

Pemetaan Produksi Tanaman Tomat di Indonesia Berdasarkan Provinsi Menggunakan Algoritma K-Means Clustering

Syaifuddin¹, Ramlah², Irma Hakim³, Yunida Berliana⁴, Nurhayati^{5,*}

¹ Politeknik Pembangunan Pertanian Gowa, Indonesia

² Sekolah Tinggi Pertanian Kutai Timur, Sangatta, Indonesia

³ Universitas Muhammadiyah Makassar, Makassar, Indonesia

⁴ Universitas Tjut Nyak Dhien, Medan, Indonesia

⁵ Universitas Islam Sumatera Utara, Medan, Indonesia

Email: ¹syiaifuddinanwar.gowa@gmail.com, ²ramlah.proteksi44@gmail.com, ³campus_gardenia@yahoo.co.id,

⁴yd66berliana@gmail.com, ^{5,*}nurhayati@uisu.ac.id

Submitted: 27/08/2022; Accepted: 29/08/2022; Published: 30/08/2022

Abstrak—Tomat merupakan salah satu jenis sayuran komoditas hortikultura yang penting, karena memiliki nilai ekonomis yang tinggi dan kebutuhan akan tanaman ini terus meningkat seiring dengan pertambahan jumlah penduduk, bertambahnya tingkat pendapatan penduduk serta tingginya kesadaran masyarakat terhadap pentingnya nilai gizi. Oleh sebab itu tujuan dilakukan nya penelitian ini untuk melihat dan melakukan pemetaan produksi tanaman tomat di Indonesia berdasarkan provinsi berupa cluster (pengelompokkan). Data penelitian yang digunakan dalam makalah ini adalah data produksi tomat di Indonesia berdasarkan provinsi dalam kurun waktu lima tahun terakhir (tahun 2017- tahun 2021) yang diperoleh dari Dinas Pertanian Kabupaten/Kota masing-masing provinsi dan Badan Pusat Statistik Indonesia. Algoritma yang diusulkan pada penelitian ini adalah K-Means Clustering dengan bantuan RapidMiner. Hasil dari usulan makalah ini berupa pengelompokan dan pemetaan produksi tanaman tomat di Indonesia yang terbagi menjadi 5 (lima) zona, antara lain Zona Hitam (daerah yang memiliki produksi tomat sangat tinggi) yang terdiri dari 1 provinsi, Zona Hijau (daerah yang memiliki produksi tinggi) yang terdiri dari 2 provinsi, Zona Biru (daerah yang memiliki produksi sedang) yang terdiri dari 4 provinsi, Zona Biru Muda (daerah yang memiliki produksi rendah) yang terdiri dari 8 provinsi, dan Zona Oranye (daerah yang memiliki produksi cukup rendah) yang terdiri dari 18 provinsi.

Kata Kunci: Pemetaan; Tomat; Pengelompokkan; Data Mining; K-Means

Abstract—Tomato is one of the essential horticultural commodity vegetables because it has high economic value. The need for this plant continues to increase along with the increase in population, income levels, and heightened public awareness of the importance of nutritional value. Therefore, this research aims to see and map the production of tomato plants in Indonesia by the province in the form of clusters (grouping). The research data used in this paper is data on tomato production in Indonesia by the province in the last five years (2017-2021) obtained from the District/City Agriculture Service of each province and the Indonesian Central Statistics Agency. The algorithm proposed in this study is K-Means Clustering with the help of RapidMiner. The results of the proposed paper are grouping and mapping of tomato production in Indonesia, which is divided into 5 (five) zones, including the Black Zone (areas with very high tomato production), which consists of 1 province, Green Zone (areas with high production of tomatoes). Which consists of 2 provinces, the Blue Zone (areas with moderate production), which consists of 4 provinces. The Light Blue Zone (areas with low production), which consists of 8 provinces, and the Orange Zone (areas with moderately low production), which consists of 18 provinces.

