

**ANALISIS ALGORITMA *K-MEDOIDS* UNTUK
PENGELOMPOKAN HASIL PANEN SAWIT
BERDASARKAN BERAT DAN KUALITAS
(STUDI KASUS CV. RAM BINTANG MOTOR)**

SKRIPSI

Oleh :

YUSRIL IZZA HAHOLONGAN SIREGAR

198160028



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MEDAN AREA
MEDAN
2024**

UNIVERSITAS MEDAN AREA

© Hak Cipta Di Lindungi Undang-Undang

1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 29/5/24

Access From (repository.uma.ac.id)29/5/24

**ANALISIS ALGORITMA K-MEDOIDS UNTUK
PENGELOMPOKAN HASIL PANEN SAWIT
BERDASARKAN BERAT DAN KUALITAS
(STUDI KASUS : CV.RAM BINTANG MOTOR)**

SKRIPSI

Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh
Gelar Sarjana Di Fakultas Teknik
Universitas Medan Area

Oleh:

YUSRIL IZZA HAHOLONGAN SIREGAR

198160028

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MEDAN AREA
MEDAN
2024**

UNIVERSITAS MEDAN AREA

© Hak Cipta Di Lindungi Undang-Undang

1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 29/5/24

Access From (repository.uma.ac.id)29/5/24

HALAMAN PENGESAHAN

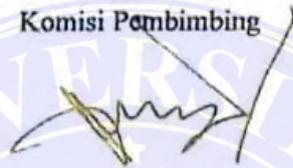
Judul Skripsi : Analisis Algoritma *K-Medoids* Untuk Pengelompokan Hasil Panen Sawit Berdasarkan Berat Dan Kualitas (Studi Kasus : CV.Ram Bintang Motor)

Nama : Yusril Izza haolongan Siregar

NPM : 198160028

Fakultas : Teknik

Disetujui Oleh
Komisi Pembimbing



Dr. Dian Noviandri, ST, M.Kom
Pembimbing I

Diketahui Oleh :

Dekan Fakultas Teknik



Enning Aputratno, ST, MT
NIDN : 0102027402

Ka. Prodi



Rizki Muliana S.Kom.M.Kom
NIDN : 010903902

Tanggal Lulus : 1 februari 2024

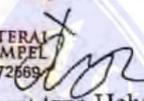
HALAMAN PERNYATAAN

Saya menyatakan bahwa skripsi yang saya susun, sebagai syarat memperoleh gelar sarjana merupakan hasil karya tulis saya sendiri. Adapun bagian – bagian tertentu dalam penulisan skripsi ini yang saya kutip dari hasil karya orang lain telah dituliskan sumbernya secara jelas sesuai dengan norma, kaidah, dan etika penulisan ilmiah.

Saya bersedia menerima sanksi pencabutan gelar akademik yang saya peroleh dan sanksi – sanksi lainya dengan peraturan yang berlaku, apabila di kemudian hari ditemukan adanya plagiat dalam skripsi ini.



Medan, 1 Februari 2024


Yusril Izza Haholongan Siregar
NPM 198160028

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI
TUGAS AKHIR/SKRIPSI/TESIS UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Medan Area, Saya bertanda tangan dibawah ini:

Nama : Yusril Izza Haholongan Siregar
NPM : 198160028
Program Studi : Teknik Informatika
Fakultas : Teknik
Jenis Karya : Skripsi

Demii pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Medan Area **Hak Bebas Royalti Noneklusif (Non-exclusive Royalty-Free Right)** atas karya ilmiah saya yang berjudul :

Analisis Algoritma K-Medoids Untuk Pengelompokan Hasil Panen Sawit Berdasarkan Berat Dan Kualitas.

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneklusif ini Universitas Medan Area berhak menyimpan, mengalih media/format-kan, mengelola dalam bentuk pangkalan data (*Database*), merawat, dan memublikasikan skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Medan
Pada Tanggal : 1 Februari 2024
Yang Menyatakan


Yusril Izza Haholongan Siregar
NPM 198160028

ABSTRAK

Kelapa sawit merupakan salah satu komoditas utama tanaman perkebunan Indonesia sebagai penghasil devisa Negara. Jenis tanaman sawit ini menduduki posisi terpenting disektor pertanian, karena kelapa sawit mampu menghasilkan nilai ekonomi yang besar tiap hektarnya jika dibandingkan dengan tanaman penghasil minyak lainnya. Ditemukan beberapa keluhan dari pengusaha sawit karena tidak mengetahui seberapa banyak sawit yang diterima dari kalangan masyarakat yang memiliki kualitas yang bagus dan kurang bagus. Maka dari itu perlu untuk mengelompokkan pengiriman hasil sawit masyarakat berdasarkan berat dan kualitas guna mempertimbangkan tujuan pabrik pengiriman buah sawit. *K-Medoids clustering* merupakan suatu algoritma yang digunakan untuk mengelompokkan hasil panen sawit karena algoritma ini dapat menemukan medoids di dalam sebuah kluster yang merupakan titik pusat dari suatu kelompok. Pengelompokan untuk klaster pertama (klaster rendah) ada 435 data sawit, sedangkan klaster kedua (klaster sedang) mendapatkan angka sebesar 444 data, dan yang ketiga(klaster tinggi) mendapatkan 121 dari total data. Sehingga dapat diketahui hasil panen sawit yang memiliki kelompok paling tinggi untuk kemudian disesuaikan dengan kriteria penerimaan dari berbagai pabrik kelapa sawit. Dengan metode *K-medoids clustering* dalam mengelompokkan hasil panen sawit, maka proses pengelompokan dapat diproses dengan cepat dan akan ditampilkan dalam laporan.

Kata Kunci : *K-medoids, clustering, pengelompokan hasil panen sawit*

ABSTRACT

Palm oil is one of Indonesia's most important plantation crops as a source of foreign exchange for the country. This type of oil palm plant occupies the most important position in the agricultural sector because oil palm can produce a high economic value per hectare compared to other oil producing plants. There have been several complaints from palm oil entrepreneurs because they did not know how much palm oil they received from the community, which had good and poor quality. Therefore, it is necessary to cluster the community's shipments of palm oil products based on weight and quality to consider the destination of the palm fruit shipping factory. K-Medoids clustering is an algorithm used to group palm oil yields because this algorithm can find medoids in a cluster which is the central point of a group. The grouping for the first cluster (low cluster) had 435 palm oil data, while the second cluster (medium cluster) had 444 data, and the third cluster (high cluster) had 121 of the total data. Thus, it was known which palm oil crop had the highest cluster and then adjusted to the acceptance criteria of different palm oil mills. By using the K-medoids clustering method to group the palm oil yields, the grouping process could be quickly processed and displayed in the report.

Keyword : K-medoids, Clustering, Grouping Palm Oil Yields



RIWAYAT HIDUP

Penulis dilahirkan di Desa Balakka pada tanggal 06 juli 2001 dari ayah Jamaluddin Siregar dan ibu Yanti Harahap penulis merupakan anak ke 4 dari 4 bersaudara.

Tahun 2019 penulis lulus dari SMK Swasta Yapim Barumun tengah dan pada tahun 2019 terdaftar sebagai mahasiswa Fakultas Teknik Universitas Medan Area. Penulis melaksanakan kerja praktek (KP) di Kantor Desa Gunung Malintang.



KATA PENGANTAR

Puji Syukur penulis ucapkan kehadiran Allah SWT, berkat Rahmat dan Karunia- NYA sehingga penulis bisa menyelesaikan skripsi yang berjudul; Analisis Algoritma *K- medoids* Untuk Pengelompokan Hasil Panen Sawit Berdasarkan Berat Dan Kualitas (Study Kasus : CV. Ram Bintang Motor).

Penyusunan Skripsi ini sebagai syarat untuk menyelesaikan Program Sarjana (S1) di Fakultas Teknik Informatika Universitas Medan Area.

Selanjutnya penulis ingin menyampaikan rasa terima kasih yang tak terhingga kepada semua pihak yang membantu kelancaran penulisan skripsi ini. Disamping itu izinkan penulis untuk menyampaikan ucapan terima kasih kepada:

1. Allah SWT dengan segala rahmat dan karunia-Nya yang memberikan kekuatan bagi penulis dalam menyelesaikan skripsi ini
2. Kedua Orang Tua penulis, Abang dan Kakak penulis yang senantiasa memberikan dukungan dan doa yang tiada henti
3. Bapak Prof. Dr. Dadan Ramdan M. Eng, M.Sc selaku Rektor Universitas Medan Area
4. Bapak Dr.Eng. Supriatno, ST, MT selaku Dekan Fakultas Teknik Universitas Medan Area
5. Bapak Rizki Muliono, S.Kom, M.Kom selaku Ketua Jurusan Informatika Universitas Medan Area beserta seluruh staffnya
6. Bapak Dr. Dian Noviandri, ST, M. Kom selaku Dosen Pembimbing yang telah memberikan bimbingan dan arahan kepada penulis.
7. Bapak Ibu dosen Fakultas Teknik yang telah memberikan ilmunya kepada penulis, semoga Bapak dan Ibu dosen selalu dalam rahmat dan lindung

