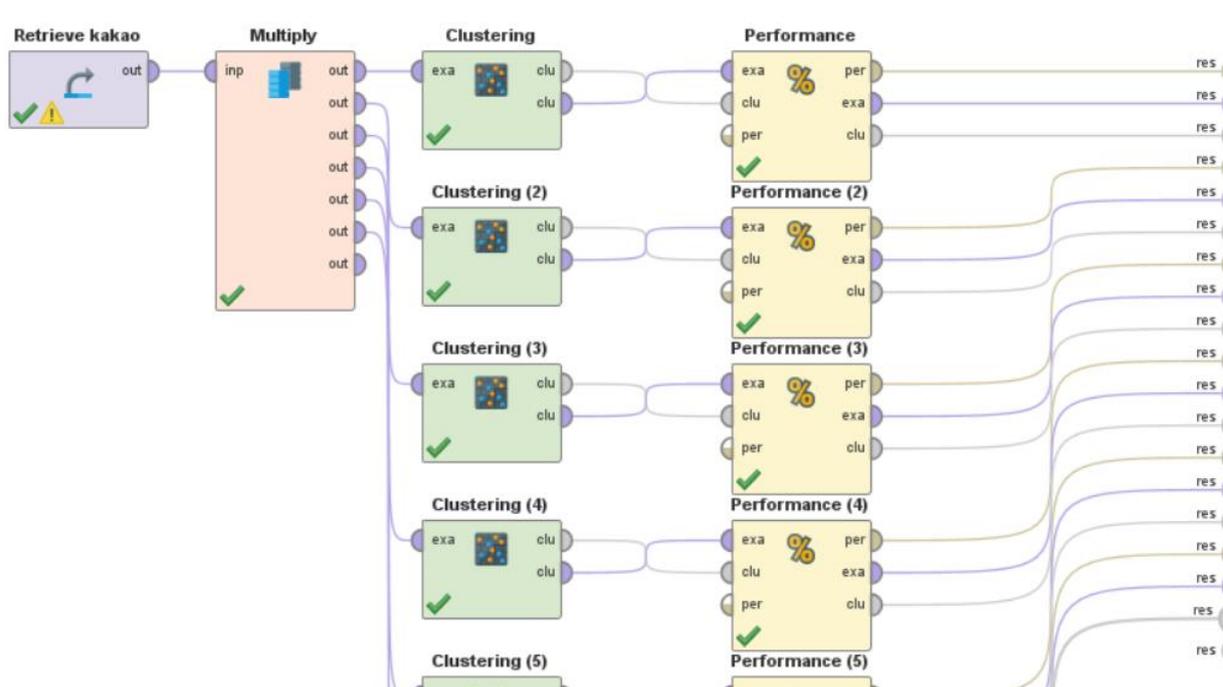
III. HASIL DAN PEMBAHASAN

Hasil perhitungan SSE pada K-Means untuk setiap nilai k dapat dilihat pada Tabel 2.