Keywords: Mapping; Tomato; Clustering; Data Mining; K-Means

1. PENDAHULUAN

Hampir seluruh masyarakat Indonesia mengetahui tanaman tomat, karena selain dikenal sebagai sayuran, tomat juga memiliki peranan penting sebagai sumber gizi bagi masyarakat [1]. Buah tomat banyak mengandung zat-zat yang berguna bagi tubuh manusia antara lain mengandung vitamin C, vitamin A (karotien) dan mineral [2]. Tomat juga sebagai salah satu jenis sayuran yang termasuk dalam komoditas tanaman hortikultura penting, baik karena harganya yang cukup murah, nilai gizi nya juga cukup tinggi apabila di konsumsi masyarakat [3]. Citra rasa buah tomat sangat khas dan digemari hampir seluruh masyarakat karena vitamin yang terkandung didalamnya dan sering dimanfaatkan sebagai campuran masakan. Budidaya tomat banyak dilakukan oleh petani petani karena nilai ekonomi yang dimiliki cukup tinggi dan meningkatnya kebutuhan terhadap tomat terus meningkat sejalan dengan pertumbuhan penduduk yang pesat, pendapatan penduduk yang meningkat serta kesadaran masyarakat yang semakin tinggi bahwa nilai gizi sangat penting bagi kesehatan. Semakin tinggi produktivitas tomat dan harga penjualan di pasaran juga tinggi, tentu semakin membuat keuntungan yang besar. Oleh sebab itu maka budidaya tomat memiliki peluang usaha yang baik bagi petani sehingga banyak dikatakan bahwa budidaya tanaman tomat termasuk budidaya pertanian yang memiliki prospek yang cerah [4]. Komoditas tomat disebut sebagai tanaman multiguna, karena selain bermanfaat sebagai buah maupun sayuran, tomat juga bermanfaat untuk campuran obat-obatan serta bahan kosmetik [5].

Begitu penting nya informasi tentang produksi tomat di Indonesia yang tersebar di 34 provinsi, sehingga perlu dilakukan pemetaan dan pengelompokkan terhadap hal tersebut, agar pemerintah Indonesia khususnya pemerintah masing-masing provinsi memperoleh informasi tentang provinsi mana saja di Indonesia yang menghasilkan produksi tomat tertinggi dan terendah agar masing-masing pemerintah dapat membuat kebijakan maupun langkah-langkah strategis agar produksi tomat dapat ditingkatkan. Algoritma pemetaan dan

pengelompokan yang diusulkan pada penelitian ini adalah algoritma K-Means Clustering yang merupakan salah satu teknik dari algoritma Data Mining. Sebagaimana diketahui bahwa Algoritma Data Mining telah banyak digunakan untuk pemecahan masalah komputasi seperti yang berhubungan dengan klasifikasi [6]–[9], hingga Pengelompokan [10]–[14]. Algoritma K-Means digunakan pada penelitian ini karena sederhana dan cepat untuk memecahkan permasalahan cluster khususnya untuk data numerik dan sangat fleksibel serta efisien untuk ukuran data yang cukup besar dan tersebar [15].

Penelitian-penelitian terkait yang menjadi rujukan penelitian ini, diantaranya: Penelitian untuk memetakan tingkat kriminalitas di kabupaten Karawang menggunakan algoritma K-Means. Hasil pengelompokan daerah rawan kriminalitas di Karawang pada 2019 didapatkan cluster tidak rawan sebanyak 23 kecamatan, cluster rawan sebanyak 3 kecamatan dan cluster sangat rawan sebanyak 4 kecamatan. Sedangkan pada 2020 didapatkan cluster tidak rawan sebanyak 22 kecamatan, cluster rawan sebanyak 4 kecamatan, dan cluster sangat rawan sebanyak 4 kecamatan. Hasil evaluasi clustering menggunakan silhouette coefficient pada tahun 2019 yaitu sebesar 0,52 dan 0,54 pada tahun 2020, keduanya masuk dalam kategori medium *structure* dengan interpretasi penempatan cluster yang wajar [16]. Selanjutnya dilakukan penelitian untuk mengelompokkan data ekspor buah-buahan berdasarkan negara tujuan menggunakan algoritma K-Means. Penelitian ini mengelompokkan data ekspor kedalam 3 cluster yaitu tinggi, sedang dan rendah. Hasil Centroid data untuk cluster tingkat ekspor tinggi sebesar 2054519.3, centroid data untuk cluster tingkat ekspor sedang sebesar 489020.3 dan centroid data untuk cluster tingkat ekspor rendah sebesar 20.2. Sehingga diperoleh hasil cluster negara tujuan ekspor yaitu 2 negara cluster tingkat ekspor tinggi yakni negara Tiongkok & Malaysia, 2 negara cluster tingkat ekspor rendah yakni Vietnam & Thailand, dan 6 negara cluster tingkat ekspor rendah yakni Hongkong, Singapura, Nigeria, India, Jepang, Uni Emirat Arab [17]. Penelitian berikutnya dilakukan untuk memetakan tambak garam serta produksinya pada Kabupaten Pamekasan. Algoritma yang digunakan sama dengan dua penelitian yang lain, yakni algoritma K-Means Clustering. Penelitian ini menggunakan 3 (tiga) cluster yaitu tidak perlu meningkatkan produksi sebagai cluster 1, cukup perlu meningkatkan produksi sebagai cluster 2, dan perlu meningkatkan produksi sebagai cluster 3. Berdasarkan 19 sampel data maka didapat hasil 1 data berada pada cluster 1 (5%), 5 data berada pada cluster 2 (26%), dan 13 data ada pada cluster 3 (69%), dan tingkat akurasi perhitungannya adalah 89,47% [18].