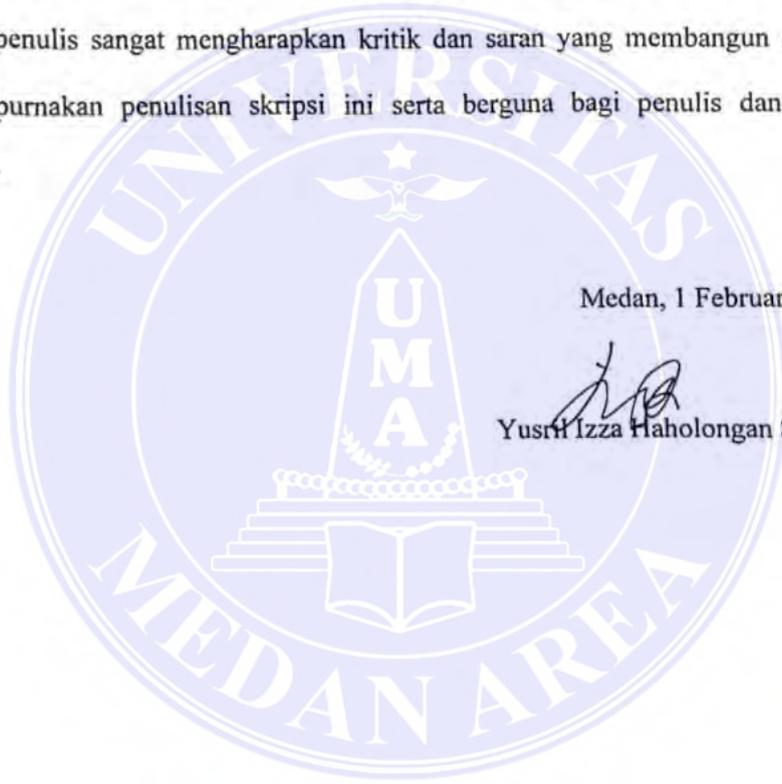
Allah SWT. Sehingga ilmu yang telah diajarkan dapat bermanfaat dikemudian hari

8. Seluruh teman-teman informatika angkatan 2019 yang tidak dapat disebutkan satu persatu. Terima kasih atas pertemanan selama ini
9. Semua pihak yang tidak dapat disebutkan satu persatu, yang telah ikhlas memberikan doa dan motivasi sehingga dapat terselesaikan skripsi ini.

Dalam penulisan skripsi ini masih banyak kekurangan dan kesalahan maka dari itu penulis sangat mengharapkan kritik dan saran yang membangun untuk menyempurnakan penulisan skripsi ini serta berguna bagi penulis dan para pembaca

Medan, 1 Februari 2024


Yusri Izza Haholongan Siregar



DAFTAR ISI

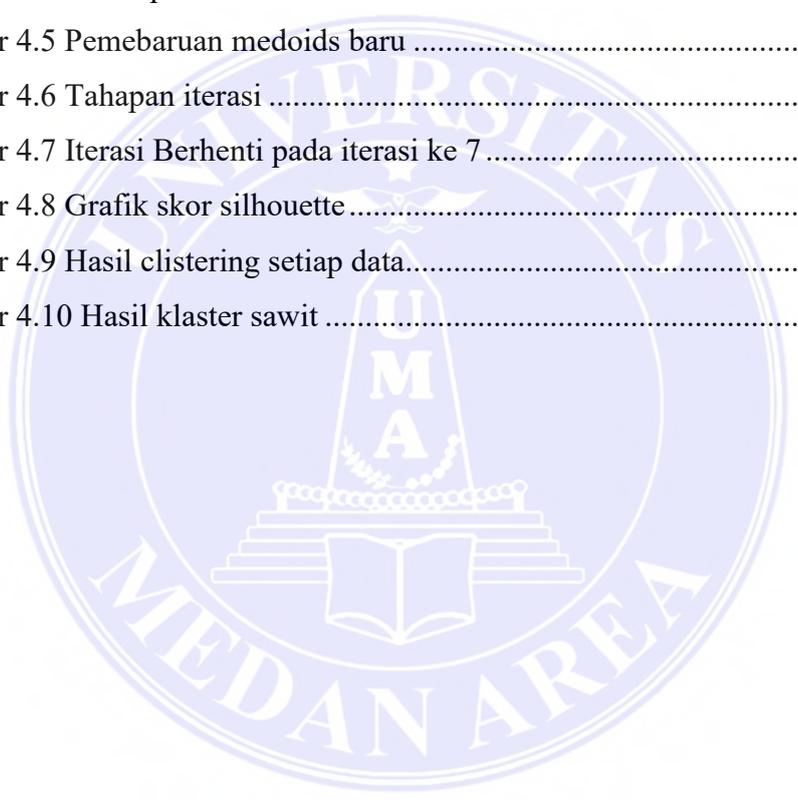
ABSTRACT	Error! Bookmark not defined.
RIWAYAT HIDUP	ii
KATA PENGANTAR.....	iv
DAFTAR ISI.....	vi
DAFTAR GAMBAR	viii
DARTAR TABEL	ix
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian.....	4
1.6 Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA.....	5
2.1 Hasil Panen Sawit	6
2.2 Kualitas dan berat Sawit.....	7
2.3 Ram Bintang Motor.....	9
2.4 Data Mining.....	9
2.5 Data Mining Sebagai Proses Dalam <i>Knowledge Discovery in Data</i>	11
2.6 Tugas-tugas Data Mining	13
2.7 Algoritma <i>K-MEDOIDS</i>	15
2.8 Silhouette.....	17
2.9 Normaslisasi Min-Max.....	18
2.10 Google Colab.....	19
2.11 Penelitian Terdahulu	19
BAB III METODE PENELITIAN.....	22
3.1 Subjek Dan Objek Penelitian	22
3.2 Alat Penelitian	22
3.3 Prosedur Penelitian.....	23
3.3.1 Studi Literatur	25
3.3.2 Analisis Masalah.....	26

3.3.3 Mengumpulkan Data.....	26
3.3.4 Menentukan Metode Penelitian.....	28
3.3.5 Melakukan <i>Clustering</i>	28
BAB IV HASIL DAN PEMBAHASAN	37
4.1 Hasil dan Pembahasan.....	37
BAB V KESIMPULAN DAN SARAN.....	45
5.1 Kesimpulan	45
5.2 Saran.....	45
DAFTAR PUSTAKA.....	46



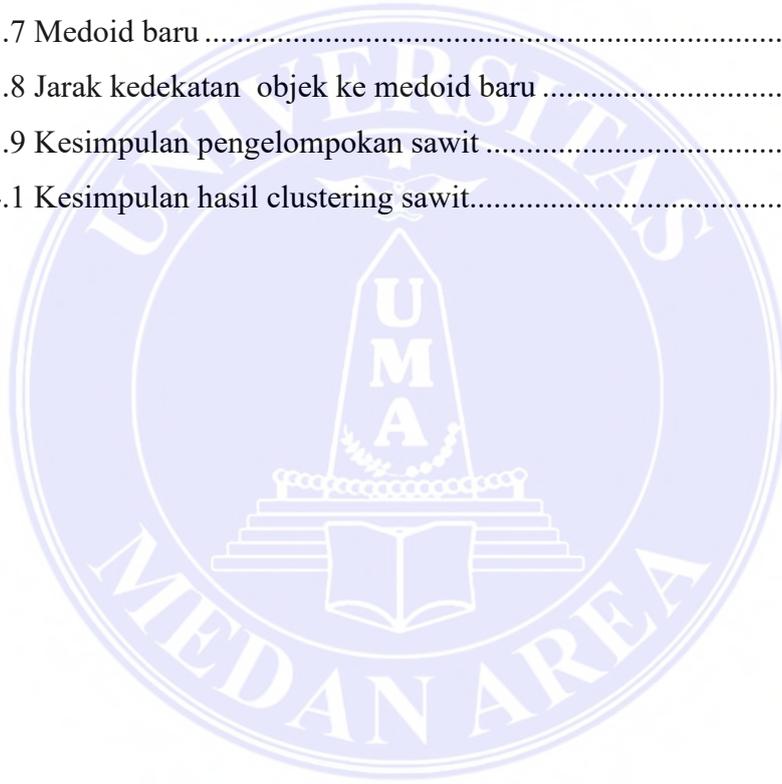
DAFTAR GAMBAR

Gambar 2.1 Aliran tahapan data mining	10
Gambar 2. 2 Proses KDD dalam Data Mining.....	12
Gambar 3.1 Tahapan dalam penelitian.....	23
Gambar 3.2 Flowchart K-medoids clustering	28
Gambar 4.1 Upload Data.....	38
Gambar 4.2 Inisialisasi Medoid	39
Gambar 4.3 Kedekatan data setiap kluster	39
Gambar 4.4 Penetapan kluster.....	40
Gambar 4.5 Pembaruan medoids baru	40
Gambar 4.6 Tahapan iterasi	41
Gambar 4.7 Iterasi Berhenti pada iterasi ke 7	41
Gambar 4.8 Grafik skor silhouette	41
Gambar 4.9 Hasil clustering setiap data.....	42
Gambar 4.10 Hasil kluster sawit	43



DARTAR TABEL

Tabel 2.1 Perbandingan varietas sawit.....	8
Tabel 2.2 Penelitian terdahulu.....	20
Tabel 3.1 Variabel Perhitungan.....	29
Tabel 3.2 Data Penelitian	29
Tabel 3.3 Data tanpa Transformasi	30
Tabel 3.4 Normalisasi data.....	31
Tabel 3.5 Medoid awal.....	32
Tabel 3.6 Jarak kedekatan pbjek ke medoid sementara	32
Tabel 3.7 Medoid baru	33
Tabel 3.8 Jarak kedekatan objek ke medoid baru	34
Tabel 3.9 Kesimpulan pengelompokan sawit	35
Tabel 4.1 Kesimpulan hasil clustering sawit.....	44



BAB I

PENDAHULUAN

1.1 Latar Belakang

Kelapa sawit merupakan salah satu komoditas utama tanaman perkebunan Indonesia sebagai penghasil devisa Negara. Jenis tanaman sawit ini menduduki posisi terpenting disektor pertanian, karena kelapa sawit mampu menghasilkan nilai ekonomi yang besar tiap hektarnya jika dibandingkan dengan tanaman penghasil minyak lainnya. Selain buah kelapa sawit juga memiliki manfaat lain seperti bahan bakar, bahan makanan, kosmetik dan pupuk. Indonesia berpeluang untuk mengembangkan pertanian sawit karena permintaan tiap tahun mengalami peningkatan yang cukup besar baik dari dalam maupun luar Negeri (Nurmalita & Bowo, 2019). Tanaman kelapa sawit di Indonesia pertama kali di perkenalkan oleh pemerintah Belanda pada tahun 1880. Kelapa sawit adalah tanaman biji minyak penting yang berasal dari Afrika Barat (Fernando dkk. 2020). Kelapa sawit awalnya ditanam sebagai tanaman hias. Penanaman pohon secara komersial baru dimulai pada tahun 1911. Penanaman yang dipimpin oleh Adrien Hallet dari Belgia. Penanaman yang dipimpin oleh Adrian Hallet ini diikuti oleh K.Shadt yang menandai lahirnya perkebunan kelapa sawit pertama di pantai timur Sumatera dengan luas 5.123 hektar (Saputra dkk. 2020). Keluhan-keluhan dari pengusaha sawit karena sering mendapatkan sedikit keuntungan atau malah mendapat kerugian karena hasil timbangan lapangan dan timbangan pabrik berbeda. Penyusutan berat dan kualitas dikarenakan buah sawit yang sudah restan membuat pihak pengolahan (pabrik) tidak dapat menerima buah sawit tersebut, dan tentu ini berdampak pada hasil yang didapatkan oleh pengusaha sawit.

