

Available online at: <http://inventory.poltekappdg.ac.id/>

INVENTORY

Industrial Vocational E-Journal on Agroindustry

| ISSN Online 2723-1895 |



Segmentasi Perkebunan Kelapa Sawit dengan *Data Mining* Teknik *K-Means Clustering* Berdasarkan Luas Areal, Produksi dan Produktivitas

Trisna Yuniarti ¹, Dahliyah Hayati ¹

¹ Politeknik APP Jakarta, Jalan Timbul No.34, Jakarta Selatan, 12630, Indonesia

ARTICLE INFORMATION

Received: October 22, 2021

Revised: December 24, 2021

Available online: December 31, 2021

KEYWORDS

Centroid, Clustering, Data Mining, K-Means, Oil Palm

CORRESPONDENCE

Name: Trisna Yuniarti

E-mail: trisna.yuniarti@poltekapp.ac.id

A B S T R A C T

The oil palm is the most productive plantation product in Indonesia. Government strategies and policies related to oil palm plantations continue to be carried out considering that the plantation area is increasing every year. Segmentation of oil palm plantations based on area, production, and productivity aims to identify groups of potential oil palm plantations in the territory of Indonesia. This segmentation can provide consideration in formulating strategies and policies that will be made by the government. The segmentation method for grouping oil palm plantations uses the K-Means Clustering Data Mining technique with 3 clusters specified. Data mining stages start from data collection until representation is carried out, where 34 data sets are collected, only 25 data sets can be processed further. The results of this grouping obtained three plantation segments, namely 72% of the plantation group with low potential, 20% of the plantation group with medium potential, and 8% of the plantation group with high potential.

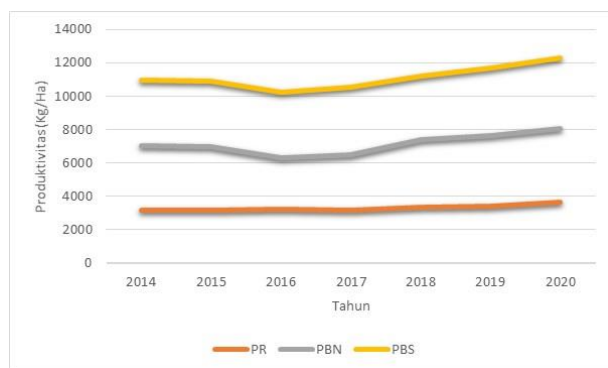
PENDAHULUAN

Indonesia merupakan negara agraris yang banyak menghasilkan komoditas perkebunan. Hasil komoditas tersebut memiliki kontribusi yang sangat tinggi bagi perekonomian negara. Salah satu barang dagangan ekspor yang bernilai tinggi dihasilkan dari perkebunan kelapa sawit. Berdasarkan status perusahaan, jenis perkebunan kelapa sawit dilakukan oleh Perkebunan Besar Negara (PBN), Perkebunan Besar Swasta (PBS), dan Perkebunan Rakyat (PR). Perkebunan tersebut hampir banyak ditemukan di Indonesia, terutama di Pulau Sumatera dan Kalimantan. Produk yang dihasilkan dari perkebunan kelapa sawit berupa minyak kelapa sawit atau sering disebut CPO dan minyak inti kelapa sawit atau sering disebut PKO. Produk tersebut dapat diolah kembali untuk keperluan beberapa industri makanan maupun non makanan, bahkan sebagai sumber energi terbarukan [1].

Berdasarkan hasil studi literatur diketahui bahwa dari beberapa periode lahan dan produksi kelapa sawit mengalami perkembangan yang signifikan. Kegiatan

pengembangan perkebunan kelapa sawit memiliki pengaruh bagi perekonomian masyarakat, yaitu terciptanya lapangan kerja dan lapangan kerja bagi masyarakat setempat. Selain berdampak positif, perluasan perkebunan ini juga memiliki dampak negatif, yaitu adanya perluasan dan perusahaan perkebunan tanpa legalitas yang pada akhirnya merusak keanekaragaman ekosistem serta menyebabkan deforestasi [2]. Berdasarkan data lima tahun Direktorat Jenderal Perkebunan, yaitu tahun 2014 sampai dengan 2018, luas areal meningkat dengan rata-rata pertumbuhan sebesar 7,89%. Areal perkebunan diprediksi akan terus naik seiring dengan perkembangan industri dan pemenuhan minyak nabati di dunia. Gambar 1. memperlihatkan produktivitas perkebunan kelapa sawit mengalami kenaikan rata-rata sebesar 0,37% dari tahun 2014 sampai dengan tahun 2020 [3]. Produktivitas tersebut diperlihatkan dari hasil PBN, PBS, dan PR. Indonesia merupakan negara pengekspor minyak kelapa sawit tertinggi ke berbagai negara di dunia. Pada saat ini, pemerintah berusaha untuk meningkatkan produktivitas kelapa sawit guna memenuhi permintaan pelanggan dari dalam negeri maupun luar negeri. Adanya peningkatan

produktivitas diharapkan dapat meningkatkan pendapatan para petani.