Harapan dari penelitian ini mampu berkontribusi bagi pemerintah RI dan bagi bidang pertanian, khususnya bagi masing-masing pemerintah Provinsi sebagai referensi dan informasi mengenai pemetaan dan pengelompokan produksi tomat pada 34 provinsi yang ada di Indonesia. Selain itu bagi dunia pendidikan, hal ini diharapkan bermanfaat sebagai cara lain yang lebih baik dalam penyelesaian kasus pemetaan dan pengelompokan serta dapat digunakan oleh pihak akademisi untuk keilmuan dan pengembangan lebih lanjut.

2. METODOLOGI PENELITIAN

2.1 Algoritma dan Data Penelitian

Penelitian ini menggunakan Algoritma K-Means yang merupakan salah satu algoritma Data Mining yang digunakan untuk mengklaster dan melakukan pemetaan data produksi tomat. Sedangkan Data penelitian yang digunakan dalam makalah ini adalah data produksi tomat di Indonesia berdasarkan provinsi dalam kurun waktu lima tahun terakhir (tahun 2017- tahun 2021) yang diperoleh dari Dinas Pertanian Kabupaten/Kota masing-masing provinsi dan Badan Pusat Statistik Indonesia (dapat dilihat pada Tabel 1).

Tabel 1. Produksi Tanaman Tomat di Indonesia

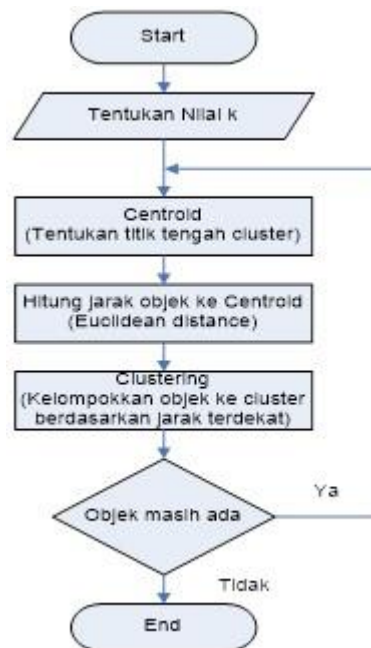
No	Provinsi	Produksi Tanaman Tomat (Ton)				
		Tahun 2017	Tahun 2018	Tahun 2019	Tahun 2020	Tahun 2021
1	Aceh	26.136	19.682	20.821	20.781	11.706
2	Sumatera Utara	97.358	103.650	118.583	162.744	203.162
3	Sumatera Barat	101.292	131.819	146.829	113.491	97.271
4	Riau	293	241	117	158	151
5	Jambi	24.450	11.621	12.348	19.652	23.890
6	Sumatera Selatan	13.559	14.050	12.487	10.620	8.836
7	Bengkulu	18.545	18.284	18.485	23.033	30.868
8	Lampung	25.432	19.604	18.669	19.096	15.934
9	Kepulauan Bangka Belitung	704	685	473	536	1.147
10	Kepulauan Riau	2	6	285	235	425
11	DKI Jakarta	0	0	0	0	0
12	Jawa Barat	295.321	268.448	284.948	299.267	292.309
13	Jawa Tengah	71.772	90.404	81.710	79.832	77.297
14	DI Yogyakarta	871	821	1.372	1.531	949
15	Jawa Timur	66.758	65.585	74.558	83.920	93.121