Data Mining untuk mengolah data yang besar dan membantu untuk mengambil keputusan. Data mining adalah serangkaian proses untuk menggali nilai informasi yang selama ini tidak di ketahui secara manual dari suatu basis data (Boy, 2020). Informasi yang dihasilkan diperoleh dengan cara mengekstraksi dan mengenali pola yang penting atau menarik dari data yang terdapat dalam basis data. Algoritma *K-Medoids* dapat diterapkan pada pengelompokan hasil pengiriman sawit masyarakat berdasarkan berat dan kualitas pada setiap kali panen. Data hasil pengelompokan tersebut dapat diketahui *cluster* sedang, dan *cluster* tinggi.

K-Medoids merupakan suatu algoritma yang digunakan untuk menemukan *medoids* di dalam sebuah kluster yang merupakan titik pusat dari suatu kelompok. (Agustini dkk. 2022). Variable yang digunakan pada penelitian ini adalah kualitas dan berat sawit. Dengan mengidentifikasi masalah, mempelajari literatur, mengumpulkan masalah, mengolah data, menguji data, dan menyimpulkan hasil penelitian.

Penelitian terdahulu yang berjudul Pengelompokan Data Janjang Panen Kelapa Sawit Menggunakan Algoritma *K-Medoids* Pada PT SIR MANDAU (Agustini dkk. 2022) dengan hasil penelitian yaitu perhitungan algoritma *K-Medoids* dengan pengujian *tools google colab 5.3*. diperoleh hasil yang sama dengan perhitungan manual.

1.2 Rumusan Masalah

Berdasarkan latar belakang ditemukan beberapa keluhan dari pengusaha sawit karena tidak mengetahui seberapa banyak sawit yang diterima dari kalangan masyarakat yang memiliki kualitas yang bagus dan kurang bagus. Maka rumusan masalah pada penelitian ini adalah bagaimana menganalisis algoritma *K-Medoids* untuk pengelompokan hasil panen sawit berdasarkan berat dan kualitas yang akan diterapkan di *google colab*.

1.3 Batasan Masalah

Batasan masalah dalam penelitian ini adalah:

1. Data yang digunakan dalam proses penelitian ini merupakan data primer yang diperoleh dari Cv. Ram Bintang Motor. Data yang digunakan merupakan data tanggal 3 - 29 maret 2023.
2. Variabel yang digunakan dalam mengelompokkan hasil panen sawit adalah *Nett1*, dan *Sortasi_Persen*.
3. Hasil pada penelitian ini akan mendapatkan hasil cluster rendah, sedang, dan tinggi.

1.4 Tujuan Penelitian

Adapun tujuan dari penelitian ini dilakukan anatara lain:

1. Menguji Algoritma *K-Medoids* untuk di lakukan penelitian, apakah algoritma tersebut bisa untuk dilakukan sebuah *clustering* atau pengelompokan hasil panen sawit.
2. Untuk mengelompokkan pengiriman hasil sawit masyarakat berdasarkan berat dan kualitas guna mempertimbangkan tujuan pabrik pengiriman buah sawit.

1.5 Manfaat Penelitian

Adapun manfaat yang dapat di peroleh dari penelitian ini yaitu:

1. Untuk mengetahui kualitas sawit agar bisa melakukan pengiriman ke pabrik yang tepat.
2. Untuk membantu pihak terkait seperti para pengusaha sawit dan masyarakat agar bisa meminimalisir kerugian akibat surut tonase dan kualitas sawit yang buruk.



1.6 Sistematika Penulisan

Sistematika penulisan pada penelitian ini ditunjukkan kepada pembaca untuk lebih memahami isi dari penelitian ini. Berikut sistematis penulisan penelitian ini secara garis besar yaitu :

BAB I PENDAHULUAN

Pendahuluan membahas seputar latar belakang dari penelitian ini, dimana untuk mengetahui sebab penelitian ini dilakukan, selanjutnya membahas rumusan masalah, tujuan penelitian dan manfaat penelitian.

BAB II TINJAUAN PUSTAKA

Tinjauan Pustaka membahas hal-hal yang mendasar berisi teori-teori yang berkaitan dengan analisis metode *clustering* dengan menggunakan algoritma *k-medoids* pengelompokan pengiriman hasil sawit masyarakat berdasarkan berat dan kualitas dan hal-hal yang mendukung dalam melakukan penelitian.

BAB III METODE PENELITIAN

Metode penelitian membahas tentang sumber data tahapan yang dilakukan dalam melakukan penelitian ini serta gambaran umum system yang akan diselesaikan.

BAB IV HASIL DAN PEMBAHASAN

Hasil dan pembahasan menampilkan hasil dari penelitian yang telah dilakukan.

BAB V KESIMPULAN DAN SARAN

Kesimpulan dan saran adalah menarik kesimpulan dari hasil penelitian yang telah dilakukan dan memberikan saran-saran yang nantinya berguna untuk penelitian selanjutnya.

BAB II

TINJAUAN PUSTAKA

2.1 Hasil Panen Sawit

Penen hasil kelapa sawit adalah proses pemotongan tandan buah segar (TBS) dari pohon sampai dengan pengangkutan ke pabrik sawit. Tandan buah segar hasil panen tersebut harus segera diangkut ke pabrik sebab semakin lama proses pengiriman maka kadar minyaknya semakin berkurang sehingga dapat menyebabkan hasil yang diterima petani juga semakin berkurang (Thoharudin dkk. 2020). Keterlambatan dalam pengangkutan hasil sawit akan mempengaruhi proses pengolahan dan mutu akhir. Kandungan FFA dalam buah akan terus meningkat seiring lamanya buah tidak terangkut. Sangat penting untuk mempertahankan panen pada interval yang pendek pada tanaman muda, karena buah akan membrondol lebih dari 10% dalam waktu 5-7 hari, interval panen yang lama mengakibatkan banyaknya buah busuk dan jumlah brondolan yang banyak (Agustini dkk. 2022). Upaya untuk mengoptimalkan produksi maka kegiatan panen, angkut harus dikelola dengan baik, karena kegiatan panen, angkut, olah merupakan kegiatan yang tidak terpisah dan menentukan kualitas panen.

Sangat penting untuk mempertahankan panen pada interval yang pendek pada tanaman muda, karena buah akan membrondol lebih dari 10% dalam waktu 5-7 hari, interval panen yang lama mengakibatkan banyaknya buah busuk dan jumlah brondolan yang banyak. Upaya untuk mengoptimalkan produksi maka kegiatan panen angkut harus dikelola dengan baik, karena kegiatan angkut, panen, olah merupakan kegiatan yang tidak terpisah dan menentukan kualitas panen. Hal ini yang menjadi acuan pemberian nilai kualitas oleh pihak pembeli (pabrik) kepada

petani dengan memberikan sortasi persen sebagai simbol.

Secara umum pengangkutan atau pengiriman buah sawit dibagi menjadi dua yaitu: (1) pengumpulan buah dari pohon ke tempat yang bisa dijangkau mobil (TPH) dan kedua pengiriman buah sawit dari TPH ke pabrik pengolahan sawit (PKS). Transportasi menjadi hal penting dalam menyalurkan hasil panen dari tempat penampungan hingga menuju loading ramp di pabrik (Waskitha dkk. 2018). Kendala yang sering ditemui dalam proses pengiriman hasil sawit masyarakat berupa perubahan cuaca membuat berat dan kualitas menurun yang dapat merugikan pemilik sawit dan pengangkut mendapat kerugian terutama kepada pihak pengusaha.

2.2 Kualitas dan berat Sawit

Beragamnya umur tanaman kelapa sawit di Provinsi Sumatera Utara menjadi suatu kendala yang sering dijumpai dikalangan masyarakat, karena masing-masing umur tanaman memiliki produksi dan kualitas buah yang berbeda. produksi minyak kelapa sawit mulai meningkat saat umur tanaman 4-6 tahun dan akan mencapai produksi maksimal pada saat umur tanaman mencapai 8-10 tahun (Hasibuan dkk. 2019). Kualitas buah sawit pada saat masih muda memiliki kualitas yang rendah, hal ini dilihat dari kandungan minyak sawit yang masih rendah, berat buah yang masih rendah, dan produksi buah yang masih kecil. Factor penting yang harus terpenuhi dalam mengelola usaha tani, menurut (Chaira dkk. 2022) petani perlu menguasai penggunaan atau pengolahan factor produksi yang diperlukan secara benar, dengan jumlah standart rasio yang ideal dan dengan memperhitungkan kebutuhan input secara tepat.