TABEL 2
 PERHITUNGAN ELBOW

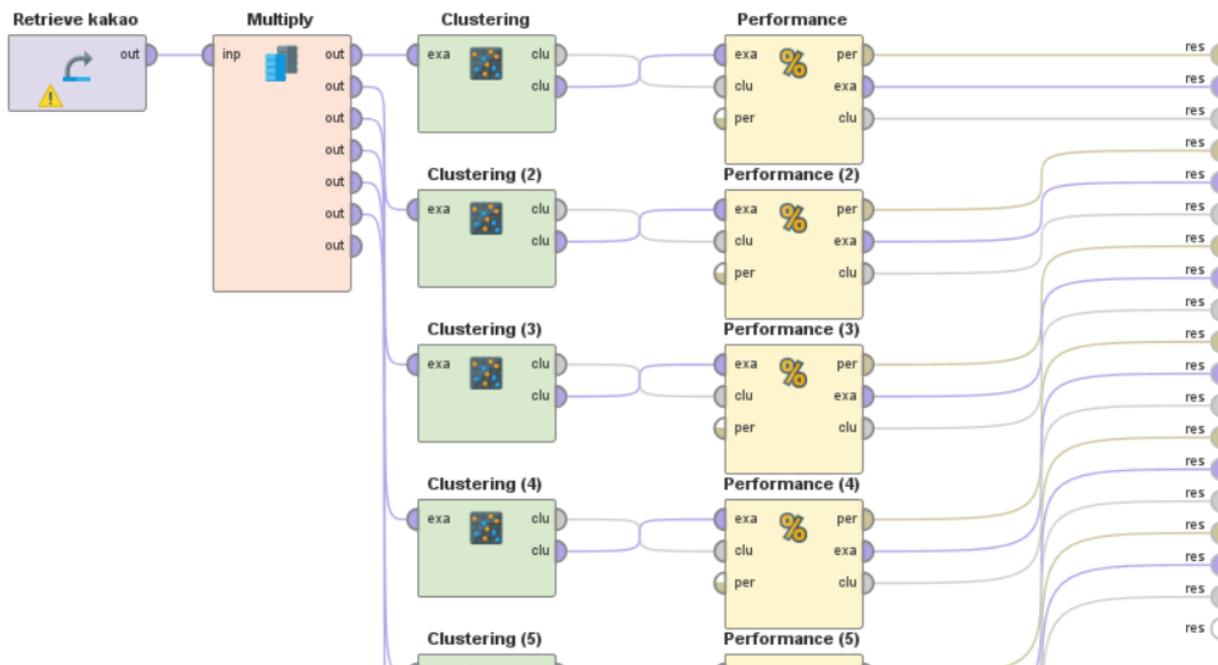
Jumlah Klaster	SSE
3	18075252,3
4	870,25
5	621732,25
6	19496640,3
7	19611612,3
8	10794510,3

Pemrosesan data untuk riset ini menggunakan software RapidMiner 9.10. Data yang telah didapatkan diinputkan ke dalam excel terlebih dahulu lalu data tersebut diinputkan ke aplikasi RapidMiner. Berikut ini Gambar 2 dan Gambar 3 yang merupakan desain input proses serta Gambar 4 dan Gambar 5 yang menunjukkan grafik dari hasil proses klusterisasi pada RapidMiner.



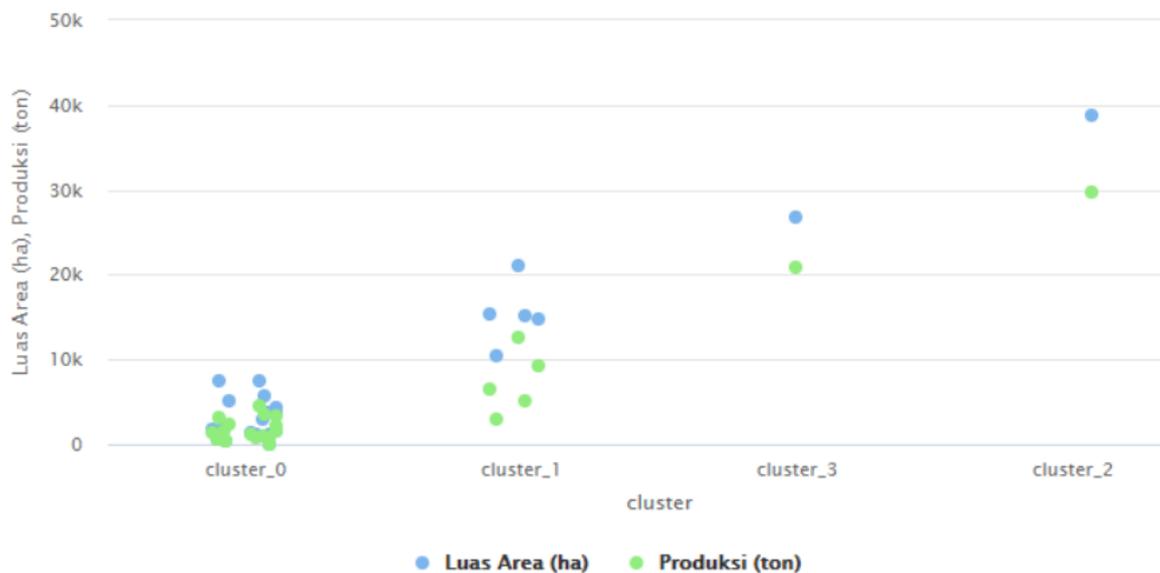
Gambar 2. Gambar desain input untuk K-Means