No	Provinsi	Produksi Tanaman Tomat (Ton)				
		Tahun 2017	Tahun 2018	Tahun 2019	Tahun 2020	Tahun 2021
16	Banten	1.017	784	830	1.894	1.190
17	Bali	24.520	25.602	15.171	13.811	12.172
18	Nusa Tenggara Barat	22.970	20.871	29.215	28.609	28.514
19	Nusa Tenggara Timur	6.716	5.466	9.950	9.907	10.605
20	Kalimantan Barat	1.805	1.128	2.088	1.857	4.108
21	Kalimantan Tengah	1.909	1.831	2.410	4.352	3.505
22	Kalimantan Selatan	9.153	7.813	7.399	7.409	5.067
23	Kalimantan Timur	6.429	7.151	7.430	8.210	9.101
24	Kalimantan Utara	2.754	2.843	2.688	2.367	6.008
25	Sulawesi Utara	30.276	53.075	42.392	57.331	66.711
26	Sulawesi Tengah	22.490	16.163	16.516	26.706	17.634
27	Sulawesi Selatan	64.917	67.373	58.513	60.435	63.373
28	Sulawesi Tenggara	3.464	3.792	5.608	4.720	4.800
29	Gorontalo	2.574	3.148	3.543	2.721	2.808
30	Sulawesi Barat	1.924	1.378	1.662	760	1.483
31	Maluku	2.761	3.549	3.324	4.110	4.024
32	Maluku Utara	1.834	877	7.548	6.785	5.449
33	Papua Barat	525	826	2.511	1.311	2.214
34	Papua	12.314	8.220	9.850	6.804	8.569
Jumlah		962.845	976.790	1.020.333	1.084.995	1.114.401

Pada tabel 1 dapat dilihat, bahwa DKI Jakarta tidak menghasilkan produksi tanaman tomat sama sekali (0). Sehingga nanti nya data DKI Jakarta tidak akan ikut di proses. Hal ini disebut juga dengan *Cleaning* yang berarti membersihkan data penelitian dari data-data yang dapat mengganggu hasil penelitian.

2.2 Diagram Alir dengan Algoritma K-Means

K-Means merupakan salah satu metode data Clustering non-hierarchical atau Partitional Clustering. Algoritma K-Means berusaha mengelompokkan data yang ada ke dalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama antara satu dengan yang lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain. Alur diagram dari algoritma K-Means tersaji pada gambar 1.



Gambar 1. Diagram Alir Algoritma K-Means [19]

- Penentuan jumlah *cluster* (*k*) pada data set.
- Pentuan nilai *centroid* (nilai pusat).
- Penentuan awal nilai pusat dilakukan secara acak, sedangkan rumus (1) digunakan pada step iterasi seperti berikut ini.

$$V_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj} \tag{1}$$

Keterangan :

V_{ij} = *Centroid* rata-rata *cluster* ke-I untuk variabel ke-j

N_i = Jumlah anggota *cluster* ke-i

i, k = Indeks dari *cluster*

j = Indeks dari variabel

X_{kj} = nilai data ke-k variabel ke-j untuk *cluster* tersebut

d. Pada tiap-tiap *record*, hitung jarak terdekat dengan *Centroid*.

Beberapa teknik yang bisa dilakukan untuk menghitung jarak data ke kelompok pusat, diantaranya *Minkowsky*, *Manhattan/City Block*, dan *Euclidean*. Masing-masing teknik mempunyai kekurangan dan kelebihan tersendiri. Pada penulisan makalah ini, jarak nilai pusat yang digunakan *Euclidean Distance*, dengan rumus:

$$De = \sqrt{(xi - si)^2 + (yi - ti)^2} \quad (2)$$

Keterangan :

De = *Euclidean Distance*

i = Jumlah objek ²

(x, y) = objek Koordinat

(s, t) = *Centroid* Koordinat

e. Objek dikelompokkan berdasarkan jarak dari *Centroid* yang terdekat

f. Tahap ke-3 hingga ke-4 diulangi, selanjutnya lakukan perulangan hingga nilai *Centroid* optimal.

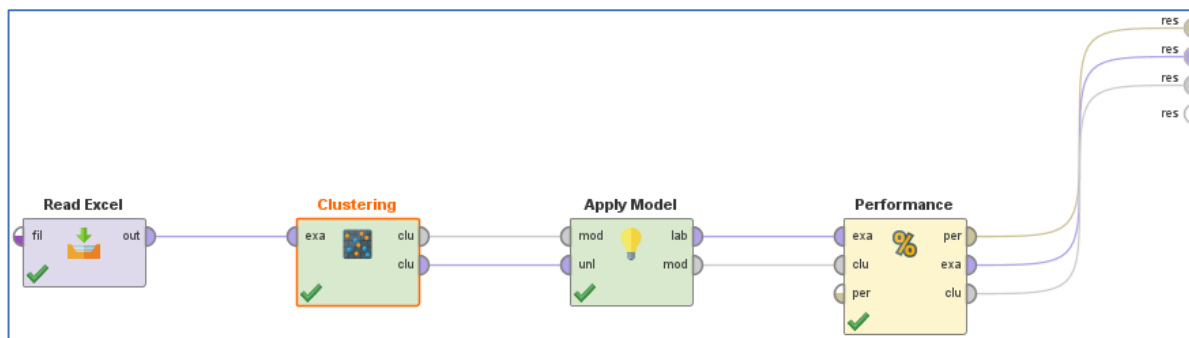
3. ANALISA DAN PEMBAHASAN

3.1 Penentuan Cluster

Pada penerapan algoritma K-Means menggunakan Rapidminer, terlebih dahulu adalah menentukan nilai K (Jumlah minimal cluster yang harus dideteksi). Cluster yang ditentukan pada makalah ini sebanyak 5 cluster, antara lain: cluster sangat tinggi, cluster tinggi, cluster sedang, cluster rendah dan cluster sangat rendah.