Menurut (Sari dkk. 2022) Estimasi kematangan tandan buah segar kelapa

sawit merupakan komponen penting dalam pengelolaan panen kelapa sawit, karena akan mengarah pada profitabilitas dan daya jual produk. Laju ekstraksi minyak (OER) tandan buah segar (TBS) kelapa sawit sangat bergantung pada tahap kematangannya TBS kelapa sawit yang matang akan menghasilkan OER yang tinggi sedangkan TBS kelapa sawit yang kurang matang akan menghasilkan lebih sedikit minyak. Jenis-jenis sawit yang sering didapati pada perkebunan kelapa sawit di Sumatera Utara adalah jenis dura, pisifera, tenera. Pada umumnya jenis dura memiliki ciri-ciri daging buah relative tipis dan cangkang yang tebal jika dibandingkan dengan jenis lainnya sehingga dianggap dapat memperpendek umur mesin pengolahan namun biasanya tandan buahnya besar dan kandungan minyak per tandannya berkisar 18%, jenis fisifera memiliki ciri-ciri fisik ketebalan cangkang sangat tipis bahkan hampir tidak ada, sedangkan pada jenis sawit tenera daging buah sangat tebal (Hasibuan, 2023).

Tabel 2.1 Perbandingan varietas sawit

Varietas	Descripsi
Dura	Cangkang tebal (2-8 mm) Daging buah relatif tipis , yaitu 35-50% terhadap buah
Pisifera	Ketebalan cangkang sangat tipis, bahkan hampir tidak ada, daging buah tebal,
Tenera	Tempurung tipis (0,5-4 mm) Daging buah sangat tebal (60-96% dari buah) Tandan buah lebih banyak, tetapi ukurannya relative lebih kecil

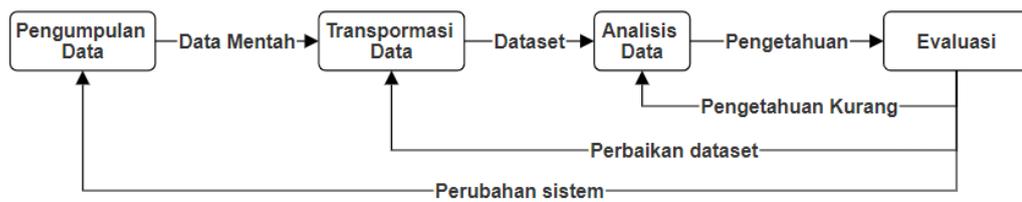
2.3 Ram Bintang Motor

Ram Bintang Motor suatu perusahaan industry yang bernaung dalam bidang perkebunan, yakni tempat penjualan Tandan Buah Segar (TBS) yang dihasilkan oleh hasil perkebunan sawit masyarakat dan Pt. Instansi ini berlokasi di Desa Binanga Kecamatan Barumun Tengah Kabupaten Padang Lawas Provinsi Sumatera Utara.

2.4 Data Mining

Data mining adalah suatu proses pengerukan atau pengumpulan informasi penting dari suatu data yang besar. Menurut (Yansa dkk, 2023) *Data mining* bekerja dengan menemukan pola dan informasi dari data terpilih menggunakan metode tertentu. *Data mining* biasa disebut *knowledge discovery in database* (KDD), merupakan proses yang meliputi pengumpulan, pemakaian data *historis* guna menemukan keteraturan serta pemakaian *data historis* untuk menemukan keteraturan dalam set data berukuran besar. Keluaran pada *data mining* dapat dipakai untuk memperbaiki pengambilan keputusan dimasa mendatang (Yansa dkk, 2023). *Data mining* adalah proses menemukan pola dan pengetahuan menarik dari data dalam jumlah yang besar.

Istilah *data mining* memiliki beberapa pandangan, seperti *knowledge discovery* ataupun *pattern recognition*. Istilah *knowledge discovery* atau penemuan pengetahuan tepat digunakan karena tujuan utama dari data mining memang untuk mendapatkan pengetahuan yang masih tersembunyi di dalam bongkahan data. Sedangkan istilah untuk *pattern recognition* atau pengenalan pola tepat untuk digunakan karena guna menemukan pola yang tersembunyi di dalam bongkahan data (Nabila dkk, 2021).



Gambar 2.1 Aliran tahapan data mining

Secara umum defenisi *data mining* dapat diartikan sebagai berikut:

1. Proses penemuan pola yang menarik dari data yang tersimpan dalam jumlah besar.
2. Ekstrasi dari suatu informasi yang berguna atau menarik (non-trivial, implicit, sebelumnya belum diketahui potensial kegunaannya) pola atau pengetahuan dari data yang tersimpan dalam jumlah besar.
3. Eksplorasi dari Analisa secara otomatis atau semiotomatis terhadap data-data dalam jumlah besar untuk mencari pola aturan yang berarti

Berikut adalah Langkah-langkah yang diperlukan dalam proses data mining

1. *Identify the Business Problem*

Hal utama adalah mengetahui masalah bisnis yang dihadapi, karena data tidak dapat diolah jika belum mengetahui ada permasalahan yang dihadapi, dengan mengetahui masalah yang dihadapi kemudian dapat ditentukan data-data apa saja yang dibutuhkan untuk dapat dilakukan tahap Analisa.

2. *Mine the Data For actionable Infromatios*

Langkah selanjutnya adalah dengan mengidentifikasi masalah.

Kumpulkan semua data yang diperlukan untuk tujuan analitis. Baru setelah itu tahap analisis data dilakukan dan wawasan baru diperoleh dari analisis ini untuk pengambilan keputusan.

3. *Take the Actios*

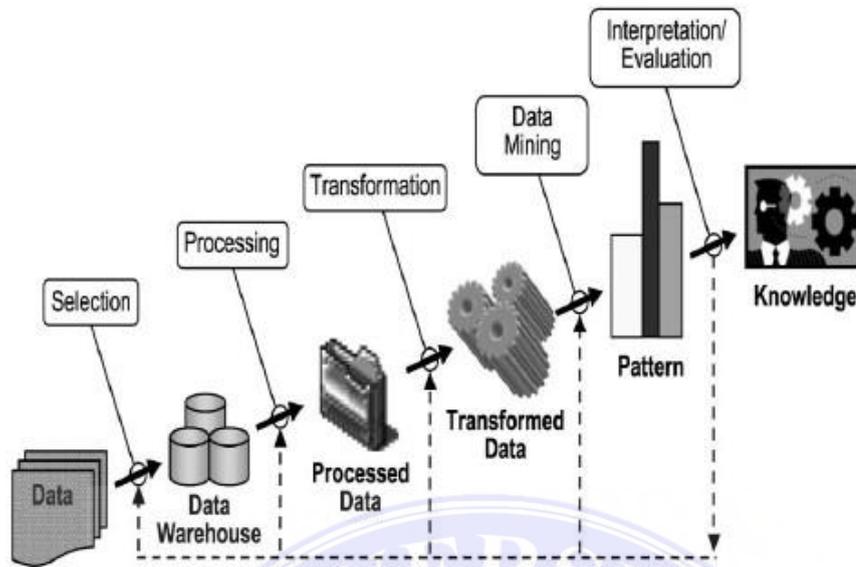
Keputusan dan pedoman yang diperoleh dari proses data mining diimplementasikan dalam bentuk langkah-langkah konkrit dalam proses bisnis.

4. Hasil Pengukuran

Selanjutnya, kita perlu memantau hasil dari langkah-langkah yang dijalankan. Apakah memenuhi tujuan yang diharapkan dan mengatasi masalah yang ada

2.5 Data Mining Sebagai Proses Dalam *Knowledge Discovery in Data*

Menurut (Lesmana, 2021) *Knowledge Discovery in Database* (KDD) ini adalah keseluruhan proses kritis untuk menemukan dan mengidentifikasi pola dalam data yang valid, baru, berguna, dan dapat dipahami. KDD berurusan dengan teknik integrasi dan penemuan ilmiah, interpretasi, dan visualisasi pola dalam satu set kumpulan data. *Knowledge Discovery in Database* (KDD) adalah suatu ilmu dari percabangan ilmu statistik, database, AI, visualisasi dan komputer paralel yang mempengaruhi pengetahuan *interdisciplinary knowledge* yang melibatkan hasil ekstraksi kecenderungan suatu pola, sehingga mengubah hasilnya secara tepat dan akurat serta menjadi informasi yang mudah dipahami (Oktarian dkk, 2020)



Gambar 2. 2 Proses KDD dalam *Data Mining*

Proses dalam KDD Data Mining adalah sebagai berikut:

1. *Data Selections*

Buat set data dan pilih set data target atau fokuskan pada subset variabel atau sampel data untuk ditemukan. Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses data mining disimpan dalam suatu berkas, terpisah dari basis data operasional (Oktarian dkk, 2020).

2. *Pre-processing/Cleaning*

Pemrosesan pendahuluan dan pembersihan data merupakan operasi dasar seperti penghapusan noise dilakukan. Sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses cleaning pada data yang menjadi focus KDD. Proses cleaning mencakup antara lain membuang duplikasi data seperti kesalahan cetak (tipografi) (Oktarian dkk, 2020)

3. *Transformation*

Pencarian fitur-fitur yang berguna untuk mempresentasikan data bergantung kepada goal yang ingin dicapai. Merupakan proses transformasi pada data yang telah dipilih sehingga data tersebut sesuai untuk proses data mining. Proses ini merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data (Lesmana, 2021).