Pada Gambar 2 dapat dilihat bahwa desain input untuk K-Means pada aplikasi RapidMiner.



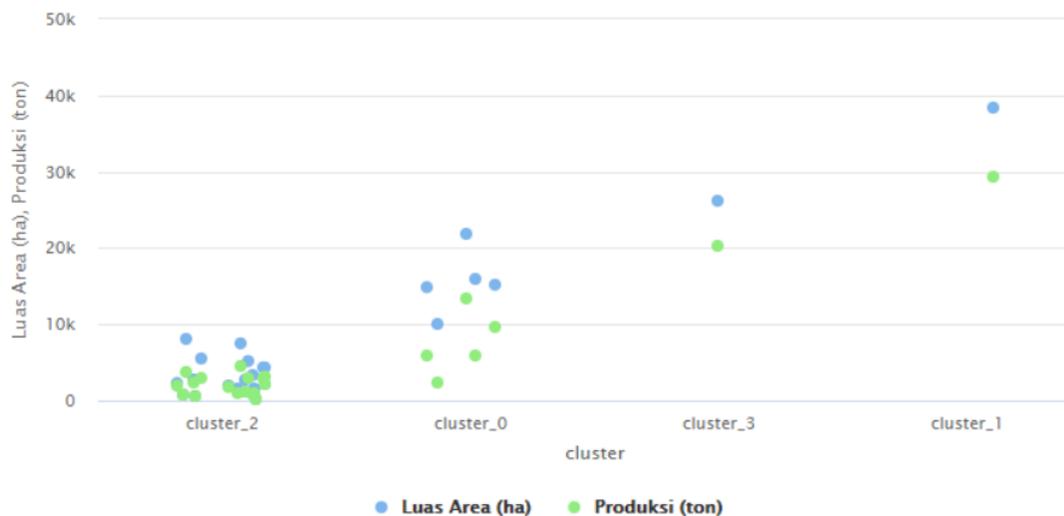
Gambar 3. Gambar desain input untuk K-Medoids

Pada Gambar 3 dapat dilihat bahwa desain input untuk K-Medoids pada aplikasi RapidMiner.



Gambar 4. Gambar grafik pembagian cluster K-Means

Pada Gambar 4 dapat dilihat bahwa cluster 0 memiliki anggota yang paling banyak, disusul oleh cluster 1, cluster 2, dan cluster 3. Cluster 0 merupakan daerah dengan luas area serta produksi kakao yang terkecil, cluster 1 ialah daerah dengan luas area yang terbilang cukup dengan produksi yang lumayan, cluster 3 adalah daerah produksi kakao yang memiliki luas area yang cukup besar dan produksi yang banyak pula, cluster 2 adalah daerah produksi kakao yang luas area dan produksi kakao terbesar.



Gambar 5. Gambar grafik pembagian cluster K-Medoids

Pada Gambar 5 terdapat cluster 2 yang merupakan daerah dengan luas area serta produksi kakao yang terkecil, cluster 0 yaitu daerah dengan luas area yang terbilang cukup dengan produksi yang lumayan, cluster 3 merupakan daerah produksi kakao dengan luas area yang cukup besar dan produksi yang banyak, cluster 1 adalah daerah produksi kakao yang luas area dan produksi kakao terbesar atau sangat berlimpah.

Akurasi dari clustering dapat dilihat berdasarkan dari nilai Davies Bouldin yang diperoleh dari aplikasi RapidMiner, lebih jelasnya perhatikan Tabel 3.

TABEL 3
NILAI PERBANDINGAN DBI BERDASARKAN JUMLAH KLASTER

Jumlah Klaster	K-Means	K-Medoids
3	0.429	0.491
4	0.292	0.365
5	0.406	0.631
6	0.341	0.801
7	0.342	0.656
8	0.349	0.566

Dari tabel 2 dapat dilihat bahwa untuk jumlah klaster atau nilai $k = 4$ merupakan nilai k yang terbaik karena memiliki nilai DBI terendah, dengan kata lain, untuk nilai $k = 4$ merupakan nilai k yang sangat optimal baik untuk K-Means ataupun untuk K-Medoids. Nilai DBI dari K-Means senilai 0,292 dan K-Medoids memiliki nilai DBI senilai 0,365. Namun dapat dilihat ternyata nilai DBI untuk K-Means lebih kecil jika disandingkan dengan nilai DBI K-Medoids. Hal ini berarti pengelompokan menggunakan algoritma K-Means lebih efektif dan akurat daripada menggunakan K-Medoids dalam pengelompokan daerah produksi kakao Provinsi Sulawesi Selatan.

Kemudian untuk data dari hasil pengelompokan Kabupaten/Kota dengan nilai k yang terbaik yaitu sebanyak 4 clusters menggunakan algoritma K-Means bisa diperhatikan pada tabel 4.

TABEL 4
DATA HASIL PENGELOMPOKAN DAERAH DENGAN ALGORITMA K-MEANS

Cluster	Jumlah Kabupaten/Kota	Anggota
0	17	Kepulauan Selayar, Bulukumba, Bantaeng, Jeneponto, Takalar, Gowa, Sinjai, Maros, Pangkajene dan Kepulauan, Barru, Sidenreng Rappang, Enrekang, Tana Toraja, Toraja Utara, Kota Makassar, Kota Parepare, Kota Palopo

Cluster	Jumlah Kabupaten/Kota	Anggota
1	5	Bone, Soppeng, Wajo, Pinrang, Luwu Timur
2	1	Luwu Utara
3	1	Luwu

Berdasarkan tabel 3 diketahui bahwa cluster 0 (C0) adalah cluster dengan anggota terbanyak yang merupakan daerah dengan produksi kakao yang masih kurang. Cluster 1 (C1) merupakan daerah dengan produksi kakao yang lumayan banyak. Cluster 2 (C2) merupakan daerah penghasil kakao yang berlimpah, dan merupakan yang tertinggi di Sulawesi Selatan. Serta cluster 3 (C3) yang juga merupakan daerah penghasil kakao yang terbilang sukses. Adapun kriteria sukses yaitu produksi kakao yang sangat berlimpah didukung dengan area produksi yang sangat luas, sehingga C3 memiliki potensi yang sangat besar sebagai daerah produksi kakao. Diharapkan dengan adanya pembagian daerah ini mampu menjadi solusi agar produksi kakao lebih ditingkatkan lagi khususnya pada daerah yang notabene memiliki potensi yang besar untuk produksi kakao.

IV. SIMPULAN

Perbandingan terhadap K-Means serta K-Medoids pada pengelompokan daerah produksi kakao di Sulawesi Selatan menunjukkan bahwa K-Means mempunyai nilai Davies-Bouldin lebih kecil bila disandingkan dengan nilai Davies-Bouldin yang didapatkan algoritma K-Medoids, yakni senilai 0,292 untuk nilai $k = 4$. Sehingga diperoleh C0 sebanyak 17 daerah, C1 beranggotakan 5 daerah, dan untuk C2 juga C3 sama-sama beranggotakan 1 daerah. Hal ini berarti pengelompokan menggunakan algoritma K-Means lebih efektif daripada menggunakan K-Medoids dalam pengelompokan daerah produksi kakao Provinsi Sulawesi Selatan. Untuk penelitian selanjutnya disarankan agar dapat melakukan pengembangan terhadap penelitian ini seperti melakukan pengelompokan daerah dengan menggunakan algoritma ataupun metode lain dan dapat dibandingkan sehingga dapat ditentukan metode yang lebih tepat dalam pengelompokan daerah produksi kakao.

UCAPAN TERIMA KASIH

Ucapan terima kasih dipersembahkan kepada Universitas Dipa Makassar yang sudah mendukung penelitian ini hingga dapat diselesaikan dengan optimal. Terima kasih kami sampaikan pula kepada dosen pembimbing atas pedomannya dalam penyusunan jurnal, juga kepada seluruh pihak terlibat yang sudah membantu dan tak dapat diucapkan satu persatu.

DAFTAR PUSTAKA

- [1] E. Nurhadi, S. I. Hidayat, P. N. Indah, S. Widayanti and G. I. Harya, "Keberlanjutan Komoditas Kakao Sebagai Produk Unggulan Agroindustri dalam Meningkatkan Kesejahteraan Petani," *AGRIEKONOMIKA*, vol. 8, no. 1, pp. 51-61, 2019.
- [2] H. SY, Rismayani and A. Syam, "Data Mining Menggunakan Algoritma K-Means Pengelompokan Penyebaran Diare Di Kota Makassar," in *Prosiding Seminar Ilmiah Sistem Informasi Dan Teknologi Informasi*, Makassar, 2019.
- [3] A. Wanto, *Data Mining : Algoritma dan Implementasi*, Yayasan Kita Menulis, 2020.
- [4] I. Kamila, U. Khairunnisa and M. Mustakim, "Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau," *Rekayasa dan Manajemen Sistem Informasi*, vol. 5, no. 1, pp. 119-125, 2019.
- [5] H. D. Tampubolon, M. Safii and D. Suhendro, "Penerapan Algoritma K-Means dan K-Medoids Clustering untuk Mengelompokkan Tindak Kriminalitas Berdasarkan Provinsi," *Jurnal Ilmu Komputer dan Teknologi*, vol. 2, no. 2, pp. 6-12, 2021.
- [6] N. L. Anggreini and S. Tresnawati, "Komparasi Algoritma K-Means dan K-Medoids Untuk Menangani Strategi Promosi Di Politeknik TEDC Bandung," *Jurnal TEDC*, vol. 14, no. 2, pp. 120-127, 2020.
- [7] R. Gustrinda and D. I. Mulyana, "Penerapan Data Mining Dalam Pemilihan Produk Unggulan dengan Metode Algoritma K-Means Dan K-Medoids," *Media Informatika Budidarma*, vol. 6, no. 1, pp. 27-34, 2022.
- [8] M. Herviany, S. P. Delima, T. Nurhidayah and K. Kasini, "Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Daerah Rawan Tanah Longsor Pada Provinsi Jawa Barat: Comparison of K-Means and K-Medoids Algorithms for Grouping Landslide Prone Areas in West Java Province," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 1, no. 1, pp. 34-40, 2021.
- [9] R. Siagian, P. Sirait and A. Halim, "Penerapan Algoritma K-Means dan K-Medoids untuk Segmentasi Pelanggan pada Data Transaksi E-Commerce," *Jurnal Sistem Informasi*, vol. 11, no. 2, pp. 260-270, 2022.
- [10] S. W. a. M. Y. Fathoni, "Pengelompokan Produktivitas Tanaman Padi di Jawa Tengah Menggunakan Metode Clustering K-Means," *JUPITER (Jurnal Penelitian Ilmu dan Teknologi Komputer)*, vol. 13, no. 2, pp. 212-219, 2021.
- [11] F. Firzada and Y. Yuhandri, "Klasterisasi Tingkat Masa Studi Tepat Waktu Mahasiswa Menggunakan Algoritma K-Medoids," *Jurnal Sistem Informasi dan Teknologi*, vol. 3, no. 3, p. 162-168, 2021.
- [12] M. Wahyudi, M. Masitha, R. Saragih and S. Solikhun, *Data Mining: Penerapan Algoritma K-Means Clustering dan K-Medoids Clustering*, Yayasan Kita Menulis, 2020.