Gambar 1. Produktivitas Kelapa Sawit Indonesia pada Tahun 2014 - 2020

Indonesia menjadi negara tujuan pengimpor kelapa sawit ke beberapa negara di dunia. Guna menyeimbangkan *supply* dan *demand* hasil produk kelapa sawit tersebut, pemerintah melakukan kajian proyeksi kebutuhan hasil perkebunan kelapa sawit untuk beberapa tahun ke depan. Dengan demikian, dapat dilakukan prediksi ketersediaannya tanpa mengabaikan faktor lingkungan sebagai dampak lain dari hasil pengolahan produksi kelapa sawit, walau secara ekonomi akan memberikan manfaat [4]. Pengembangan berupa perluasan diiringi dengan peningkatan produktivitas harus dilakukan secara bertanggung jawab. Oleh karena itu, Kementerian Pertanian menetapkan suatu pedoman Perkebunan Kelapa Sawit Berkelanjutan Indonesia atau *Indonesian Sustainable Palm Oil (ISPO)* guna setiap perusahaan perkebunan sawit mematuhi perundang-undangan yang tercantum pada pedoman tersebut.

Hampir seluruh wilayah di Indonesia terdapat perkebunan kelapa sawit, akan tetapi belum diketahui secara pasti besarnya potensi perkebunan tersebut dilihat dari segi produktivitas, luas areal dan jumlah produksi yang dimiliki pada masing-masing propinsi. Metode yang dapat digunakan untuk mengetahui potensi perkebunan kelapa sawit yang memiliki karakteristik yang sama berdasarkan luas areal, jumlah produksi, dan produktivitas adalah dengan menggunakan *Data Mining* teknik *K-Means Clustering*. Beberapa penelitian yang menggunakan metode ini telah dilakukan untuk pengelompokan data penjualan [5,6,7,8]. Penelitian lainnya menggunakan metode ini untuk melakukan pengelompokan distribusi suatu produk [9,10]. Selain itu, teknik *K-Means Clustering* dapat juga dikombinasikan dengan metode lainnya, seperti pada penelitian dalam memprediksi profil pemakaian beban panas skala besar di sektor komersial, industri, publik atau residensial yang fokus pada industri manufaktur [11]. Pada penggunaan teknik yang sama, penelitian

pengelompokan yang memfokuskan pada tanaman kelapa sawit dilakukan untuk pemetaan hasil produksi buah sawit, mengetahui blok tanaman kelapa sawit paling produktif dan tidak produktif, serta untuk mengetahui pengelompokan jenis tanaman kelapa sawit produktif sehingga dapat meningkatkan produksi dan kesejahteraan karyawan [12, 13, 14].

Clustering merupakan salah satu teknik untuk melakukan pengelompokan data dengan membagi data sesuai dengan kelompok objek yang sama, dimana objek tersebut memiliki kemiripan satu sama lain [15]. Adanya perkembangan teknologi dan internet memberikan kemudahan akses dalam pengumpulan data. Data tersebut dapat dimanfaatkan lebih dalam untuk memperoleh pengetahuan atau informasi yang baru bagi penggunaannya. Pada *data mining*, teknik *clustering* yang sering digunakan adalah *K-Means*, yaitu algoritma pengelompokan yang berdasarkan pada kesamaan antara objek data statis, sederhana diimplementasikan, serta memiliki kompleksitas waktu linier yang rendah [16].

Penelitian ini bertujuan untuk melakukan pengelompokan perkebunan kelapa sawit di Indonesia berdasarkan data luas areal, produksi, dan produktivitas. Adanya pengelompokan ini diharapkan akan diketahui informasi lebih dalam dan dapat digunakan untuk membuat kebijakan atau strategi dalam peningkatan perkebunan kelapa sawit berkelanjutan di seluruh wilayah Indonesia. Pengelompokan pada perkebunan tersebut dilakukan dengan menerapkan metode *data mining* teknik *K-Means Clustering*. Penggunaan metode tersebut akan menghasilkan kelompok data berdasarkan data-data yang mirip dan memiliki karakteristik yang sama.

METODOLOGI

Metode *data mining* dengan teknik *K-Means Clustering* digunakan pada penelitian segmentasi perkebunan kelapa sawit di Indonesia. Teknik ini digunakan untuk mendapatkan hasil pengelompokan atau segmentasi perkebunan kelapa sawit dari berbagai propinsi yang terdapat di Indonesia. Pendekatan yang diusulkan terdiri dari beberapa langkah untuk menghasilkan segmentasi yang terdiri dari: pengumpulan data, pengolahan data, dan analisis data.

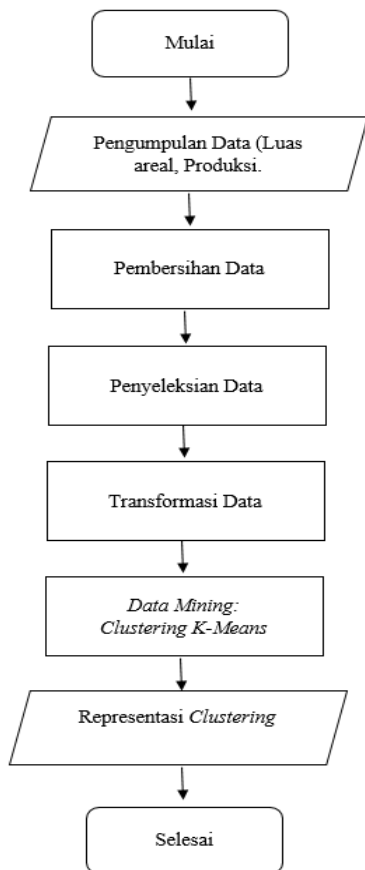
Proses Pengumpulan Data

Data yang digunakan untuk melakukan pengelompokan perkebunan kelapa sawit di Indonesia merupakan data sekunder, yaitu: data luas areal, data produksi, dan data produktivitas. Data tersebut merupakan jenis data kuantitatif yang didapatkan dari hasil publikasi

Direktorat Jenderal Perkebunan, Kementerian Pertanian tahun 2021 melalui *website* <https://ditjenbun.pertanian.go.id/>. Data yang telah terkumpul berupa data luas areal, data produksi, dan data produktivitas pada perkebunan sawit yang terdapat di berbagai propinsi di Indonesia. Berdasarkan data tersebut diberikan variabel untuk setiap data yang digunakan, X_1 = data luas areal, X_2 = data produksi, dan X_3 = data produktivitas. Pada data luas areal pertanian dilakukan penjumlahan antara tanaman yang menghasilkan (TM) dan tanaman rusak (TR). Tanaman rusak tetap diperhitungkan karena secara di lapangan tanaman tersebut masih menghasilkan walau tidak memadai lagi.

Proses pengolahan data

Data luas areal, data produksi, dan data produktivitas selanjutnya diolah menggunakan metode *data mining* dengan tahapan: pembersihan data; pengintegrasian data; penyeleksian data; transformasi data; data mining; dan representasi pengetahuan yang didapatkan dari hasil pengolahan *data mining* tersebut [17]. Gambar 1. merupakan proses untuk setiap tahapan dalam melakukan pengolahan data. Proses komputasi dilakukan dengan menggunakan *software excel* dan *Rapid Miner* untuk mempermudah proses pengolahan data.



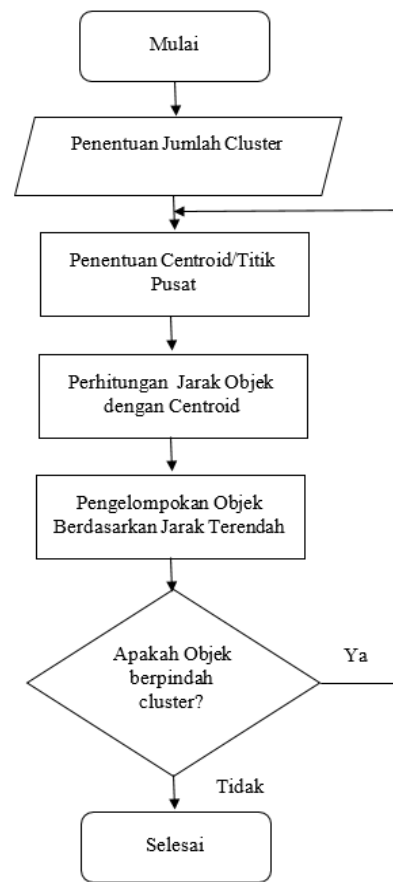
Gambar 2. Pengolahan Data Dengan *Data Mining K-Means*

Data mining dengan teknik *K-Means Clustering* akan dilakukan setelah data ditransformasikan. Pengelompokan data tersebut dilakukan dengan tahapan yang terdapat pada Gambar 2. Penentuan *centroid* dapat menggunakan rumus Euclidean sebagai berikut [18]:

$$\sqrt{\sum_{i=1}^n (X_i - Y_j)^2} \tag{1}$$

Keterangan:

- d = jarak antara suatu titik dengan titik pusat,
- X_i = nilai variabel ke i pada sampel X,
- Y_j = nilai variabel ke j dari pusat *clustering* Y.



Gambar 3. Alur Proses Segmentasi dengan *K-Means*

Segmentasi dengan teknik *K-Means Clustering* pada *data mining* terdiri dari beberapa tahapan. Alur setiap tahapan pada teknik ini dapat dilakukan seperti pada Gambar 3. Hasil komputasi dari pengolahan data menggunakan teknik ini akan menghasilkan segmentasi perkebunan kelapa sawit pada wilayah Indonesia menjadi beberapa kelompok berdasarkan hasil kemiripan dari data luas areal, data produksi, dan data produktivitas.

Analisis Data

Tahap analisis merupakan tahap terakhir dalam proses segmentasi perkebunan tersebut. Setelah proses pengolahan data dengan teknik *K-Means Clustering*

dilakukan, maka akan dilanjutkan dengan melakukan analisis lebih dalam guna mengetahui secara terperinci segmen perkebunan kelapa sawit yang terdapat diberbagai wilayah Indonesia berdasarkan karakteristik

yang sama dari data luas areal, data produksi, dan data produktivitas.

Tabel 1. Data Luas Areal dan Produksi Kelapa Sawit Menurut Provinsi dan Keadaan Tanaman Tahun 2020

Propinsi	Luas Areal TBM (Ha)	Luas Areal TM (Ha)	Luas Areal TR (Ha)	Produksi (Ton)	Produktivitas (Kg/Ha)
Aceh	97.261	360.302	30.440	1.134.606	3.149
Sumut	119.461	1.176.501	29.117	5.776.781	4.910
Sumbar	41.451	337.604	14.254	1.312.253	3.887
Riau	330.605	2.441.069	82.268	9.984.315	4.090
Kepri	2.079	5.140	178	20.020	3.895
Jambi	192.179	840.481	41.939	3.022.565	3.596
Sumsel	141.012	1.022.994	33.958	4.267.023	4.171
Kep. Babel	34.190	204.379	1.244	843.047	4.125
Bengkulu	52.991	268.190	4.069	1.063.404	3.965
Lampung	18.601	172.710	5.001	384.948	2.229
D.K.I. Jakarta	-	-	-	-	-
Jawa Barat	599	12.865	-	33.093	2.572
Banten	1.587	14.990	2.666	27.423	1.829
Jawa Tengah	-	-	-	-	-
D.I.	-	-	-	-	-
Yogyakarta	-	-	-	-	-
Jawa Timur	-	-	-	-	-
Bali	-	-	-	-	-
NTB	-	-	-	-	-
NTT	-	-	-	-	-
Kalbar	316.531	1.666.521	56.152	5.471.407	3.283
Kalteng	216.123	1.754.028	48.509	7.685.770	4.382
Kalsel	67.055	410.288	19.917	1.561.147	3.805
Kaltim	205.300	1.088.044	20.262	3.823.221	3.514
Kaltara	22.897	130.395	4.385	301.607	2.313
Sulawesi	-	-	-	-	-
Utara	-	-	-	-	-
Gorontalo	1.781	11.517	-	4.975	432
Sulteng	26.255	113.890	5.727	371.717	3.264
Sulsel	9.559	26.830	8.347	100.317	3.739
Sulbar	38.225	100.732	17.223	348.015	3.455
Sultra	11.030	93.695	5.576	76.300	814
Maluku	794	9.799	259	19.145	1.954
Maluku Utara	5.541	-	-	-	-
Papua	28.674	111.855	19.191	557.559	4.985
Papua Barat	5.126	45.892	-	106.413	2.319

(Sumber: Ditjenbun, 2021)

HASIL DAN PEMBAHASAN

Pengumpulan Data

Data didapatkan dari publikasi Kementerian Pertanian yang terdiri dari data luas areal dan produksi kelapa sawit pada perkebunan rakyat (PR), besar negara (PBN) dan besar swasta (PBS) menurut propinsi serta keadaan tanaman pada tahun 2020 [3]. Dataset perkebunan kelapa sawit untuk setiap propinsi yang terdapat di wilayah

Indonesia yang berjumlah 34 data yang diperlihatkan pada Tabel 1. Data tersebut terdiri dari data propinsi, luas areal, produksi, dan produktivitas.

Pembersihan Data

Data yang terdapat pada Tabel 1. menuju ke langkah selanjutnya, yaitu melakukan pembersihan data. Pembersihan data sangat penting dilakukan agar data yang akan diolah menjadi lebih valid hasilnya.

Pembersihan pada data mentah yang terdapat pada Tabel 1. diidentifikasi dan ditemukan terdapat data perkebunan kelapa sawit pada beberapa propinsi yang tidak memiliki nilai. Data yang tidak memiliki nilai tersebut, yaitu untuk Propinsi: D.K.I. Jakarta, Jawa Tengah, D.I. Yogyakarta, Jawa Timur, Bali, Nusa Tenggara Barat, Nusa Tenggara

Timur, Sulawesi Utara, dan Maluku Utara akan dihapus sebelum ke tahap penyeleksian data. Pembersihan data telah dilakukan dan menghasilkan data baru seperti yang terdapat pada Tabel 2. Data mengalami perubahan atau mengalami reduksi dimana pada awalnya data berjumlah 34 dataset menjadi 26 dataset.

Tabel 2. Tahapan Pembersihan Data

Propinsi	Luas Areal TBM (Ha)	Luas Areal TM (Ha)	Luas Areal TR (Ha)	Produksi (Ton)	Produktivitas (Kg/Ha)
Aceh	97.261	360.302	30.440	1.134.606	3.149
Sumut	119.461	1.176.501	29.117	5.776.781	4.910
Sumbang	41.451	337.604	14.254	1.312.253	3.887
Riau	330.605	2.441.069	82.268	9.984.315	4.090
Kepri	2.079	5.140	178	20.020	3.895
Jambi	192.179	840.481	41.939	3.022.565	3.596
Sumsel	141.012	1.022.994	33.958	4.267.023	4.171
Kep. Babel	34.190	204.379	1.244	843.047	4.125
Bengkulu	52.991	268.190	4.069	1.063.404	3.965
Lampung	18.601	172.710	5.001	384.948	2.229
Jawa Barat	599	12.865	-	33.093	2.572
Banten	1.587	14.990	2.666	27.423	1.829
Kalbar	316.531	1.666.521	56.152	5.471.407	3.283
Kalteng	216.123	1.754.028	48.509	7.685.770	4.382
Kalsel	67.055	410.288	19.917	1.561.147	3.805
Kaltim	205.300	1.088.044	20.262	3.823.221	3.514
Kaltara	22.897	130.395	4.385	301.607	2.313
Gorontalo	1.781	11.517	-	4.975	432
Sulteng	26.255	113.890	5.727	371.717	3.264
Sulsel	9.559	26.830	8.347	100.317	3.739
Sulbar	38.225	100.732	17.223	348.015	3.455
Sultra	11.030	93.695	5.576	76.300	814
Maluku	794	9.799	259	19.145	1.954
Maluku Utara	5.541	-	-	-	-
Papua	28.674	111.855	19.191	557.559	4.985
Papua Barat	5.126	45.892	-	106.413	2.319

Penyeleksian Data

Pembersihan data yang telah dilakukan akan diseleksi kembali, yaitu dilakukan pemilihan data yang tepat digunakan untuk proses segmentasi perkebunan kelapa sawit yang ditemukan di wilayah Indonesia. Hasil dari penyeleksian data yang terdapat pada Tabel 3. didapatkan bahwa bahwa data luas areal tanaman menghasilkan (TM), luas areal tanaman rusak (TR), produksi, dan produktivitas selanjutnya akan digunakan pada proses transformasi data. Pada data luas areal tanaman belum menghasilkan (TBM) tidak digunakan karena mengindikasikan bahwa pada areal tersebut kelapa sawit masih belum bereproduksi. Sedangkan, data tanaman rusak (TR) tetap dipakai mengingat tanaman tersebut terkadang masih menghasilkan kelapa sawit walaupun hasilnya tidak optimal.

Hasil seleksi menunjukkan bahwa Propinsi Maluku tereliminasi pada tahap ini. Hal ini disebabkan karena perkebunan kelapa sawit pada propinsi ini masih dalam

bentuk tanaman belum menghasilkan, tidak ada tanaman yang menghasilkan maupun tanaman rusak. Indikasi dari hal tersebut menyatakan bahwa pada Propinsi Maluku perkebunan sawit belum bereproduksi atau tidak ada produktivitasnya, sehingga dilakukan penghapusan untuk propinsi ini. Dengan demikian data semula yang berjumlah 26 dataset berubah menjadi 25 dataset.

Transformasi Data

Data yang dihasilkan pada tahap penyeleksian data akan ditransformasi agar menghasilkan data yang tepat saat digunakan pada proses *data mining k-means*. Transformasi data dilakukan dengan melakukan seleksi atribut, dimana Pada Tabel 3. diketahui bahwa data luas areal pada tanaman menghasilkan (TM) dan luas areal pada tanaman rusak (TR) masih dalam bentuk data terpisah. Kedua data tersebut akan diintegrasikan menjadi satu data sehingga menghasilkan data baru yang selanjutnya disebut data luas areal total. Pada tahap transformasi data akan dihasilkan tiga variabel, yaitu:

data luas areal total, data produksi, dan data produktivitas.

Hasil transformasi data dihasilkan sebanyak 25 dataset yang akan diproses menggunakan teknik *K-Means Clustering*. Sebelum melakukan eksekusi menggunakan teknik tersebut, setiap variabel diberikan atribut khusus untuk mempermudah proses pengolahannya. Pada Tabel

4. memperlihatkan nama propinsi diberikan atribut ID dengan memberi nomor id 1 sampai dengan 25. Lalu variabel lainnya seperti data luas areal total diberi atribut X_1 , data produksi diberi atribut X_2 dan data produktivitas diberi atribut X_3 . Data awal yang semula terdiri dari empat variabel mengalami perubahan bentuk menjadi tiga variabel.

Tabel 3. Penyeleksian Data

Propinsi	Luas Areal TM (Ha)	Luas Areal TR (Ha)	Produksi (Ton)	Produktivitas (Kg/Ha)
Aceh	360.302	30.440	1.134.606	3.149
Sumut	1.176.501	29.117	5.776.781	4.910
Sumbar	337.604	14.254	1.312.253	3.887
Riau	2.441.069	82.268	9.984.315	4.090
Kepri	5.140	178	20.020	3.895
Jambi	840.481	41.939	3.022.565	3.596
Sumsel	1.022.994	33.958	4.267.023	4.171
Kep. Babel	204.379	1.244	843.047	4.125
Bengkulu	268.190	4.069	1.063.404	3.965
Lampung	172.710	5.001	384.948	2.229
Jawa Barat	12.865	-	33.093	2.572
Banten	14.990	2.666	27.423	1.829
Kalbar	1.666.521	56.152	5.471.407	3.283
Kalteng	1.754.028	48.509	7.685.770	4.382
Kalsel	410.288	19.917	1.561.147	3.805
Kaltim	1.088.044	20.262	3.823.221	3.514
Kaltara	130.395	4.385	301.607	2.313
Gorontalo	11.517	-	4.975	432
Sulteng	113.890	5.727	371.717	3.264
Sulsel	26.830	8.347	100.317	3.739
Sulbar	100.732	17.223	348.015	3.455
Sultra	93.695	5.576	76.300	814
Maluku	9.799	259	19.145	1.954
Papua	111.855	19.191	557.559	4.985
Papua Barat	45.892	-	106.413	2.319

Clustering K-Means

Data yang telah mengalami transformasi diolah lebih lanjut pada proses *data mining* guna melakukan segmentasi perkebunan kelapa sawit di wilayah Indonesia. *Data mining* digunakan pada segmentasi ini adalah teknik *clustering k-means*, yaitu membagi perkebunan kelapa sawit menjadi beberapa segmen berdasarkan dari tiga variabel luas areal (X_1), produksi (X_2), dan produktivitasnya (X_3). Segmentasi dilakukan dengan menggunakan *software Rapid Miner*. Hal pertama yang dilakukan adalah membuat proses desain *k-means* dengan cara melakukan impor data yang terdapat pada Tabel 4. dengan memilih algoritma *k-means* dan *performance*. Segmentasi ditentukan sebanyak 3 kelompok atau $K=3$ yang memberikan informasi bahwa kelompok perkebunan terdiri dari kelompok perkebunan yang memiliki potensi hasil perkebunan kelapa sawit tinggi, kelompok perkebunan yang memiliki potensi hasil perkebunan kelapa sawit

yang sedang, dan kelompok perkebunan yang memiliki potensi hasil perkebunan kelapa sawit yang rendah.

Tabel 4. Transformasi Data

ID	X_1	X_2	X_3
1	390.742	1.134.606	3.149
2	1.205.618	5.776.781	4.910
3	351.858	1.312.253	3.887
4	2.523.337	9.984.315	4.090
5	5.318	20.020	3.895
6	882.420	3.022.565	3.596
7	1.056.952	4.267.023	4.171
8	205.623	843.047	4.125
9	272.259	1.063.404	3.965
10	177.711	384.948	2.229
11	12.865	33.093	2.572
12	17.656	27.423	1.829
13	1.722.673	5.471.407	3.283
14	1.802.537	7.685.770	4.382
15	430.205	1.561.147	3.805
16	1.108.306	3.823.221	3.514
17	134.780	301.607	2.313
18	11.517	4.975	432

ID	X ₁	X ₂	X ₃
19	119.617	371.717	3.264
20	35.177	100.317	3.739
21	117.955	348.015	3.455
22	99.271	76.300	814
23	10.058	19.145	1.954
24	131.046	557.559	4.985
25	45.892	106.413	2.319

Penentuan nilai *centroid* titik pusat didapatkan secara random dan otomatis dengan tipe pengukuran yang dipakai berupa *numerical measures euclidian distance*. Nilai *centroid* pada masing-masing segmentasi/cluster berdasarkan luas areal (X₁), produksi (X₂), dan produktivitas (X₃) yang digunakan sesuai dengan *centroid table* yang ditampilkan seperti pada Tabel 5.

Tabel 5. Data *Centroid* pada *Centroid Table*

Variabel	Cluster 0	Cluster 1	Cluster 2
X ₁	142.753,7	1.195.193,8	2.162.937
X ₂	459.221,6	4.472.199,4	8.835.042,5
X ₃	2.929,5	3.894,8	4.236,0

Tabel 6. merupakan hasil komputasi yang menampilkan nilai statistik maksimum, minimum, dan rata-rata untuk variabel X₁, X₂, dan X₃.

Tabel 6. Statistik Nilai Maksimum, Minimum, dan Rata-rata Setiap Variabel

Variabel	Minimum	Maksimum	Rata-rata
X ₁	5.318	2.523.337	514.855,72
X ₂	4.975	9.984.315	1.931.882,84
X ₃	432	4.985	3.227,08

Hasil proses *k-means clustering* didapatkan tiga model segmentasi seperti yang terdapat pada Tabel 7. Hasil segmentasi tersebut memberikan informasi bahwa segmen pertama atau *cluster 0* terdiri dari 18 item, segmen kedua atau *cluster 1* terdiri dari 2 item, dan segmen ketiga atau *cluster 2* terdiri dari 2 item.

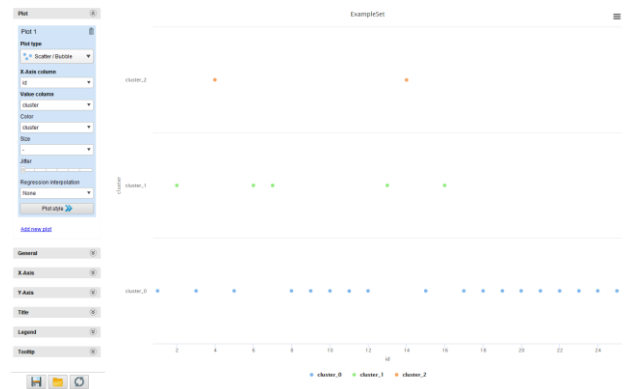
Tabel 7. Hasil Segmentasi Perkebunan Kelapa Sawit Tahun 2020 Menggunakan *K-Means Clustering*

Cluster	Jumlah
Cluster 0	18 items
Cluster 1	5 items
Cluster 2	2 items

Representasi Clustering K-Means

Hasil komputasi memperlihatkan jumlah item setiap segmentasi perkebunan kelapa sawit yang terdapat pada propinsi di wilayah Indonesia. Representasi *K-Means Clustering* dapat dilihat berdasarkan Gambar 4. Pada gambar tersebut diketahui bahwa segmentasi 1 atau *cluster 0* diperlihatkan dengan warna biru, yaitu merupakan kelompok perkebunan yang memiliki hasil

perkebunan kelapa sawit yang rendah. Segmentasi 2 atau *cluster 1* diperlihatkan dengan warna hijau, yaitu merupakan kelompok perkebunan yang memiliki hasil perkebunan kelapa sawit sedang. Segmentasi 3 atau *cluster 2* diperlihatkan dengan warna oranye, merupakan kelompok perkebunan yang memiliki hasil perkebunan kelapa sawit yang tinggi.



Gambar 4. Plot Hasil Segmentasi Perkebunan Kelapa Sawit di Indonesia

Analisis Data

Hasil pengolahan data menggunakan *K-Means Clustering* ditinjau dari tiga variabel luas areal (X₁), produksi (X₂), dan produktivitasnya (X₃) untuk perkebunan kelapa sawit yang terdapat di wilayah Indonesia menghasilkan tiga segmentasi/*cluster*. Segmentasi 1 atau *cluster 0* merupakan kelompok perkebunan yang memiliki potensi hasil perkebunan kelapa sawit yang rendah, terdiri dari id: 1, 3, 5, 8, 9, 10, 11, 12, 15, 17, 18, 19, 20, 21, 22, 23, 24, 25. Pada segmen pertama dapat diketahui bahwa luas areal total terletak pada rentang 5.318 Ha sampai dengan 430.205 Ha, produksi pada rentang 4.975 Ton sampai dengan 1.561.147 Ton , dan produktivitas pada rentang 432 Kg/Ha sampai dengan 4.985 Kg/Ha.

Segmentasi 2 atau *cluster 1* merupakan kelompok perkebunan yang memiliki potensi hasil perkebunan kelapa sawit sedang, terdiri dari id: 2, 6, 7, 13, 16. Pada segmen kedua dapat diketahui bahwa luas areal total terletak pada rentang 882.420 Ha sampai dengan 1.722.673 Ha, produksi pada rentang 3.022.565 Ton sampai dengan 5.776.781 Ton , dan produktivitas pada rentang 3.283 Kg/Ha sampai dengan 4.910 Kg/Ha.

Segmentasi 3 atau *cluster 2* merupakan kelompok perkebunan yang memiliki potensi hasil perkebunan kelapa sawit yang tinggi, terdiri dari id: 4 dan 14. Pada segmen kedua dapat diketahui bahwa luas areal total terletak pada rentang 1.802.537 Ha sampai dengan 2.523.337 Ha, produksi pada rentang 7.685.770 Ton sampai dengan 9.984.315 Ton , dan produktivitas pada rentang 4.090 Kg/Ha sampai dengan 4.382 Kg/Ha.

Hasil performa diperlihatkan dari hasil *Davies Bouldin Index* (DBI) yang memiliki nilai sebesar 0,145. Hal ini menunjukkan kinerja pada *centroid* 3 segmentasi/cluster memberikan hasil yang terbaik karena nilai DBI mendekati nilai 0. Perincian masing masing segmentasi/cluster dapat dilihat pada Tabel 8. yang memperlihatkan segmentasi untuk setiap propinsi. Segmentasi ini didapatkan dari kesamaan karakteristik perkebunan kelapa sawit suatu propinsi di wilayah Indonesia berdasarkan kemiripan dari data luas areal, data produksi, dan data produktivitas. Dari Tabel 8. Dapat diketahui bahwa dari 100% data didapatkan 72% provinsi yang memiliki karakteristik yang sama pada *cluster 0*, sebanyak 20% provinsi yang memiliki karakteristik yang sama pada *cluster 1*, dan 8% % provinsi yang memiliki karakteristik yang sama pada pada *cluster 2*.

Tabel 8. Hasil Segmentasi Perkebunan Kelapa Sawit di Wilayah Indonesia

Propinsi	Cluster 0	Cluster 1	Cluster 2
Aceh	√		
Sumut		√	
Sumbar	√		
Riau			√
Kepri	√		
Jambi		√	
Sumsel		√	
Kep. Babel	√		
Bengkulu	√		
Lampung	√		
Jawa Barat	√		
Banten	√		
Kalbar		√	
Kalteng			√
Kalsel	√		
Kaltim		√	
Kaltara	√		
Gorontalo	√		
Sulteng	√		
Sulsel	√		
Sulbar	√		
Sultra	√		
Maluku	√		
Papua	√		
Papua Barat	√		

KESIMPULAN

Segmentasi perkebunan kelapa sawit di wilayah Indonesia dilakukan dengan menggunakan *data mining* teknik *K-Means Clustering* guna mengelompokkan perkebunan kelapa sawit berdasarkan data luas areal, produksi, dan produktivitas. Hasil dari tiga cluster yang ditentukan, diperoleh tiga segmentasi perkebunan kelapa

sawit, yaitu 18 propinsi atau 72% termasuk dalam kelompok perkebunan yang memiliki potensi hasil perkebunan kelapa sawit rendah, 5 propinsi atau 20% termasuk dalam kelompok perkebunan yang memiliki potensi hasil perkebunan kelapa sawit sedang, dan 2 propinsi atau 8% termasuk dalam kelompok perkebunan yang memiliki potensi hasil perkebunan kelapa sawit kategori tinggi, yaitu Propinsi Riau dan Propinsi Kalimantan Tengah. Pada penelitian berikutnya dapat menambahkan variabel lainnya, seperti jumlah petani atau jumlah pekerja guna menghasilkan segmentasi yang lebih optimal pada perkebunan kelapa sawit yang terdapat di berbagai wilayah Indonesia.

DAFTAR PUSTAKA

- [1] F. Verneau, F. La Barbera, M. Amato, and V. Sodano, "Consumers' concern towards palm oil consumption an empirical study on attitudes and intention in italy," *Br. Food J.*, vol. 121, no. 9, pp. 1982–1997, 2019.
- [2] S. Hutabarat, "Optimalisasi Pemanfaatan Lahan Perkebunan Kelapa Sawit di Riau," *Unri Conf. Ser. Agric. Food Secur.*, vol. 1, pp. 46–57, 2019.
- [3] Direktorat Jenderal Perkebunan, "Statistik Perkebunan Unggulan Nasional 2019-2021," *Angew. Chemie Int. Ed. 6(11)*, 951–952., 2021.
- [4] R. Utami, E. I. Kumala Putri, and M. Ekayani, "Economy and Environmental Impact of Oil Palm Palm Plantation Expansion (Case Study: Panyabungan Village, Merlung Sub-District, West Tanjung Jabung Barat District, Jambi)," *J. Ilmu Pertan. Indones.*, vol. 22, no. 2, pp. 115–126, 2017.
- [5] M. H. Siregar, "Data Mining Klasterisasi Penjualan Alat-Alat Bangunan Menggunakan Metode K-Means (Studi Kasus Di Toko Adi Bangunan)," *J. Teknol. Dan Open Source*, vol. 1, no. 2, pp. 83–91, 2018.
- [6] S. Handoko, F. Fauziah, and E. T. E. Handayani, "Implementasi Data Mining Untuk Menentukan Tingkat Penjualan Paket Data Telkomsel Menggunakan Metode K-Means Clustering," *J. Ilm. Teknol. dan Rekayasa*, vol. 25, no. 1, pp. 76–88, 2020.
- [7] S. P. Tamba, F. T. Kesuma, and Feryanto, "Penerapan Data Mining Untuk Menentukan Penjualan Sparepart Toyota Dengan Metode K-Means Clustering," *J. Sist. Inf. dan Ilmu Komput. Prima(JUSIKOM PRIMA)*, vol. 2, no.

- 2, pp. 67–72, 2020.
- [8] F. Indriyani and E. Irfiani, “Clustering Data Penjualan pada Toko Perlengkapan Outdoor Menggunakan Metode K-Means,” *JUITA J. Inform.*, vol. 7, no. 2, p. 109, 2019.
- [9] C. Astria, A. P. Windarto, A. Wanto, and E. Irawan, “Metode K-Means pada Pengelompokan Wilayah Pendistribusian Listrik,” *Semin. Nas. Sains Teknol. Inf.*, pp. 306–312, 2019.
- [10] U. B. Mulia, “Jumlah Ritel,” vol. XI, no. 1, pp. 32–44.
- [11] M. Jesper, F. Pag, K. Vajen, and U. Jordan, “Annual Industrial and Commercial Heat Load Profiles: Modeling Based on k-Means Clustering and Regression Analysis,” *Energy Convers. Manag. X*, vol. 10, no. March, p. 100085, 2021.
- [12] I. M. Pulungan, M. Fauzan, and A. P. Windarto, “Implementasi Algoritma K-Means Clustering dalam Menentukan Blok Tanaman Sawit Paling Produktif,” in *Prosiding Seminar Nasional Riset Information Science (SENARIS)*, 2019, no. September, pp. 338–348.
- [13] R. A. D. Saragih, M. Safii, and H. S. Tambunan, “Penerapan Metode K-Means Clustering Untuk Mengelompokkan Kelapa Sawit Produktif,” in *Prosiding SiManTap: Seminar Nasional Matematika dan Terapan*, 2019, pp. 362–370.
- [14] D. F. Pasaribu, I. S. Damanik, E. Irawan, H. S. Tambunan, and K. Kunci, “Memanfaatkan Algoritma K-Means Dalam Memetakan Potensi Hasil Produksi Kelapa Sawit PTPN IV Marihat,” *BIOS J. Teknol. Inf. dan Rekayasa Komput.*, vol. 2, no. 1, pp. 11–20, 2021.
- [15] P. Deepak and A. Jurek-loughrey, “Multi-View Clustering,” pp. 27–53.
- [16] Y. Liu, X. Du, and S. Ma, *Innovative study on clustering center and distance measurement of K-means algorithm: mapreduce efficient parallel algorithm based on user data of JD mall*, no. 0123456789. Springer US, 2021.
- [17] T. Yuniarti, I. Surjandari, E. Muslim, and E. Laoh, “Data Mining Approach for Short Term Load Forecasting by Combining Wavelet Transform and Group Method of Data Handling (WGMDH),” in *In 2017 3rd International Conference on Science in Information Technology (ICSITech)*, 2017, pp. 53–58.
- [18] B. Haihong, Z. Yiqun, S. Jianshuo, and Sh. Fangchu, “Study on power consumption load forecast based on K-means.pdf,” in *2020 7th International Conference on Power and Energy Systems Engineering (CPESE 2020), 26–29 September 2020, Fukuoka, Japan*, 2020, pp. 693–700.