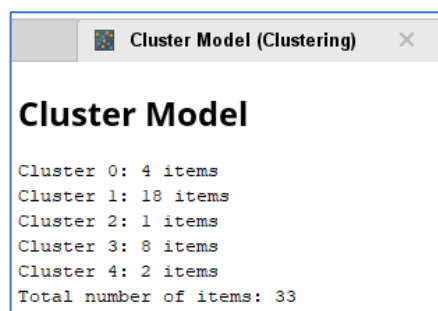
3.2 Proses K-Means dengan RapidMiner

Proses pemetaan produksi tomat dalam bentuk klasterisasi dan hasil dari algoritma K-Means dengan menggunakan RapidMiner dapat dilihat pada gambar 2, gambar 3, gambar 4, gambar 5 dan gambar 6.



Gambar 2. Process K-Means (k = 4 dan max runs = 10)

Gambar 2 menjelaskan proses pengklasteran K-Means dengan RapidMiner yang diawali dengan memasukkan data Produksi tomat di Indonesia, dilanjutkan pemilihan operator K-Means dengan nilai k = 5 dan nilai max runs = 10. Selanjutnya dihubungkan ke Apply Model untuk menerapkan model yang sudah dilatih atau dipelajari. Selanjutnya menghubungkan ke Cluster Distance Performance untuk mengevaluasi kinerja model yang memberikan daftar nilai kriteria kinerja secara otomatis sesuai dengan tugas yang diberikan.



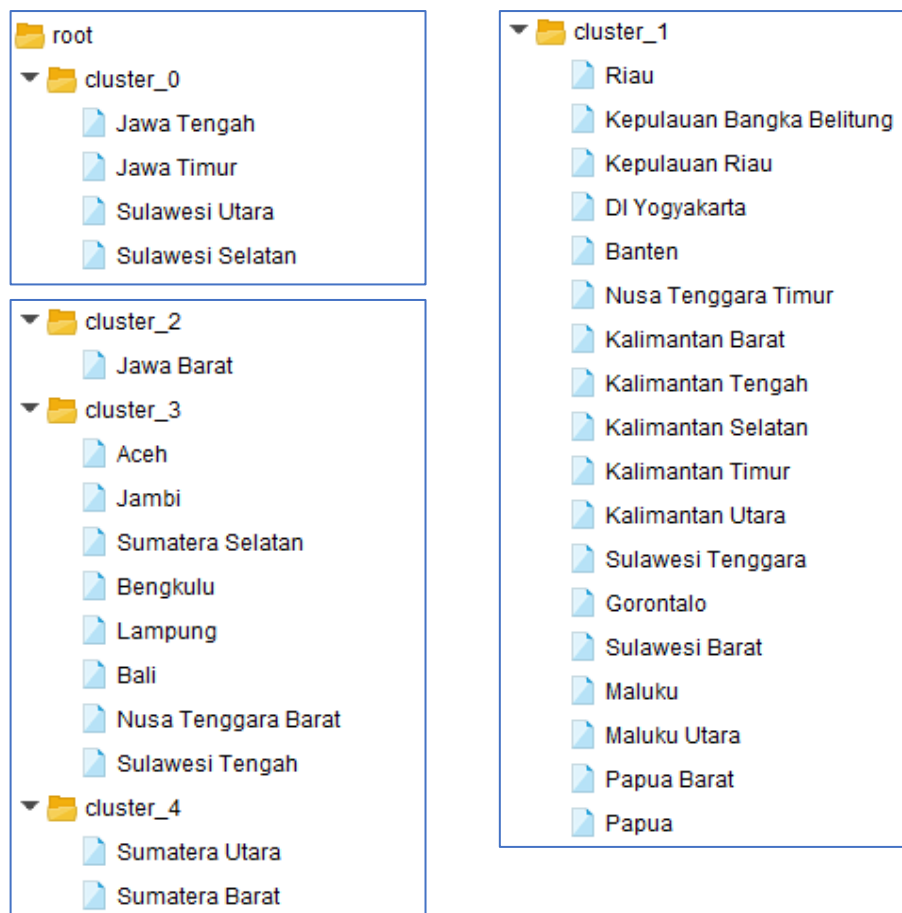
Gambar 3. Cluster (K-Means)

Gambar 3 merupakan cluster model yang dihasilkan oleh RapidMiner, Cluster 0: 4 items, Cluster 1: 18 items, Cluster 2: 1 items, Cluster 3: 8 items dan Cluster 4: 2 items. Hasil akhir Centroid table dapat dilihat pada gambar 4.

Attribute	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4
2017.0	58430.750	3169.389	295321	22262.750	99325
2018.0	69109.250	2808.833	268448	18234.625	117734.500
2019.0	64293.250	3838.222	284948	17964	132706
2020.0	70379.500	3648.167	299267	20288.500	138117.500
2021.0	75125.500	3977.944	292309	18694.250	150216.500

Gambar 4. Tabel Centroid yang Dihasilkan dengan RapidMiner

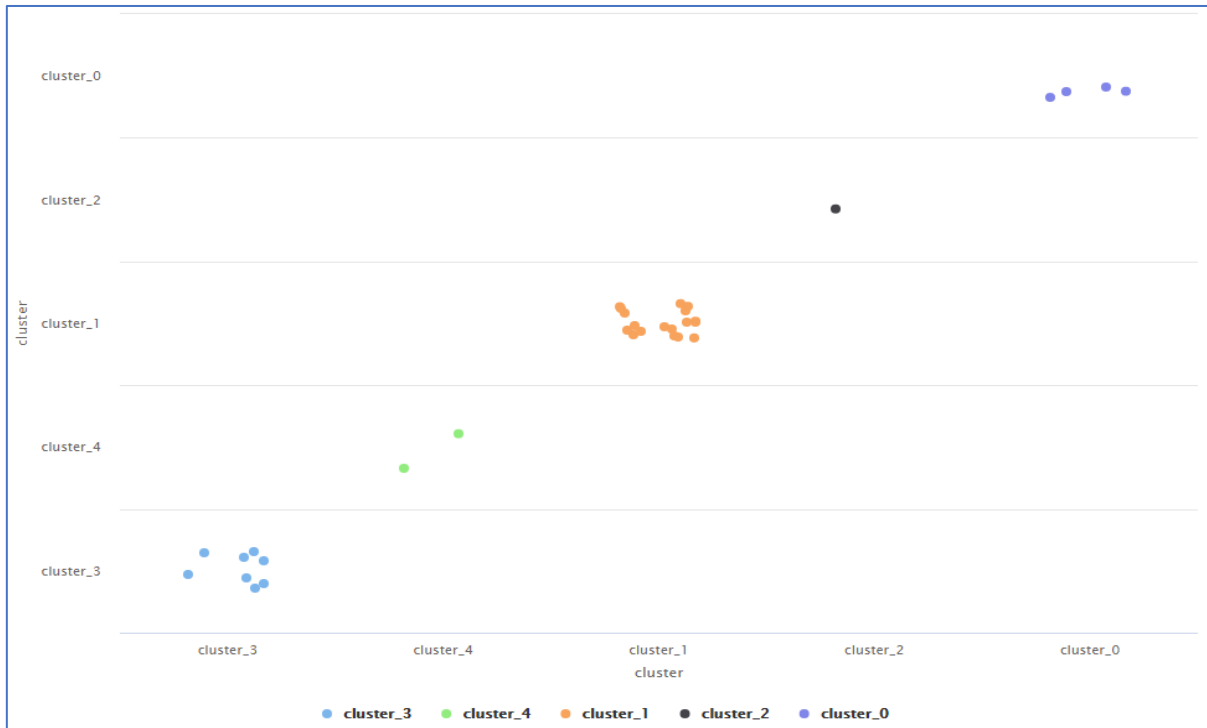
Hasil detail pemetaan produksi tanaman tomat di Indonesia berdasarkan 34 provinsi dapat dilihat pada gambar 5 berikut.



Gambar 5. Folder View Pengelompokan Produksi Tomat di Indonesia Berdasarkan Provinsi

Berdasarkan gambar 5 dapat dilihat hasil pengelompokan produksi tanaman tomat di Indonesia berdasarkan 34 provinsi, dikurangi provinsi DKI Jakarta yang tidak dimasukkan dalam data penelitian karena tidak menghasilkan produksi tomat sedikitpun.

Cluster_0 berdasarkan gambar 5 terdiri dari 4 (empat) provinsi yakni: Jawa Tengah, Jawa Timur, Sulawesi Utara dan Sulawesi Selatan. Cluster_1 terdiri dari 18 (delapan belas) provinsi yakni: Riau, Kepulauan Bangka Belitung, Kepulauan Riau, DI Yogyakarta, Banten, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat dan Papua. Cluster_2 hanya terdiri dari 1 (satu) provinsi, yakni Jawa Barat. Cluster_3 terdiri dari 8 (delapan) provinsi (Aceh, Jambi, Sumatera Selatan, Bengkulu, Lampung, Bali, Nusa Tenggara Barat dan Sulawesi Tengah). Sedangkan Cluster_4 terdiri dari 2 (dua) provinsi, yakni Sumatera Utara dan Sumatera Barat.



Gambar 6. Pemetaan Produksi Tanaman Tomat di Indonesia Berdasarkan Provinsi

Gambar 6 merupakan hasil pemetaan dari produksi tanaman tomat di Indonesia dari 34 provinsi. Berdasarkan hasil analisis, pemetaan produksi tanaman tomat di Indonesia terbagi menjadi 5 (lima) zona, antara lain: (1) Titik warna hitam merupakan pemetaan produksi tomat sangat tinggi di Indonesia (Cluster_2) yang masuk dalam Zona Hitam. (2) Titik warna hijau merupakan pemetaan produksi tomat tinggi di Indonesia (Cluster_4) yang masuk dalam Zona Hijau. (3) Titik warna biru merupakan pemetaan produksi tomat sedang di Indonesia (Cluster_0) yang masuk dalam Zona Biru. (4) Titik warna biru muda merupakan pemetaan produksi tomat rendah di Indonesia (Cluster_3) yang masuk dalam Zona Biru Muda. (5) Terakhir titik warna oranye merupakan pemetaan produksi tomat sangat rendah di Indonesia (Cluster_1) yang masuk dalam Zona Oranye.

4. KESIMPULAN

Data Mining K-Means dapat dimanfaatkan untuk melakukan pemetaan produksi tanaman tomat di Indonesia yang tersebar di 34 Provinsi. Berdasarkan hasil analisa pemetaan produksi tanaman tomat dengan menggunakan RapidMiner terbagi menjadi 5 (lima) zona, antara lain Zona Hitam (daerah yang memiliki produksi tomat sangat tinggi) yang terdiri dari 1 provinsi (Jawa barat). Zona Hijau (daerah yang memiliki produksi tinggi) yang terdiri dari 2 provinsi (Sumatera Utara dan Sumatera Barat). Zona Biru (daerah yang memiliki produksi sedang) yang terdiri dari 4 provinsi (Jawa Tengah, Jawa Timur, Sulawesi Utara dan Sulawesi Selatan). Zona Biru Muda (daerah yang memiliki produksi rendah) yang terdiri dari 8 provinsi (Aceh, Jambi, Sumatera Selatan, Bengkulu, Lampung, Bali, Nusa Tenggara Barat dan Sulawesi Tengah), dan Zona Oranye (daerah yang memiliki produksi cukup rendah) yang terdiri dari 18 provinsi (Riau, Kepulauan Bangka Belitung, Kepulauan Riau, DI Yogyakarta, Banten, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat dan Papua).

REFERENCES

- [1] W. Wardana and W. O. Alzarliani, "Faktor-Faktor yang Mempengaruhi Minat Petani Menerapkan Teknologi Pengolahan Buah Tomat di Desa Wakuli Kecamatan Kapontori Kabupaten Buton," *Agrikan: Jurnal Agribisnis Perikanan*, vol. 12, no. 1, pp. 145–151, 2019.
- [2] F. Fitriani, Y. Evayanti, and N. Isnaini, "Pemberian Jus Tomat Terhadap Kadar Hemoglobin Pada Ibu Hamil Trimester Iii Tahun 2019," *Jurnal Kebidanan Malahayati*, vol. 6, no. 2, pp. 230–235, 2020.
- [3] Chitra Anggriani Salingkat, A. Noviyanty, and Syamsiar, "Pengaruh Jenis Bahan Pengemas, Suhu dan Lama Penyimpanan terhadap Karakteristik Mutu Buah Tomat," *Agroland: Jurnal Ilmu-ilmu Pertanian*, vol. 27, no. 3, pp. 274–286, 2020.
- [4] D. Septiadi and A. I. Mundiayah, "Karakteristik dan Analisis Finansial Usaha Tani Tomat di Kabupaten Lombok Timur," *Agroteksos*, vol. 31, no. 3, pp. 180–188, 2021.
- [5] S. A. Assagaf, "Pengaruh Pemberian Mulsa Alang-Alang dan Pupuk NPK Phonska Terhadap Pertumbuhan dan

- Produksi Tanaman Tomat (*Solanum lycopersicum*),” *Jurnal Biosainstek*, vol. 2, no. 1, pp. 40–46, 2020.
- [6] T. H. Sinaga, A. Wanto, I. Gunawan, S. Sumarno, and Z. M. Nasution, “Implementation of Data Mining Using C4.5 Algorithm on Customer Satisfaction in Tirta Lihou PDAM,” *Journal of Computer Networks, Architecture, and High-Performance Computing*, vol. 3, no. 1, pp. 9–20, 2021.
- [7] A. Pradipta, D. Hartama, A. Wanto, S. Saifullah, and J. Jalaluddin, “The Application of Data Mining in Determining Timely Graduation Using the C45 Algorithm,” *IJISTECH (International Journal of Information System & Technology)*, vol. 3, no. 1, pp. 31–36, 2019.
- [8] I. Parlina *et al.*, “Naive Bayes Algorithm Analysis to Determine the Percentage Level of visitors the Most Dominant Zoo Visit by Age Category,” *Journal of Physics: Conference Series*, vol. 1255, no. 1, pp. 1–5, 2019.
- [9] S. F. Damanik, A. Wanto, and I. Gunawan, “Penerapan Algoritma Decision Tree C4.5 untuk Klasifikasi Tingkat Kesejahteraan Keluarga pada Desa Tiga Dolok,” *Jurnal Krisnadana Volume*, vol. 1, no. 2, pp. 21–32, 2022.
- [10] N. Arminarahmah, A. D. GS, G. W. Bhawika, M. P. Dewi, and A. Wanto, “Mapping the Spread of Covid-19 in Asia Using Data Mining X-Means Algorithms,” *IOP Conf. Series: Materials Science and Engineering*, vol. 1071, no. 1, p. 012018, 2021.
- [11] F. S. Napitupulu, I. S. Damanik, I. S. Saragih, and A. Wanto, “Algoritma K-Means untuk Pengelompokan Dokumen Akta Kelahiran pada Tiap Kecamatan di Kabupaten Simalungun,” *Building of Informatics, Technology and Science (BITS) Volume*, vol. 2, no. 1, pp. 55–63, 2020.
- [12] J. Hutagalung, N. L. W. S. R. Ginantra, G. W. Bhawika, W. G. S. Parwita, A. Wanto, and P. D. Panjaitan, “COVID-19 Cases and Deaths in Southeast Asia Clustering using K-Means Algorithm,” *Journal of Physics: Conference Series*, vol. 1783, no. 1, p. 012027, 2021.
- [13] N. A. Febriyati, A. D. GS, and A. Wanto, “GRDP Growth Rate Clustering in Surabaya City uses the K- Means Algorithm,” *International Journal of Information System & Technology*, vol. 3, no. 2, pp. 276–283, 2020.
- [14] M. A. Hanafiah and A. Wanto, “Implementation of Data Mining Algorithms for Grouping Poverty Lines by District/City in North Sumatra,” *International Journal of Information System & Technology*, vol. 3, no. 2, pp. 315–322, 2020.
- [15] C. Kamila, M. Adiyatma, G. R. Namang, R. Ramadhan, F. Syah, and D. Redaksi, “Pendidikan Teknik Informatika dan Komputer, Fakultas Teknik,” *INTECH - Informatika Dan Teknologi*, vol. 2, no. 1, pp. 19–24, 2021.
- [16] R. N. Fahmi, M. Jajuli, and N. Sulistiyowati, “Analisis Pemetaan Tingkat Kriminalitas di Kabupaten Karawang menggunakan Algoritma K-Means,” *INTECOMS: Journal of Information Technology and Computer Science*, vol. 4, no. 1, pp. 67–79, 2021.
- [17] H. A. Negara, A. R. Putra, and U. Enri, “Clustering Data Ekspor Buah-Buahan Berdasarkan Negara Tujuan Menggunakan Algoritma K-Means,” *Bina Insani Ict Journal*, vol. 8, no. 1, pp. 73–82, 2021.
- [18] W. Maulidi Molyono, S. Achmadi, and Y. Agus Pranoto, “Pemetaan Tambak Garam Serta Produksi Garam Pada Kabupaten Pamekasan Menggunakan K-Means Clustering,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 5, no. 2, pp. 794–799, 2021.
- [19] A. Wanto *et al.*, *Data Mining : Algoritma dan Implementasi*. Yayasan Kita Menulis, 2020.