4. *Data Mining*

Pemilihan tugas data *mining* pemilihan goal dari proses KDD misalnya *klasifikasi*, *regresi*, *clustering*, dan lain-lain. Pemilihan algoritma data mining untuk pencarian (*searching*). Proses *Data Mining* yaitu proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan (Oktarian dkk, 2020).

5. *Evaluasi Interpretasi*

Transformasi pola yang dihasilkan oleh penambangan data. Sampel informasi yang dihasilkan dari proses data mining harus disajikan dalam format yang mudah dipahami oleh pemangku kepentingan. Fase ini merupakan bagian dari proses KDD dan melibatkan melihat apakah ada pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya. (Oktarian dkk, 2020).

2.6 Tugas-tugas Data Mining

Tugas-tugas data mining menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu:

1. Deskripsi

Deskripsi adalah pemaparan atau penggambaran suatu objek dengan jelas terhadap suatu masalah dalam konteks data mining

2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variable target lebih ke arah numeric dari pada ke arah kategori. Model dibangun menggunakan record lengkap yang menyediakan nilai dari variable target sebagai nilai prediksi

3. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan datang di masa mendatang.

4. Klasifikasi

Dalam klasifikasi, terdapat target variable kategori. Sebagai contoh penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah (Setio dkk, 2020).

5. Pengklusteran

Clustering merupakan suatu metode yang digunakan untuk mengelompokkan dan mencari suatu data yang memiliki karakteristik antara satu sama lain. Selain itu *clustering* juga salah satu metode *data mining* yang bersifat *unsupervised* yang artinya metode ini diterapkan tanpa adanya target untuk outputnya nanti. Terdapat dua jenis metode *clustering* yang sering digunakan dalam pengelompokan data yaitu *hierarchical clustering* dan *non-hierarchical clustering*. Metode pengelompokan data, seperti *clustering*, memainkan peran penting dalam analisis data dan penambangan informasi. *Clustering* adalah teknik yang digunakan untuk mengelompokkan data menjadi beberapa kelompok atau kluster berdasarkan

kesamaan karakteristik atau pola yang dimiliki oleh data tersebut. Tujuan utama clustering adalah mengidentifikasi pola tersembunyi dalam data dan mengelompokkan data yang serupa ke dalam kelompok yang sama. Pengelompokan data melalui teknik *clustering* dapat memberikan pemahaman yang lebih baik tentang pola dan hubungan antara data yang serupa, yang pada gilirannya dapat digunakan untuk pengambilan keputusan, segmentasi pasar, pengenalan pola, dan banyak aplikasi lainnya dalam berbagai bidang seperti ilmu sosial, ilmu kesehatan, dan ilmu komputer.

6. Asosiasi

Tugas asosiasi dalam data mining adalah menemukan atribut yang muncul bersamaan. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja.

2.7 Algoritma K-MEDOIDS

Algoritma PAM (*Partitioning Around Medoids*) atau biasa juga disebut dengan algoritma *K-Medoids*, merupakan algoritma yang diwakili oleh cluster yaitu medoid. Perbedaan antara algoritma *K-Medoids* dengan algoritma *K-Means* yaitu algoritma *K-Medoids* menggunakan objek sebagai perwakilan (medoid) pusat cluster untuk tiap *cluster*, sementara algoritma *K-Means* membutuhkan nilai rata-rata (mean) sebagai pusat cluster (Harahap, 2021). *K-Medoids* merupakan salah satu metode clustering yang berfungsi untuk memecah dataset menjadi kelompok-kelompok (Andini & Arifin, 2020). Kelebihan algoritma *K-Medoids* adalah hampir dapat bekerja pada setiap jenis data matriks dan mampu mengatasi outlier. Metode *K-Medoids* cukup efisien untuk data set yang kecil, Kelebihan dari metode *K-Means* yang sensitive terhadap outlier dan hasil clustering tidak tergantung pada urutan masuk dataset (Agustini dkk, 2022).

Tahapan algoritma *K-Medoids* adalah sebagai berikut :

1. *Preprocessing* data
 - a. *Cleaning* data Membersihkan data
 - b. *Transformasi* data
 - c. Mengubah data yang non numeric menjadi numeric
 - d. *Normalisasi* data
 2. Inisialisasi pusat *Cluster* sebanyak k (jumlah cluster)
 3. Kumpulkan setiap data (objek) ke cluster terdekat dengan menggunakan rumus menghitung jarak *euclidian distance* berikut ini :

$$D(X_2, X_1) = ||X_2 - X_1||_2 = \sqrt{\sum^p |X_2 - X_{ij}|^2} \dots\dots\dots(2.1)$$

4. Pilih secara acak objek pada masing-masing *cluster* sebagai kandidat *medoid* baru
5. Hitung jarak setiap objek yang berada pada masing-masing *cluster* dengan kandidat *medoid* baru.
6. Hitung total simpangan (S)dengan menghitung nilai total *distance* baru –total *distance* lama. Jika $S < 0$, maka tukar.
7. objek dengan data *cluster* untuk membentuk sekumpulan objek baru sebagai *medoid*.
8. Ulangi langkah 3 sampai 5 hingga tidak terjadi perubahan *medoid*, sehingga didapatkan *cluster* beserta anggota *cluster* masing-masing.

2.8 Silhouette

Validasi indeks *silhouette* merupakan metode yang digunakan untuk melihat kualitas dari sebuah *cluster*. Validasi indeks *silhouette* pertama kali diperkenalkan sebagai teknik yang digunakan untuk mengevaluasi objek secara visual yang berada baik didalam *cluster* maupun yang berada diluar *cluster*, berdasarkan rentang nilai *silhouettenya* (Okta dkk, 2019). Nilai *silhouette* objek mendefinisikan kedekatannya dengan *clusternya* sendiri dengan kelompok lain.

Rentang nilai *silhouette* dirumuskan sebagai berikut:

$$s_i = \frac{b_i - a_i}{\max[b, a]} \quad (2.2)$$

s_i : nilai koefisien *silhouette* objek ke-i dengan $i = 1, 2, \dots, n$

a_i : rata-rata Jarak antara objek ke-i dengan objek lainnya dalam satu *cluster*

b_i : minimum rata-rata Jarak antara objek ke-i dengan objek lainnya di masing-masing *cluster*.

Jika nilai koefisien *silhouette* < 0 maka objek berada dalam kelompok yang salah, sedangkan jika nilai koefisien *silhouette* > 0 maka objek berada dalam kelompok yang benar dan jika nilai koefisien *silhouette* $= 0$ maka objek berada diantara dua *cluster* sehingga objek tersebut belum dapat ditentukan masuk didalam kelompok yang benar atau salah. Untuk menentukan rentang objek secara keseluruhan dapat dirumuskan sebagai berikut:

$$SC = \frac{\sum_{i=1}^n s_i}{n} \quad (2.3)$$

dengan:

SC : rentang nilai dari koefisien *silhouette*

n : banyaknya data

Ketika rentang nilai dari koefisien *silhouette* > 0,51 menunjukkan bahwa struktur hasil pengelompokan sudah dianggap baik, sedangkan ketika rentang nilai dari koefisien *silhouette* > 0,71 menunjukkan bahwa struktur hasil pengelompokan dianggap sangat baik.

2.9 Normalisasi Min-Max

Normalisasi *min-max* adalah metode yang digunakan dalam analisis data untuk mengubah nilai-nilai dalam suatu atribut menjadi rentang yang lebih kecil, seperti 0 hingga 1. Tujuan dari normalisasi *min-max* adalah untuk menjaga konsistensi dan kesamaan skala antara atribut-atribut yang memiliki rentang nilai yang berbeda.

Pada penelitian kali ini, Peneliti menggunakan normalisasi *Min-Max*. Normalisasi dilakukan bertujuan untuk mendapatkan nilai jarak yang lebih kecil antar atribut dengan menentukan rentang minimum dan maksimum. Rumus yang digunakan untuk menghitung normalisasi *Min-Max* adalah sebagai berikut:

$$X_{ni} = \frac{x - \min A}{\max A - \min A} \dots \dots \dots (2.4)$$

Diketahui : X_{ni} = nilai hasil normalisasi

x = data yang akan dinormalisasi

$\min A$ = nilai minimum dari data asli

$\max A$ = nilai maksimum dari data asli

2.10 Google Colab

Google Colaboratory atau disebut juga dengan *google colab* merupakan sebuah platform yang dimiliki oleh *google* dan bertujuan untuk memudahkan pengguna dalam melakukan pekerjaan terutama yang berkaitan dengan *data science* dan *machine learning*. *Google colab* ini menggunakan bahasa pemrograman *Python* yang telah dikonfigurasi sebelumnya (Kurniawan dkk, 2023). Didalamnya juga terdapat beberapa *library* seperti *TensorFlow*, *Matplotlib* dan *Keras*. Mesin virtual akan dinonaktifkan secara otomatis dalam jangka waktu tertentu serta konfigurasi pengguna hilang. Namun, untuk notebook tetap ada tidak ikut menghilang. *Google Colaboratory* ini dihosting di platform *Google Cloud*.

2.11 Penelitian Terdahulu

Berikut merupakan beberapa penelitian terdahulu yang menjadi referensi pada penelitian analisis metode *clustering* menggunakan algoritma *K-Medoids* pengelompokan pengiriman hasil sawit masyarakat berdasarkan berat dan kualitas, yaitu:

Tabel 2.2 Penelitian terdahulu

No	Penulis (tahun)	Topik	Hasil
1	Lila Agustini, Sumarno , Ika Okta Kirana (2022)	Pengelompokan Data Janjang Panen Kelapa Sawit Menggunakan Algoritma <i>K-Medoids</i> Pada PT SIR MANDAU	Pengelompokan data janjang panen kelapa sawit menggunakan algoritma <i>k-medoids</i> menunjukkan bahwa data janjang panen pada PT SIR MANDAU pada tahun 2020 berhasil mengalami peningkatan panen yang tinggi.
2	Deny Pranata Pasaribu, Irfan Sudahri Damanik, Eka Irawan, Suhada, Heru Satria Tambunan (2021)	Memanfaatkan Algoritma <i>K-means</i> dalam Memetakan Potensi Hasil Produksi Kelapa sawit PTPN IV Marihata	Hasil dari penelitian menggunakan algoritma <i>K-means</i> ini berupa 2 buah <i>cluster</i> , yaitu : <i>cluster</i> tinggi dan <i>cluster</i> rendah. Penerapan algoritma <i>k-means</i> dapat diimplementasikan terhadap pemetaan hasilproduksi buah sawit di PTPN IV Marihat.
3	Beri Lesmana (2021)	Pengelompokan Pengiriman Hasil Sawit Berdasarkan Tonase dan Kwalitas Menggunakan Metode <i>Clustering</i>	Berdasarkan pengujian metode <i>Clustering</i> dengan algoritma <i>k-means</i> ini, dapat diketahui kelompok pabrik, kelompok tonase dan kelompok kwalitas mana saja memiliki kelompok paling tinggi

4	Siti Hajar, Aril Andi, Novany, Agus Perdana Wirdanto, Anjar Wanto, Eka Irawa (2020)	Penerapan <i>K-Means Clustering</i> Pada Ekspor Minyak Kelapa Sawit Menurut Negara Tujuan	Dapat Menentukan Ekspor Minyak Kelapa sawit, data yang digunakan pada penelitian ini adalah data tahun 2010, 2011, 2012, 2013, 2014, dan data 2015
5	Agustina Lili, Suhada, Saputra Widodo (2022)	Pengelompokan Hasil Panen Kelapa Sawit Dalam Produksi Per Blok Menggunakan Algoritma <i>K-Means</i>	Data diolah menggunakan Microsoft Excel untuk ditentukan nilai centroid dalam 3 cluster yaitu cluster tinggi, cluster sedang dan cluster rendah

BAB III

METODE PENELITIAN

3.1 Subjek Dan Objek Penelitian

Subjek penelitian ini adalah data pengiriman sawit masyarakat pada CV Ram Bintang Motor. Objek penelitian ini adalah melakukan pengelompokan hasil panen sawit untuk menentukan buah kelapa sawit yang rendah, sedang, dan tinggi.

3.2 Alat Penelitian

Alat yang digunakan untuk pengembangan penelitian yaitu berupa perangkat keras dan perangkat lunak.

3.2.1 Perangkat Keras

Perangkat keras yang digunakan untuk melakukan proses pengembangan penelitian yaitu sebuah laptop-2FNH12M0 dengan ram 4,00 GB,dan Processor AMD Athlon Silver 3050U.

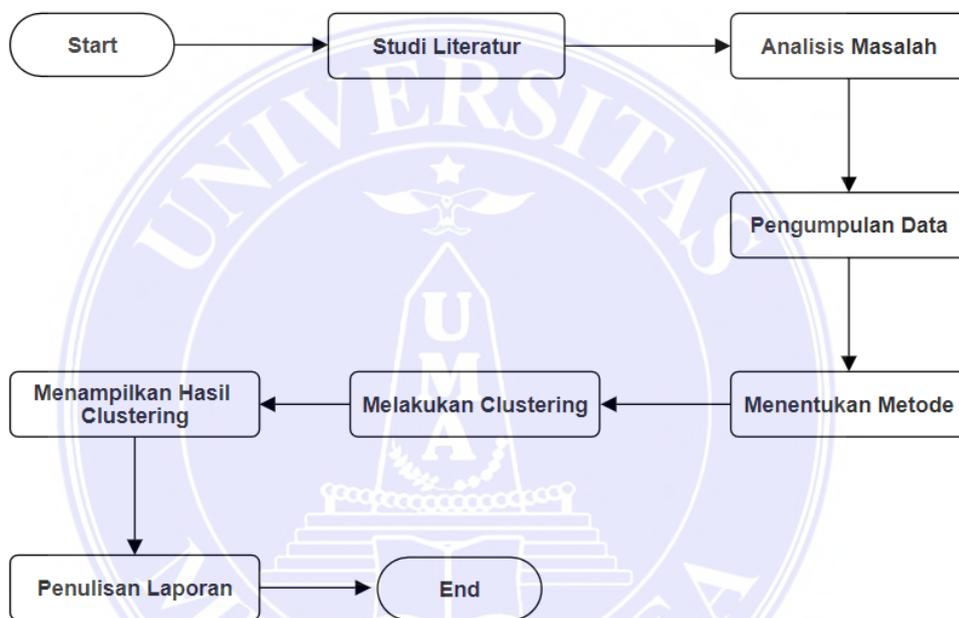
3.2.2 Perangkat Lunak

Perangkat lunak yang digunakan pada saat proses penelitian yaitu:

1. Sistem Operasi Windows 10
2. *Browser berupa Mozilla Firefox*
3. *Microsoft Excel*
4. *Paint*
5. *Google Colab*

3.3 Prosedur Penelitian

Penelitian ini menggunakan prosedur kerja yang meliputi beberapa prosedur yang di mulai dengan melakukan persiapan penelitian, Studi Literatur, Analisis masalah, pengumpulan data, menentukan metode penelitian, melakukan *clustering*, menampilkan hasil klaster, dan penulisan laporan Untuk lebih jelas, berikut tahapan penelitian yang dilakukan dapat dilihat pada gambar 3.1



Gambar 3.1 Tahapan dalam penelitian

Penelitian ini dimulai dengan melakukan studi literatur untuk memperoleh pemahaman yang mendalam tentang topik penelitian dan penelitian terkait yang telah dilakukan sebelumnya (Sugiono dkk, 2021). Studi literatur ini bertujuan untuk membangun dasar pengetahuan yang kuat sebelum melangkah ke tahap berikutnya.

Setelah studi literatur selesai, langkah selanjutnya adalah melakukan analisis masalah. Pada tahap ini, peneliti menganalisis permasalahan yang akan diteliti dengan mengidentifikasi faktor-faktor yang relevan dan menentukan tujuan penelitian secara lebih spesifik. Analisis masalah ini membantu peneliti dalam merumuskan pertanyaan penelitian yang jelas dan tujuan yang ingin dicapai.

Setelah itu, peneliti melanjutkan ke tahap pengumpulan data yang diperlukan untuk penelitian. Pengumpulan data dapat dilakukan melalui berbagai metode, seperti survei, wawancara, observasi, atau pengumpulan data sekunder. Data yang dikumpulkan harus relevan dengan tujuan penelitian dan dapat mendukung analisis clustering yang akan dilakukan.

Selanjutnya, peneliti menentukan metode penelitian yang akan digunakan. Dalam konteks ini, metode penelitian yang dipilih adalah metode *clustering* menggunakan metode *K-Medoids*. Pemilihan metode penelitian harus didasarkan pada karakteristik data dan tujuan penelitian yang telah ditetapkan sebelumnya.

Setelah metode penelitian ditentukan, peneliti melaksanakan proses clustering menggunakan metode *K-Medoids*. Metode ini digunakan untuk mengelompokkan data ke dalam kluster berdasarkan kesamaan karakteristik. Proses clustering melibatkan perhitungan jarak antara data dan pemilihan pusat kluster yang optimal.

Setelah clustering selesai, hasil clustering perlu ditampilkan agar dapat dipahami dengan jelas. Hal ini dapat dilakukan melalui visualisasi data, seperti diagram scatter plot atau diagram dendrogram, yang membantu dalam

memperlihatkan pola dan hubungan antar data dalam kluster yang terbentuk.

Terakhir, peneliti melakukan penulisan laporan yang berisi hasil penelitian, analisis data, kesimpulan, dan saran. Laporan penelitian tersebut harus disusun dengan jelas dan sistematis, mengikuti struktur yang telah ditetapkan.

3.3.1 *Studi Literatur*

Melakukan studi literatur merupakan langkah penting dalam penelitian untuk mendapatkan pemahaman mendalam tentang landasan teori yang akan digunakan. Dalam konteks pengelompokan data menggunakan metode *K-Medoids Clustering*, peneliti perlu mengacu pada sumber-sumber literatur yang relevan yang membahas tentang metode *clustering*, khususnya *K-Medoids*, serta konsep dan teori yang terkait. Sumber literatur yang dapat digunakan oleh peneliti meliputi buku, jurnal ilmiah, artikel, makalah konferensi, dan sumber informasi terpercaya lainnya. Peneliti dapat mencari literatur melalui basis data akademik, perpustakaan digital, jurnal ilmiah terkemuka, dan mesin pencari seperti *Google Scholar*.

Dalam *literatur*, peneliti dapat menemukan penjelasan tentang konsep dasar *clustering*, *algoritma K-Medoids*, teknik pengukuran jarak, evaluasi *clustering*, serta aplikasi dan penelitian terkait yang telah dilakukan sebelumnya. Buku teks tentang data *mining*, *machine learning*, atau *statistika* juga dapat memberikan pemahaman yang komprehensif tentang metode *clustering* dan konsep yang terkait. Selama proses *studi literatur*, peneliti perlu membaca, menganalisis, dan menyimpulkan informasi yang relevan untuk memperoleh

pemahaman yang mendalam tentang landasan teori dan konsep-konsep yang

terkait dengan pengelompokan data menggunakan metode *K-Medoids Clustering*. Informasi yang diperoleh dari literatur ini akan menjadi dasar peneliti dalam menerapkan metode *clustering* secara efektif dan memahami *implikasi* dan *interpretasi* hasil yang dihasilkan.

Adapun sumber-sumber yang diambil untuk dijadikan *literatur* pada penelitian ini adalah yang berkaitan dengan pengiriman sawit, kualitas dan berat sawit, *data mining*, Ram Bintang Motor, metode *clustering*, algoritma *K-Medoids*, dan penelitian terdahulu.

3.3.2 Analisis Masalah

Identifikasi pada penelitian adalah langkah untuk menentukan fokus dan fokus dan tujuan penelitian. Identifikasi masalah pada penelitian melibatkan pengidentifikasian kekosongan pengetahuan atau permasalahan yang belum terpecahkan dalam bidang tertentu. Tahapan penelitian ini, penulis mengidentifikasi permasalahan yang timbul dikalangan masyarakat, khususnya pada pengusaha yang berkecimbung dibidang pertanian kelapa sawit. karena sering mendapatkan sedikit keuntungan atau malah mendapat kerugian karena hasil timbangan lapangan dan timbangan pabrik berbeda. Hal ini dikarenakan penyusutan berat dan kualitas buah sawit yang sudah restan membuat pihak pengolahan (pabrik) tidak dapat menerima buah sawit tersebut, dan tentu ini berdampak pada hasil yang didapatkan oleh pengusaha atau toke sawit. Penyebab pihak pabrik tidak dapat menerima beberapa buah adalah karena tidak sesuai dengan kriteria penerimaan buah oleh pihak pabrik, dan beberapa pabrik memang memiliki perbedaan penerimaan buah sawit masyarakat.

3.3.3 Mengumpulkan Data

Penelitian ini menggunakan data primer sebagai sumber data. Data Primer adalah data yang bersumber internal yang didapatkan secara langsung melalui pelaksanaan observasi, yaitu pengamatan secara langsung, dan lain-lain (Sugiono dkk, 2021). Menurut Yaredi (2022) Data primer adalah data yang diperoleh dari sumber informan dikumpulkan secara langsung dari sumbernya oleh peneliti itu sendiri. Pada penelitian ini peneliti akan mengambil dataset dari sebuah instansi yaitu : CV Ram Bintang Motor yang dimana data tersebut bertujuan untuk bahan sebuah penelitian. Data yang didapatkan meliputi kualitas dan berat sawit sama seperti judul pada penelitian ini, yang dimana kualitas sebagai (Sortasi_Persen) dan berat sebagai (Nett1).

a. Sortasi_Persen

Sortasi persen pada data penelitian pada CV Ram Bintang Motor dilakukan untuk mengamati mutu buah yang diterima di pabrik kelapa sawit. Semakin tinggi sortasi yang diberi tanda persen (%) maka semakin buruk kualitas sawit tersebut, sebaliknya semakin rendah sortasi pada sawit maka semakin bagus kualitasnya. Tinggi sortasi persen dikarenakan beberapa faktor yaitu : buah sawit mulai restan, banyaknya sampah yang ada pada buah sawit, terdapat pasir dan lumpur, dan masih banyak lagi alasan pemberian sortasi persen.

b. Nett 1

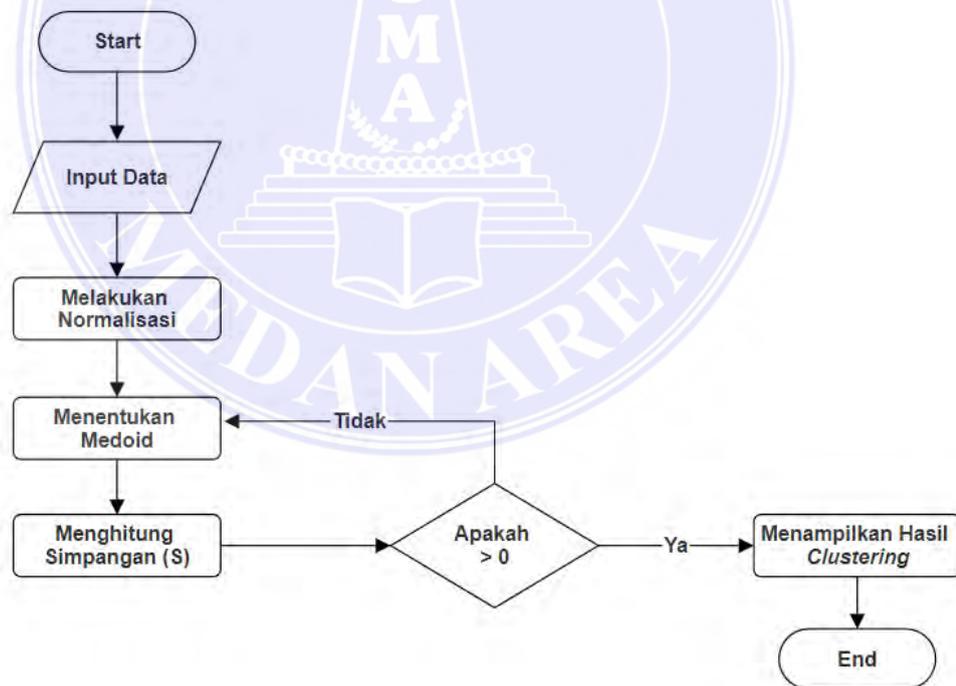
Nett 1 merupakan berat buah sawit yang didapatkan setelah dilakukan penimbangan.

3.3.4 Menentukan Metode Penelitian

Pada tahap ini, setelah peneliti melakukan pengumpulan kebutuhan maka selanjutnya menentukan metode yang akan digunakan dalam penelitian berdasarkan permasalahan yang telah ada. Pada penelitian ini penulis menggunakan metode *K-Medoids Clustering* untuk melakukan pengelompokan untuk mendapatkan hasil kelompok sawit rendah, sedang, dan tinggi.

3.3.5 Melakukan Clustering

Tahap ini, penulis akan melakukan pengolahan data yang telah diperoleh dari Cv Bintang Motor dan data yang digunakan merupakan data *real*. Berikut adalah flowchart tahapan perhitungan menggunakan algoritma *K-Medoids Clustering*, yaitu :



Gambar 3.2 Flowchart K-medoids clustering

1. Preprocessing Data

Variable penilaian yang digunakan dalam proses pemberian nilai pada setiap tandan buah sawit adalah sebagai berikut :

Tabel 3.1 Variabel Perhitungan

No	Variabel	Keterangan
1	NETT1	Nett1 adalah berat buah kelapa sawit
2	Sortasi_persen	Kualitas sawit

Data yang dimaksud adalah data penilaian untuk pengelompokan Hasil panen sawit berdasarkan berat dan kualitas yang diperoleh dari CV. Ram Bintang Motor. Dataset terdiri dari 4 kolom, kolom pertama adalah nomor data, kolom kedua adalah nama dari masing-masing pemilik buah kelapa sawit, kolom ketiga adalah *sortasi_persen* yang merupakan kualitas buah kelapa sawit, dan yang keempat adalah *nett1* merupakan berat buah kelapa sawit.

Tabel 3.2 Data Penelitian

No	Supir	SORT_PERSEN	NETT1
1	Sigoring-goring	4	1354
2	Binmor	4	3245
3	Andriani	3,5	7411
4	Sepeda motor 1	6	178
5	Sepeda motor 2	6	205
6	Betmen	3,5	5394
7	Amar	4	1805
8	Ritonga	4	5462
9	Harahap	4	874
10	Sigoring-goring	4	3149
1000	Kiki	4	2470

2. *Cleaning Data*

Data diatas merupakan data asli (*real*) yang akan diuji menggunakan metode *K-Medoids Clustering*, dimana data tersebut telah bersih dan tidak perlu melakukan proses *cleaning*.

3. *Transformasi Data*

Transformasi data merupakan proses *transformasi* pada data yang telah dipilih sehingga data tersebut sesuai untuk proses data mining. *Transformasi* data menggunakan *k-medoids* adalah teknik dimana data diubah atau dimodifikasi dengan memanfaatkan *k-medoids*. *Tranformasi* data dapat membantu mengidentifikasi pola atau struktur dalam data yang mungkin sulit terlihat pada awalnya (Kurniawan dkk, 2023). Data yang dipeoleh dari instansi penelitian tidak perlu *transformasi* karena tidak terdapat data yang bukan *numerik*.

Tabel 3.3 Data tanpa *Transformasi*

No	Supir	Sortasi_Persen	Nett1
1	Sigoring-Goring	4	1354
2	Binmor	4	3245
3	Andriani	3.5	7411
4	Sepeda Motor 1	6	178
5	Sepeda Motor 2	6	205
6	Betmen	3.5	5394
7	Amar	4	1805
8	Ritonga	4	5462
9	Harahap	4	874
10	Sigoring-Goring	4	3149
11	Sepeda Motor 3	6	75
12	Sepeda Motor 3	6	55
13	Sepeda Motor 1	6	55
14	Sepeda Motor 3	6	65
15	Putra D.	4	1814
16	Sepeda Motor 3	6	140
17	Sepeda Motor 1	6	800
18	Sepeda Motor 3	6	122
19	Sepeda Motor 2	4	1354
1000	Kiki	4	2470

4. *Normalisasi Data*

Normalisasi data dapat dilakukan dengan menerapkan rumus persamaan (2.4). *Normalisasi data* adalah proses mengubah nilai-nilai dalam suatu dataset ke dalam skala tertentu atau rentang yang standar (Okta dkk., 2019). Tujuan normalisasi data adalah untuk memastikan bahwa data memiliki properti tertentu atau distribusi yang tepat memudahkan analisis pemrosesan. Hasil *normalisasi data* dapat dilihat pada tabel dibawah ini :

Tabel 3.4 *Normalisasi data*

No	Supir	Sortasi_Persen	Nett1
1	Sigoring-Goring	0.4	0.083087
2	Binmor	0.4	0.202205
3	Andriani	0.3	0.46463
4	Sepeda Motor 1	0.8	0.009008
5	Sepeda Motor 2	0.8	0.010709
6	Betmen	0.3	0.337575
7	Amar	0.4	0.111496
8	Ritonga	0.4	0.341858
9	Harahap	0.4	0.05285
10	Sigoring-Goring	0.4	0.196157
11	Sepeda Motor 3	0.8	0.00252
12	Sepeda Motor 3	0.8	0.00126
13	Sepeda Motor 1	0.8	0.00126
14	Sepeda Motor 3	0.8	0.00189
15	Putra D.	0.4	0.112063
16	Sepeda Motor 3	0.8	0.006614
17	Sepeda Motor 1	0.8	0.048189
18	Sepeda Motor 3	0.8	0.00548
19	Sepeda Motor 2	0.4	0.083087
1000	kiki	0.4	0.153386

5. Inisialisasi Pusat Cluster

Setelah data dinormalisasi maka dilanjutkan dengan melakukan perhitungan manual metode *k-medoids clustering*. Inisialisasi pusat Cluster sebanyak *k* (jumlah cluster), untuk menentukan pusat centroid awal ditentukan dengan mengacak dari data yang sudah ada, data acak yang digunakan pada penelitian ini adalah data 998, 999 dan 1000. Dapat dilihat pada table 3.5

Tabel 3.5 Medoid awal

NO	SORT_PERSEN	NETT 1
998	0.4	0.12063
999	0.4	0.15339
1000	0.3	0.37134

6. Jarak Auclidian Distance

Perhitungan jarak pada data yang telah di normalisasi dengan *cluster* menggunakan rumus pada persamaan 2.1 perhitungan jarak dari data 1 terhadap pusat *cluster* adalah:

$$d(1,1) = \sqrt{(0,000 - 0,000)^2 + (0,012 - 0,007)^2 + (0,007 - 0,166)^2} = 0,159$$

$$d(1,2) = \sqrt{(0,000 - 0,000)^2 + (0,007 - 0,007)^2 + (0,223 - 0,166)^2} = 0,057$$

Seterusnya dilakukan mencari jarak untuk data ke – 2 sampai data ke – 1000.

Kemudian akan didapatkan hasil perhitungan jarak setiap data terhadap pusat *cluster* sementara sebagai berikut :

Tabel 3.6 Jarak kedekatan objek ke *medoid* sementara

No	cos1	cos2	cos3	kedekatan	cluster
1	0.037543	0.070299	0.305105	0.037543	1
2	0.081575	0.048819	0.196485	0.048819	2
3	0.35824	0.326914	0.093291	0.093291	3
4	0.415282	0.425259	0.617482	0.415282	1

5	0.414828	0.424684	0.616485	0.414828	1
6	0.238883	0.209584	0.033764	0.033764	3
7	0.009134	0.04189	0.278421	0.009134	1
8	0.221228	0.188472	0.104255	0.104255	3
9	0.06778	0.100535	0.333818	0.06778	1
10	0.075528	0.042772	0.201714	0.042772	2
11	0.417073	0.427505	0.621311	0.417073	1
12	0.417432	0.427951	0.62206	0.417432	1
13	0.417432	0.427951	0.62206	0.417432	1
14	0.417252	0.427728	0.621685	0.417252	1
15	0.008567	0.041323	0.277892	0.008567	1
16	0.415932	0.426077	0.618889	0.415932	1
17	0.406507	0.413602	0.595337	0.406507	1
18	0.416244	0.426469	0.619558	0.416244	1
1000	0.032756	0	0.239799	0	1
		jumlah kedekatan		240.456	

Total kedekatan 1 adalah 240.456. Karena perhitungan masih iterasi 1 maka perhitungan dilanjutkan ke iterasi 2.

7. Perhitungan Pusat *Cluster Baru*

Perhitungan pusat *cluster* baru, dilakukan dengan cara yang sama dengan perhitungan jarak kedekatan objek ke *medoid* sementara. menentukan *medoid* baru, menghitung jarak kedekatan objek ke *medoid* baru, dilanjutkan dengan menghitung total simpangan. Berikut adalah table *medoid* baru.

Tabel 3.7 *Medoid* baru

NO	SORT_PERSEN	NETT 1
1	0.4	0.083087
2	0.4	0.202205
3	0.3	0.46463

Langkah selanjutnya dengan menghitung jarak kedekatan objek ke *medoid* baru, hasil perhitungan jarak setiap data terhadap pusat *cluster* Baru sebagai berikut :

Tabel 3.8 Jarak kedekatan objek ke *medoid* baru

No	cos1	cos2	cos3	kedekatan	cluster
1	0	0.119118	0.39443	0.119118	1
2	0.119118	0	0.280833	0	2
3	0.39443	0.280833	0	0	2
4	0.406802	0.444213	0.676455	0.444213	1
5	0.406495	0.443476	0.675311	0.443476	1
6	0.27343	0.1683	0.127055	0.127055	2
7	0.028409	0.090709	0.36702	0.090709	1
8	0.258772	0.139654	0.158344	0.139654	2
9	0.030236	0.149354	0.423748	0.149354	1
10	0.113071	0.006047	0.286492	0.006047	2
11	0.408033	0.447073	0.680842	0.447073	1
12	0.408284	0.447637	0.681698	0.447637	1
13	0.408284	0.447637	0.681698	0.447637	1
14	0.408158	0.447355	0.68127	0.447355	1
1000	0.070299	0.048819	0.326914	0.048819	2
	jumlah kedekatan			277.0535	

Selanjutnya dengan menghitung total simpangan (S) dengan menghitung nilai total *distance* baru -total *distance* lama. Jika $S < 0$ maka tukar objek dengan data *Cluster* untuk membentuk sekumpulan objek baru sebagai *medoids*.

$S = \text{total } distance \text{ baru} - \text{total } distance \text{ lama}$

$$S = 277.0535 - 240.456 = 36.59745$$

Sesuai dengan data yang diuji maka iterasi berhenti pada iterasi ke 2 dengan total simpangan 277.0525 sama dengan total simpangan(S) pada iterasi pertama 240.456 sehingga $(S) \geq 0$ maka iterasi selesai. Hasil pengelompokan yang diperoleh dapat dilihat pada tabel dibawah ini.

Tabel 3.9 Kesimpulan pengelompokan sawit

No	Nama	Kelompok	Kesimpulan
1	Sigoring-Goring	1	Rendah
2	Binmor	2	Sedang
3	Andriani	3	Tinggi
4	Sepeda Motor 1	1	Rendah
5	Sepeda Motor 2	1	Rendah
6	Betmen	3	Tinggi
7	Amar	1	Rendah
8	Ritonga	3	Tinggi
9	Harahap	1	Rendah
10	Sigoring-Goring	2	Sedang
11	Sepeda Motor 3	1	Rendah
12	Sepeda Motor 3	1	Rendah
13	Sepeda Motor 1	1	Rendah
14	Sepeda Motor 3	1	Rendah
15	Putra D.	1	Rendah
16	Sepeda Motor 3	1	Rendah
1000	Kiki	1	Rendah

3.3.6 Penulisan Laporan

Pada tahap ini, peneliti telah melakukan seluruh tahapan penelitian dan kini fokus pada penulisan laporan akhir. Laporan ini merupakan hasil dokumentasi yang menggambarkan semua kegiatan yang telah dilakukan sejak studi literatur hingga tahap akhir penelitian. Dalam laporan ini, peneliti akan memaparkan secara detail setiap langkah yang telah dilakukan, termasuk tujuan penelitian, metodologi yang digunakan, data yang dikumpulkan, analisis yang dilakukan, temuan-temuan, serta kesimpulan yang diperoleh. Seluruh hasil penelitian akan diuraikan dengan jelas dan disajikan secara sistematis dalam bentuk paragraf, tabel, dan grafik yang relevan.

Tahap akhir penulisan laporan ini juga mencakup penulisan kesimpulan dan saran. Kesimpulan akan berisi ringkasan temuan-temuan penting yang diperoleh dari penelitian, termasuk hasil *clustering*. Sementara itu, saran-saran yang diberikan dapat berkaitan dengan rekomendasi untuk penelitian lanjutan atau langkah-langkah yang dapat diambil untuk mendapat hasil kluster dengan masing-masing nama data dan rekomendasi pabrik pengiriman buah kelapa sawit.

BAB V KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil dan pembahasan yang dilakukan, maka penelitian ini dapat disimpulkan :

Analisis hasil panen sawit menggunakan algoritma *K-medoids* telah berhasil memperoleh nilai *cluster* rendah sebanyak 43,5%, *cluster* sedang sebanyak 44,4%, dan *cluster* tinggi sebanyak 12,1%. Sortasi_persen dan nett 1 dijadikan sebagai variabel perhitungan dan mendapatkan hasil total simpangan (S) dengan hasil 67.108. penerapan *tools google colab* berupa plot kelompok sawit 1 dengan lebel berwarna biru, kelompok sawit 2 dengan lebel berwarna orange, dan kelompok sawit 3 dengan lebel berwarna hijau. Validasi hasil pengujian menggunakan *silhouette* dengan nilai skor sebanyak 0.7198.

5.2 Saran

Saran untuk pengembangan pada penelitian ini dimasa yang akan datang sebagai berikut :

1. Penelitian selanjutnya dapat mengelompokkan hasil panen sawit dengan hasil output nama pada masing-masing cluster.
2. Penelitian selanjutnya dapat menggunakan perbandingan algoritma yang bertujuan untuk mendapat hasil yang terbaik.
3. Menambahkan data kriteria penerimaan dari pabrik kelapa sawit.

DAFTAR PUSTAKA

- Andini, A. D., & Arifin, T. (2020). Implementasi Algoritma K-Medoids Untuk Klusterisasi Data Penyakit Pasien Di Rsud Kota Bandung. *Jurnal Responsif: Riset Sains Dan Informatika*, 2(2), 128–138.
- