

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**PEMANFAATAN ALGORITMA K-MEANS DALAM
MENENTUKAN POTENSI HASIL PRODUKSI KELAPA
SAWIT**

TUGAS AKHIR

Disusun Sebagai Salah Satu Syarat
Untuk Memperoleh Gelar Sarjana Teknik
Pada Jurusan Teknik Informatika

Oleh

AYU SRI WAHYUNI

NIM. 11950125018



UIN SUSKA RIAU

UIN SUSKA RIAU

**FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU
PEKANBARU
2023**

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PERSETUJUAN

**PEMANFAATAN ALGORITMA K-MEANS DALAM
MENENTUKAN POTENSI HASIL PRODUKSI KELAPA SAWIT**

TUGAS AKHIR

Oleh

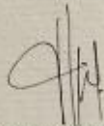
AYU SRI WAHYUNI

NIM. 11950125018

Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir

di Pekanbaru, pada tanggal 15 Januari 2024

Pembimbing I,



Dr. Elin Haeran, S.T., M.Kom
NIP. 197805042007101000

- Hak Cipta Dilindungi Undang-Undang**
1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



LEMBAR PENGESAHAN

**PEMANFAATAN ALGORITMA K-MEANS DALAM
MENENTUKAN POTENSI HASIL PRODUKSI KELAPA SAWIT**

Oleh
AYU SRI WAHYUNI
NIM. 11950125018

Telah dipertahankan di depan sidang dewan penguji
sebagai salah satu syarat untuk memperoleh gelar Sarjana Teknik
pada Universitas Islam Negeri Sultan Syarif Kasim Riau

Pekanbaru, 15 Januari 2024
Mengesahkan,
Ketua Jurusan,


Dr. Hartono, M.Pd.
NIP. 19640301 199203 1 003


Iwan Iskandar, M.T
NIP. 19821216 201503 1 003

DEWAN PENGUJI

Ketua	: Reski Mai Candra, S.T., M.Sc.		
Pembimbing I	: Dr. Elin Haerani, S.T., M.Kom		
Penguji I	: Elvia Budianita, ST, M.Cs		
Penguji II	: Liza Afriyanti, S.Kom, M.Kom.		

iii



Hak Cipta Dilindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

SURAT PERNYATAAN

Saya yang bertandatangan di bawah ini :

Nama : Ayu Sri Wahyuni
NIM : 11950125018
Tempat/Tgl.Lahir : Balam sempurna 10 Maret 2001
Fakultas : Sains dan Teknologi
Prodi : Teknik Informatika
Judul Skripsi : Pemanfaatan Algoritma K-Means Dalam Menentukan Potensi Hasil
Produksi Kelapa Sawit

Menyatakan dengan sebenar-benarnya bahwa:

1. Penulisan Skripsi dengan judul sebagaimana tersebut diatas adalah hasil pemikiran penelitian saya sendiri.
2. Semua kutipan pada karya tulis saya ini sudah disebutkan sumbernya.
3. Oleh karena itu, Skripsi saya ini, saya nyatakan bebas dari plagiat.
4. Apabila dikemudian hari terbukti terdapat plagiat dalam penulisan Skripsi saya tersebut, maka saya bersedia menerima sanksi sesuai perundang-undangan.

Demikian Surat Pernyataan ini saya buat dengan penuh kesadaran dan tanpa paksa pihak manapun juga.

Pekanbaru, 15 Januari 2024

Yang membuat pernyataan,



AYU SRI WAHYUNI

NIM. 11950125018

Pemanfaatan Algoritma K-Means Dalam Menentukan Potensi Hasil Produksi Kelapa Sawit

Ayu Sri Wahyuni*, Elin Haerani, Elvia Budianita, Liza Afrianti

Teknik Informatika, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

Email: ¹11950125018@students.uin-suska.ac.id, ²elin.haerani@uin-suska.ac.id, ³elvia.budianita@uin-suska.ac.id, ⁴liza.afriyanti@uin-suska.ac.id

Email Penulis Korespondensi: 11950125018@students.uin-suska.ac.id

Submitted: 22/12/2023; Accepted: 31/12/2023; Published: 31/12/2023

Abstrak—Mengingat pentingnya budidaya kelapa sawit saat ini dan masa depan, serta semakin meningkatnya permintaan minyak sawit oleh penduduk dunia, maka perlu dipikirkan upaya peningkatan kualitas dan kuantitas produksi minyak sawit secara tepat guna mencapai tujuan yang diinginkan dan dicapai. Berdasarkan data hasil produksi buah sawit PT Salim Ivomas Pratama Tbk terlihat di beberapa tempat produksi buahnya bervariasi. Potensi hasil buah kelapa sawit didasarkan pada luas panen, realisasi produksi dan tahun tanam nya. Pengelasteran K-Means dapat membantu mengidentifikasi potensi kelapa sawit, dengan hasil yang bervariasi dari hari ke hari. Proses ini memungkinkan lokasi dengan pola produksi serupa, yang memfasilitasi keputusan manajemen dan strategi produksi. Pada penelitian ini, wilayah potensi penanaman buah-buahan dikelompokkan menggunakan algoritma K-Means. K-Means bertujuan untuk memfasilitasi pengelompokan blok dengan produksi buah tinggi dan rendah. Data yang digunakan ialah sebanyak 180 data selama 5 tahun terakhir yakni sejak tahun 2018 hingga tahun 2022, dengan atribut Blok Panen, Luas Area, Berat janjang, dan Realisasi produk atau jumlah. Penelitian ini menggunakan bantuan software Rapidminer dan Google Colab. Hasil dari penelitian ini di dapatkan C1 (tertinggi) ialah 125 data Blok Panen dalam artian bahwa kelompok pertama termasuk kategori Hasil panen yang baik atau tinggi pada tahun 2018-2022, dan C0 (terendah) ialah 55 data Blok Panen dalam artian bahwa kelompok kedua termasuk kategori hasil panen rendah 2018-2022.

Kata Kunci: Kelapa Sawit; Hasil Panen; Data Mining; Clustering; Algoritma K-Means

Abstract—Considering the importance of oil palm cultivation now and in the future, as well as the increasing demand for palm oil by the world population, it is necessary to think about efforts to increase the quality and quantity of palm oil production appropriately in order to achieve the desired and achievable goals. Based on data on palm fruit production results from PT Salim Ivomas Pratama Tbk, it can be seen that fruit production varies in several places. The potential yield of oil palm fruit is based on the harvested area, actual production and year of planting. K-Means clustering can help identify the potential of oil palm, with results that vary from day to day. This process allows locations with similar production patterns, which facilitates management decisions and production strategies. In this research, potential fruit planting areas were grouped using the K-Means algorithm. K-Means aims to facilitate the grouping of blocks with high and low fruit production. The data used is 180 data for the last 5 years, namely from 2018 to 2022, with the attributes Harvest Block, Area, Sheet Weight, and Product Realization or quantity. This research uses the help of Rapidminer and Google Colab software. The results of this research show that C1 (the highest) is 125 Harvest Block data in the sense that the first group is included in the good or high harvest yield category in 2018-2022, and C0 (the lowest) is 55 Harvest Block data in the sense that the second group is included low yield category 2018-2022.

Keywords: Palm Oil; Harvest Results; Data Mining; Clustering; Algorithms K-Means

1. PENDAHULUAN

Kelapa sawit merupakan tanaman yang memiliki nilai ekonomi relatif tinggi karena menghasilkan minyak nabati. Kelapa sawit penting bagi Indonesia karena menciptakan lapangan kerja bagi masyarakat lokal dan merupakan sumber devisa negara[1]. Kelapa sawit merupakan tanaman tropis dan memerlukan suhu, jenis tanah, ketinggian, dan curah hujan yang tepat untuk mendapatkan hasil yang optimal[2]. Kelapa sawit (*Elaeis guineensis* jacq) saat ini merupakan tanaman perkebunan yang menempati posisi penting dalam sektor pertanian pada umumnya dan sektor perkebunan pada khususnya. Banyak pabrik yang memproduksi minyak dan lemak. Kelapa sawit menghasilkan nilai ekonomi terbesar di dunia per hektarnya[2]. Mengingat pentingnya budidaya kelapa sawit saat ini dan masa depan, serta semakin meningkatnya permintaan minyak sawit oleh penduduk dunia, maka perlu dipikirkan upaya peningkatan kualitas dan kuantitas produksi minyak sawit secara tepat guna mencapai tujuan yang diinginkan dan dicapai[3]. Salah satunya adalah pemilihan bahan tanam berupa benih yang memerlukan upaya teknis dan kultural serta pengelolaan benih yang baik. Pemilihan benih yang tepat mempengaruhi produktivitas tanaman kelapa sawit [4].

PT Salim Ivomas Pratama Tbk adalah usaha dalam bidang perkebunan kelapa sawit dalam pengelolaannya kerap mengalami penurunan hasil minyak sawit dan inti sawit karena kurangnya pemetaan lokasi potensial untuk produksi. Kendala ini disebabkan oleh luasnya perkebunan kelapa sawit di wilayah rokan hilir yang belum dipetakan dengan baik, dan pihak terkait belum menyusun rencana panen berdasarkan blok yang memiliki potensi buah yang baik [5]–[7]. Perusahaan ini terletak di wilayah Kabupaten Rokan Hilir, yang merupakan salah satu kabupaten di Provinsi Riau, Indonesia. Kabupaten Rokan Hilir dikenal sebagai salah satu daerah yang memiliki luas perkebunan kelapa sawit yang signifikan. Sebagai perusahaan kelapa sawit, PT. Salim Ivomas Pratama Tbk mungkin terlibat dalam berbagai kegiatan, termasuk penanaman, dan pengolahan kelapa sawit menjadi produk-produk turunannya seperti minyak kelapa sawit dan inti sawit [8]. PT Salim Ivomas Pratama Tbk Kabupaten Rokan

Hak cipta dilindungi Undang-Undang. Dilarang mengutip sebagian atau seluruhnya atau melakukan tindakan yang sama tanpa izin dari Universitas Islam Negeri Sultan Syarif Kasim Riau. Untuk lebih jelasnya, silakan kunjungi laman web kami di www.uin-suska.ac.id.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruhnya karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hilir diberdirikan di atas lahan seluas 7,28 hektar serta luas pembangunan pabrik 15.393 meter persegi. Kawasan ini sebagai lokasi pabrik dan kompleks perumahan para karyawan pabrik. Hingga saat ini, Pt Salim Ivomas Pratama Tbk Kabupaten Rokan Hilir memiliki dua tangki penyimpanan dan dua silo Bulk untuk menyimpan CPO (Crude palm oil), dengan kapasitas 2.000 ton CPO per tangki penyimpanan dan 5.000 ton per silo Bulk ton inti per satu unit. Jika terdapat masalah limbah cair, hasil akhir pengolahan dapat dikumpulkan dan diolah pada kolam limbah dan dialirkan dengan aman ke titik penggunaan untuk bagian dari proses. Volume kolam limbah kurang lebih 60.340 meter persegi dan luas lahan yang digunakan 142,87 hektar [9].

Clustering adalah metode yang digunakan untuk mengubah data lapangan menjadi analisis yang dikelompokkan, hal ini dapat mempermudah pengguna dalam menentukan posisi atau letak suatu titik [10], [11]. Penelitian ini bertujuan untuk mengidentifikasi lokasi potensial buah kelapa sawit berdasarkan data hasil panen dengan menggunakan teknik data mining [12]. Data mining merupakan disiplin ilmu yang memproses informasi melalui analisis data lama yang tidak dipakai [13], [14]. Pendekatan ini menggabungkan statistika, kecerdasan buatan, dan penelitian basis data yang terus berkembang.

Penelitian serupa yang dilakukan Deny Dkk pada tahun 2021 dengan judul Memanfaatkan Algoritma K-Means Dalam Memetakan Potensi Hasil Produksi Kelapa Sawit PTPN IV Marihat mendapatkan hasil C1 (tertinggi) ialah 14 data Blok Panen, dan C2 (terendah) ialah 11 data Blok Panen [15]. Selain itu juga terdapat penelitian oleh Effendi dkk pada tahun 2021 dengan judul Penerapan Metode K-Means Clustering untuk Pengelompokan Lahan Sawit Produktif pada PT Kasih Agro Mandiri, mendapatkan hasil Cluster Produktifitas Tinggi terdiri atas 26 blok lahan, Cluster Produktifitas Sedang terdiri atas 35 blok lahan, dan Cluster Produktifitas Rendah terdiri atas 39 blok lahan [16]. Selanjutnya penelitian oleh Ramadhani dan megawati pada tahun 2023 dengan judul Implementasi Algoritma K-Means Untuk Klustering Data Produktivitas Kelapa Sawit mendapatkan hasil C0 ialah 13 data, dan C1 ialah 35 data [3]. Azhari, bersama rekan-rekannya, melakukan penelitian dengan judul "Pengelompokan Hasil Produksi Tanaman Perkebunan Berdasarkan Provinsi Menggunakan Metode K-Means Clustering." Hasil penelitian ini mengindikasikan bahwa produksi tanaman perkebunan di Indonesia dapat dikelompokkan menjadi 3 Cluster, yaitu Cluster rendah, Cluster sedang, dan Cluster tinggi. Melalui penggunaan Algoritma K-Means, ditemukan bahwa Cluster tinggi terdiri dari 6 Provinsi, Cluster sedang terdiri dari 2 Provinsi, dan Cluster rendah terdiri dari 27 Provinsi. Kesimpulan yang dapat diambil dari penelitian ini adalah bahwa pengelompokan produksi tanaman perkebunan di Indonesia dapat berhasil dilakukan dengan menerapkan algoritma K-Means. [4]. Lalu, penelitian yang dilakukan oleh Fina Nasari yang berjudul "Optimalisasi Pengelompokan Data Produksi Kelapa Sawit Menggunakan Algoritma K-Medoids" menghasilkan data untuk jenis 2 kelompok, yakni 7 data pada kelompok 1 dan 5 data pada kelompok 2. Pada jenis 3 kelompok, didapatkan data pada kelompok 1, 4 data pada kelompok 2, dan 3 data pada kelompok 3. Sementara untuk jenis 5 kelompok, hasilnya adalah 3 data pada kelompok 1, 2 data pada kelompok 2 dan 4, 1 data pada kelompok 3, dan 2 data pada kelompok 4. Evaluasi menggunakan nilai Davies-Bouldin Index (DBI) menunjukkan bahwa jenis kelompok 2 menjadi jenis kelompok terbaik dengan nilai DBI sebesar -159796492242,667 [2].

Dalam penelitian ini, algoritma k-means dari data mining digunakan untuk mengelompokkan lokasi potensial kelapa sawit pada PT Salim Ivomas Pratama Tbk. K-menas adalah metode pengelompokan data dan hirarki yang membagi data kedalam dua atau lebih kelompok [6], [17], [18]. Selanjutnya peneliti menggunakan evaluasi clustering. Penelitian ini melakukan evaluasi clustering dengan maksud untuk menilai sejauh mana kualitas hasil clustering yang dihasilkan. Dalam konteks penelitian ini, penilaian hasil clustering menggunakan Davies Bouldin Index (DBI) bertujuan untuk menentukan jumlah cluster yang paling optimal. DBI, yang diperkenalkan oleh David L. Davies dan Donald W. Bouldin pada tahun 1979, adalah metode yang digunakan untuk mengukur validitas atau keoptimalan jumlah cluster pada suatu metode pengelompokan. DBI mengukur kohesi sebagai jumlah kedekatan data terhadap titik pusat cluster yang diikuti [19].

2. METODOLOGI PENELITIAN

Tahapan penelitian meliputi urutan dari awal hingga akhir penelitian. Tahapan dalam penelitian ini dapat dilihat pada gambar 1.



Gambar 1. Tahapan Penelitian

2.1 Perumusan Masalah

Dalam penelitian awal adalah menentukan fokus penelitian atau perumusan masalah dalam penelitian ini. Perumusan masalah yang didapatkan yaitu bagaimana pemanfaatan algoritma data k-means dalam menentukan potensi hasil kelapa sawit.

2.2 Pengumpulan Data

Untuk mendapatkan informasi dan data yang dibutuhkan. Peneliti melakukan pengumpulan data dengan dua tahap yaitu studi pustaka dan observasi

a. Studi Kasus

Studi pustaka merupakan tahapan untuk mengumpulkan literature yang dibutuhkan dalam penelitian. Metode ini dapat dilakukan dengan cara mencari sumber referensi yang ada di jurnal dengan berbagai jurnal yang ada [15].

b. Observasi

Dalam observasi ini dilakukan untuk melihat, mempelajari secara langsung serta mengumpulkan data produksi hasil kelapa sawit. Data yang didapat merupakan data primier, data hasil produksi di PT Salim Ivomas Pratama dengan data 5 tahun dari 2018-2022 dengan format xlsx. Atribut yang di dapat pada data ini yaitu, Blok Panen, Luas Areal, Berat janjang dan Jumlah.

2.3 Proses Data Mining (Algoritma K-Means)

Pada penelitian ini terdapat beberapa tahapan data mining yang dilakukan, yakni :

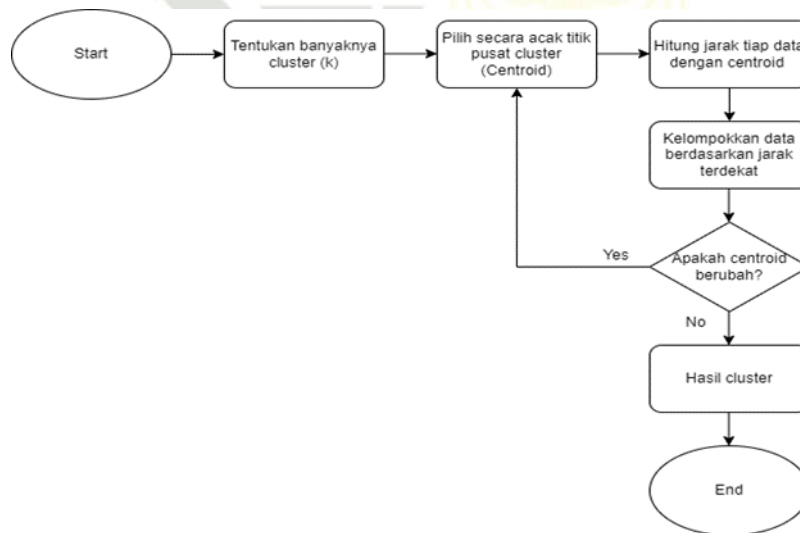
a. Data Selection

Seleksi Data adalah tahap di mana atribut-atribut yang akan digunakan dalam proses data mining dipilih, karena tidak semua atribut dari sumber data dapat digunakan sepenuhnya. Data yang digunakan dalam penelitian ini melibatkan atribut-atribut seperti Blok Panen, Luas Area, Berat Janjang, dan jumlah realisasi produk.

b. Transformation data

Membuang data yang tidak konsisten dan noise, duplikasi data, memperbaiki kesalahan data, dan bisa diperkaya dengan data eksternal yang tepat merupakan tujuan dari data Pre-Processing dan Cleaning Data. Data yang digunakan dalam penelitian ini, peneliti melakukan transformation data untuk mengatasi data yang kosong.

Selanjutnya data di olah menggunakan Algoritma K-Means. Dalam penerapan Algoritma K-Means terdapat rangkaian langkah yang dapat dilakukan melalui flowchart. Berikut gambaran perhitungan algoritma k-means.



Gambar 2. Flowchart Algoritma K-Means

Tahap yang dilakukan untuk menghitung algoritma K-Means sebagai berikut ini:

- Menetapkan jumlah cluster (k)
- Mempilih titik pusat (centroid) cluster secara acak
- Temukan jarak tiap data terhadap centroid menggunakan rumus Euclidean yaitu:

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

Keterangan:

- d(x,y) : Jarak antara titik data x dan titik data y
 x : Titik data pada objek
 y : Titik data pada centroid
 i : Jumlah objek

- d. Mengelompokkan data jadi yang dijumlahkan menurut jarak minimum (minimum) antara data dengan pusat cluster atau data centroid untuk memperoleh cluster baru.
- Menghitung ulang setiap cluster sebagai pusat cluster baru (centroid) berdasarkan data yang mengikutinya. Nilai centroid baru ditentukan berdasarkan rata-rata data tiap cluster. Rumusnya adalah:

$$CI = \frac{x_1+x_2+x_3+\dots+x_n}{\Sigma x} \tag{2}$$

Keterangan:

- CI : centroid baru
- x1 : nilai cluster ke-1
- xn : nilai cluster ke-n
- x : Total data

Setelah mendapatkan centroid baru, lakukan iterasi berikutnya atau ulangi langkah c sampai e hingga tidak ada lagi data yang dipindahkan dari cluster.

Implementasi dan Hasil

Tahapan selanjutnya adalah implementasi dan hasil. Tahapan implementasi merupakan tahapan dalam penelitian menggunakan alat Rapidminer dan Google Collab untuk mencapai hasil yang diinginkan.

2.5 Uji Validitas

Tahap ini dilakukan bertujuan sebagai membandingkan hasil dari pengerjaan pada google collab dan pada tools rapidminer. Sehingga dapat dijadikan perbandingan.

2.6 Kesimpulan

Kesimpulan diambil dari hasil pengujian yang telah dilakukan terhadap algoritma K-means.

3. HASIL DAN PEMBAHASAN

Pada penelitian ini data yang digunakan merupakan data primer yang berasal dari PT Salim Ivomas Pratama Tbk. data terdiri dari 4 atribut yaitu Blok Panen, Luas Area, Berat Janjang dan Jumlah. Data yang digunakan ialah data hasil panen selama 5 tahun terakhir yakni sejak tahun 2018 hingga tahun 2022. Berdasarkan penjelasan tersebut, dapat dilihat pada tabel 1 yang berisi data dari PT Salim Ivomas Pratama Tbk.

Tabel 1. Data Hasil Produksi Kelapa Sawit

BLOK PANEN	LUAS AREAL (HA)	BERAT JANJANG (KG)	JUMLAH (KG)
E44	30	26,828	604,292
E45	30	25,718	609,340
E46	30	26,161	608,125
E47	30	26,738	628,510
E48	30	26,783	631,508
E49	30	26,103	613,748
E50	31	27,312	644,785
E51	31	26,731	603,002
E52	30	26,598	614,705
E53	30	25,216	598,792
...
G47	30	28,350	634,122
G48	30	27,904	637,212
G49	30	28,643	629,271
G50	30	28,859	626,175
G51	30	27,835	620,335
G52	30	27,550	630,290
G53	31	29,336	636,764
G54	30	29,035	632,009
G55	30	28,187	638,973

Data di atas diolah menggunakan tahapan data mining. Proses pengolahan datanya menggunakan algoritma K-Means dengan menggunakan 10 data pertama sebagai contoh perhitungan manual.

- a. Tentukan jumlah cluster (k). Jumlah cluster ditentukan dari kebutuhan produksi. Jumlah cluster yang digunakan pada penelitian sebanyak 2 cluster.
- b. Tentukan nilai asal cluster (pusat) untuk cluster (k) sebanyak-banyaknya. Nilai titik tengah dihasilkan secara acak. Titik cluster pertama dapat dilihat pada tabel 2 sebagai berikut:

2. Diarangi mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.
 2. Diarangi mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.
 2. Diarangi mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang
 © 2023 UIN Suska Riau
 State Islamic University of Sultan Syarif Kasim Riau

Tabel 2. Titik Cluster Awal (Iterasi-1)

CENTROID	LUAS AREAL (HA)	BERAT JANJANG (KG)	JUMLAH (KG)
C0	30	25,718	609,340
C1	31	27,312	644,785

Hitung jarak terdekat data ke setiap nilai centroid adalah sebagai berikut.

- Perhitungan data ke-1 terhadap centroid cluster 0

$$d(1,0) = \sqrt{(30 - 30)^2 + (26.828 - 25,718)^2 + (604,292 - 609,340)^2} = 5168.597$$

- Perhitungan data ke-1 terhadap centroid cluster 1

$$d(1,1) = \sqrt{(30 - 31)^2 + (26.828 - 27,312)^2 + (604,292 - 644,785)^2} = 40495.8925$$

Hasil perhitungan menggunakan rumus Euclidean distance pada data awal menunjukkan bahwa data pertama merupakan data yang paling dekat dengan cluster 1, dilihat dari nilai perhitungan yang paling kecil dibandingkan dengan cluster lainnya. Dengan demikian, data 1 dapat dikategorikan ke dalam kelompok cluster 1. Proses perhitungan ini diulang untuk setiap data hingga semua data ditempatkan dalam kelompok cluster yang paling dekat.

- Setelah menghitung jarak data dari centroid, selanjutnya mengelompokkan data dari nilai terdekat (terkecil). Hasil pengelompokannya disajikan pada Tabel 3.

Tabel 3. Hasil Pengelompokan Data

Objek ke-	C0	C1	Cluster
1	5168.59788	40495.8925	1
2	0	35480.8239	1
3	1293.24166	36678.0643	1
4	19197.117	16285.119	2
5	22193.5677	13287.5344	2
6	4424.78124	31060.5385	1
7	35480.8239	0	2
8	6418.44327	41787.0393	1
9	5436.69247	30088.4728	1
10	10559.9388	46040.7349	1

Hitung nilai rata-rata setiap atribut data dalam cluster, perbarui nilai centroid baru untuk iterasi berikutnya. Selanjutnya menghitung nilai centroid baru menggunakan persamaan (2).

- Rata-rata data pada cluster 0

$$CI(Luas Areal) = \frac{211}{7} = 30,14$$

$$CI(Berat Janjang) = \frac{183,355}{7} = 26193,57$$

$$CI(Jumlah) = \frac{4,525,004}{7} = 607,429.14$$

- Rata-rata data pada cluster 1

$$CI(Luas Areal) = \frac{91}{3} = 30$$

$$CI(Berat Janjang) = \frac{80,833}{3} = 26,944$$

$$CI(Jumlah) = \frac{1,904,803}{3} = 634,934$$

Hasil dari kalkulasi untuk nilai centroid baru di iterasi-2 bisa dilihat dalam tabel 4 berikut.

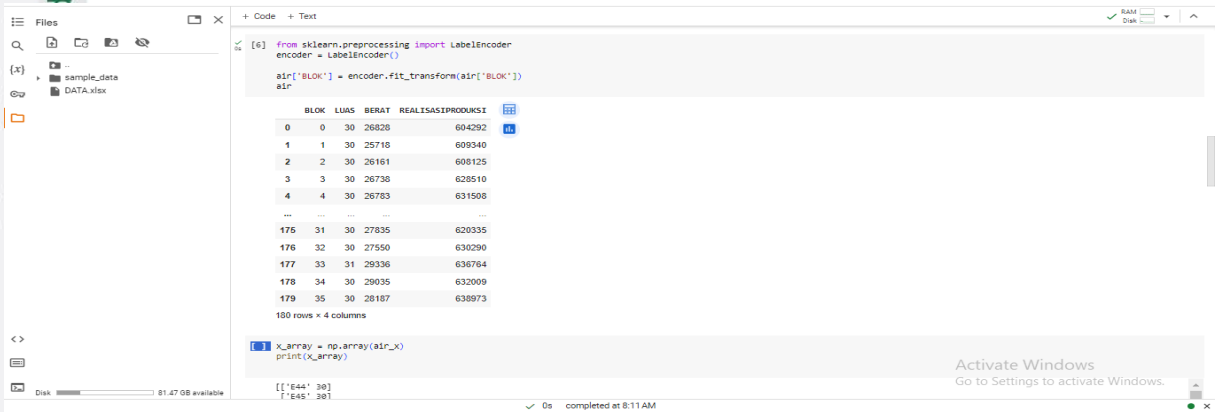
Tabel 4. Centroid Baru (Iterasi-2)

Objek ke-	C0	C1	Cluster
1	3200.65071	30642.5542	1
2	1969.14784	25623.6959	1
3	696.619036	26820.7749	1
4	21087.8861	6427.64594	2
5	24086.0704	3430.12954	2
6	6319.50621	21203.0319	1
7	37372.5962	9857.5257	2
8	4459.64394	31933.046	1
9	7287.08856	20232.2978	1
10	8692.28869	36183.6344	1

f. Jarak data dihitung ulang menggunakan nilai pusat massa yang baru. Jika tidak ada data yang berpindah antar cluster atau iterasi terakhir memiliki data yang sama dengan iterasi sebelumnya, ulangi prosedur c hingga e.

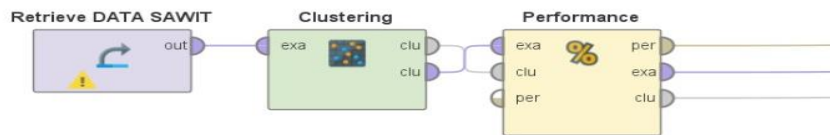
Implementasi dan Hasil

Implementasi clustering dalam penelitian kali ini menggunakan dua tools yakni Google Colab dan Rapidminer. Berikut pernyataan untuk model K-means produksi kelapa sawit dengan menggunakan algoritma K-Means.



Gambar 3. Proses pengolahan data oleh google collab

Pada gambar 3 menunjukkan proses input data dan isi data. Selanjutnya, algoritma k-means digunakan untuk mengolah data pada google collab. Setelah menjalankan pengujian pada Google Colab, selanjutnya melakukan proses clustering pada rapidminer. Berikut merupakan gambar proses pada tools rapidminer.



Gambar 4. Proses clustering

Gambar 4 merupakan gambar proses clustering dalam tools Rapidminer. Mulanya Data sawit di input ke dalam Rapidminer. Data di input tanpa melakukan transformasi, karena data sudah siap diolah pada tools rapidminer. Setelah itu, penulis menambah operator clustering dan menghubungkan data sawit dengan operator clustering dalam operator clustering, penulis memilih algoritma K-Means clustering. Kemudian penulis menambahkan operator Performance yang berguna untuk mengevaluasi kinerja pengelompokan berbasis centroid. Operator performance ini memberikan daftar nilai kinerja berdasarkan pusat Cluster. Hasil clustering menggunakan rapidminer ditunjukkan pada gambar 5.

Cluster Model

```

Cluster 0: 55 items
Cluster 1: 125 items
Total number of items: 180
  
```

Gambar 5. Cluster Model

Gambar 5 menunjukkan hasil clustering data menggunakan rapidminer, dimana dari 180 data hasil produksi kelapa sawit yang diuji hasil yang didapatkan yaitu 125 data termasuk kedalam cluster 0 dan 55 data termasuk ke dalam cluster 1. Kemudian peneliti melakukan pengujian dengan Davies Bouldien Index (DBI) yang mendapatkan hasil Sebagai berikut.

PerformanceVector

```

PerformanceVector:
Avg. within centroid distance: 173520469.413
Avg. within centroid distance_cluster_0: 252715162.026
Avg. within centroid distance_cluster_1: 138674804.664
Davies Bouldin: 0.218
  
```

Gambar 6. Nilai Davies Bouldien Index

2. Diarangi mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.
 1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

Pada gambar 6 menunjukkan hasil dari pengujian cluster menggunakan Davies Bouldien Index (DBI). Nilai DBI yang dihasilkan ialah 0,218. Selain itu juga hasil rata-rata jarak centeroid dan anggota pada cluster 1 yakni sebesar 13864804.664 dimana lebih besar dari nilai hasil rata-rata jarak centeroid dan anggota pada cluster 2 yakni sebesar 252715162.026. Hal ini menunjukkan bahwa cluster 1 lebih baik daripada cluster 2 dikarenakan cluster 1 memiliki nilai rata-rata jarak centeroid dan anggota yang lebih kecil daripada cluster 2. Nilai DBI yang terkecil menunjukkan hasil yang paling baik dan menunjukkan cluster yang optimal [8], [20]. Grafik yang didapatkan dari clustering ini dapat dilihat pada gambar berikut.



Gambar 7. Grafik Clustering

Gambar 7 menunjukkan grafik clustering beserta anggota nya. Terlihat bahwa anggota cluster 1 lebih banyak dari pada anggota cluster 0. Yakni cluster 1 sebanyak 120 anggota dan cluster 0 sebanyak 55 anggota.

4. KESIMPULAN

Berdasarkan uraian sebelumnya, dapat disimpulkan bahwa metode yang diterapkan dalam penelitian ini adalah metode K-Means Clustering dalam rangka melakukan perhitungan data mining. Hasil dari penelitian menggunakan metode K-Means ini menghasilkan dua klaster, yaitu klaster tinggi dan klaster rendah. Penerapan algoritma K-Means ini bermanfaat untuk memetakan hasil produksi buah sawit. Implementasi algoritma K-Means dapat dilakukan menggunakan perangkat lunak RapidMiner dan Google Colab. Hasil penelitian ini menunjukkan bahwa C1 (kluster tertinggi) terdiri dari 125 data Blok Panen, yang berarti bahwa kelompok ini mencakup kategori hasil panen yang baik atau tinggi pada periode tahun 2018-2022. Sementara itu, C0 (kluster terendah) terdiri dari 55 data Blok Panen, yang mengindikasikan bahwa kelompok ini termasuk dalam kategori hasil panen rendah pada periode yang sama, yaitu tahun 2018-2022.

REFERENCES

- [1] E. F. Himmah, M. Widyaningsih, and M. Maysaroh, "Identifikasi Kematangan Buah Kelapa Sawit Berdasarkan Warna RGB Dan HSV Menggunakan Metode K-Means Clustering," *J. Sains dan Inform.*, vol. 6, no. 2, pp. 193–202, 2020, doi: 10.304128/jsi.v6i2.242.
- [2] F. Nasari and U. P. Utama, "Optimalisasi Pengelompokan Data Produksi Kelapa Sawit Menggunakan Algoritma K-Medoids," vol. 5, no. 1, 2022.
- [3] I. Ramadhani and M. Megawati, "Implementasi Algoritma K-Means Untuk Klustering Data Produktivitas Kelapa Sawit: Implementation Of K-Means Algorithm For Palm Oil Productivity Data Clustering," *Indones. J. Inform. ...*, vol. 3, no. 1, pp. 56–64, 2023, [Online]. Available: <https://journal.irpi.or.id/index.php/ijirse/article/view/488%0Ahttps://journal.irpi.or.id/index.php/ijirse/article/download/488/261>
- [4] A. A. Simangunsong, I. Gunawan, and Z. M. Nasution, "Pengelompokan hasil produksi tanaman perkebunan berdasarkan provinsi menggunakan metode K-Means," *J. Mach. Learn. Artif. Intelegence*, vol. 1, no. 4, pp. 273–284, 2022, doi: 10.55123/jomlai.v1i4.1661.
- [5] D. Marlina, N. Lina, A. Fernando, and A. Ramadhan, "Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokan Wilayah Sebaran Cacat pada Anak," *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 4, no. 2, p. 64, 2018, doi: 10.24014/coreit.v4i2.4498.
- [6] Z. Nabila, A. Rahman Isnain, and Z. Abidin, "Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means," *J. Teknol. dan Sist. Inf.*, vol. 2, no. 2, p. 100, 2021, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTISI>
- [7] A. Dan Pemetaan Jumlah Penumpang Kereta Api Di, B. Wijaya, T. Maulana Fahrudin, and A. Nugroho, "Indonesia

2. Diarangi mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 a. Pengutipan hanya untuk kepentingan atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber.

- Menggunakan Metode Statistik Deskriptif Dan K-Means Clustering ARTICLEINFO ABSTRACT,” JurnalMantik, vol. 3, no. 2, pp. 1-9, 2019, [Online]. Available: <https://iocscience.org/ejournal/index.php/mantik/index>
- 8 W. I. Mawaddah, I. Gunawan, and I. P. Sari, “Implementation of Data Mining Algorithm for Clustering of Palm Oil Harvested Data,” JOMLAI J. Mach. Learn. Artif. Intell., vol. 1, no. 1, pp. 43-54, 2022, doi: 10.55123/jomlai.v1i1.163.
 - 9 W. F. Sari, M. Muslimin, A. Franz, and P. Sugiartawan, “Deteksi Tingkat Kematangan Tandan Buah Segar Kelapa Sawit dengan Algoritme K-Means,” SINTECH (Science Inf. Technol. J., vol. 5, no. 2, pp. 154-164, 2022, doi: 10.31598/sintechjournal.v5i2.1146.
 - 10 mohamad jajuli nurul rohmawati, sofi defiyanti, “Implementasi Algoritma K-Means Dalam Pengklasteran Mahasiswa Pelamar Beasiswa,” Jitter 2015, vol. I, no. 2, pp. 62-68, 2015.
 - 11 D. Anitasari and W. J. Pranot, “Clustering Penggunaan Fuel Pada PT Trasindo Murni Perkasa Menggunakan Algoritma K-Means,” J. Tek. Mesin, Ind. Elektro Dan Inform., vol. 3, no. 1, 2024.
 - 12 S. Wijayanto and Y. Fathoni, M, “Pengelompokan Produktivitas Tanaman Padi di Jawa Tengah Menggunakan Metode Clustering K-Means,” J. JUPITER, vol. 13, no. 2, pp. 212-219, 2021.
 - 13 L. Gayatri and H. Hendry, “Pemetaan Penyebaran Covid-19 Pada Tingkat Kabupaten/Kota Di Pulau Jawa Menggunakan Algoritma K-Means Clustering,” Sebatik, vol. 25, no. 2, pp. 493-499, 2021, doi: 10.46984/sebatik.v25i2.1307.
 - 14 P. L. and A. B. Warsito, “Penerapan Algoritma K-Means Untuk Clustering Big Five Personality,” J. Tek. Mesin, Ind. Elektro Dan Inform., vol. 3, no. 1, 2024.
 - 15 D. F. Pasaribu, I. S. Damanik, E. Irawan, Suhada, and H. S. Tambunan, “Memanfaatkan Algoritma K-Means Dalam Meningkatkan Potensi Hasil Produksi Kelapa Sawit PTPN IV Marihat,” BIOS J. Teknol. Inf. dan Rekayasa Komput., vol. 2, no. 1, pp. 11-20, 2021, doi: 10.37148/bios.v2i1.17.
 - 16 H. H. Hendi, A. Syahrial, S. Prayoga, and W. D. Hidayat, “Penerapan Metode K-Means Clustering untuk Pengelompokan Lahan Sawit Produktif pada PT Kasih Agro Mandiri,” Teknomatika, vol. 11, no. 02, pp. 117-126, 2021.
 - 17 A. Qurrata, A. Nazir, L. Handayani, and I. Afrianty, “Penerapan Algoritma K-Means Clustering untuk Mengetahui Pola Penclustern Beasiswa Bank Indonesia (BI),” vol. 4, no. 3, pp. 530-539, 2023, doi: 10.47065/josyc.v4i3.3343.
 - 18 A. Sulistiyawati and E. Supriyanto, “Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan,” J. Tekno Kompak, vol. 15, no. 2, p. 25, 2021, doi: 10.33365/jtk.v15i2.1162.
 - 19 E. Muningsih, I. Maryani, and V. R. Handayani, “Penerapan Metode K-Means dan Optimasi Jumlah Cluster dengan Index Davies Bouldin untuk Clustering Propinsi Berdasarkan Potensi Desa,” vol. 9, no. 1, pp. 95-100, 2021.
 - 20 E. Purwaningsih and E. Nurelasari, “Implementasi Metode K-Means Clustering Dengan Davies Bouldin Index Pada Analisis Faktor Penyebab Perceraian,” J. Inf. Manag., vol. 7, no. 2, 2023.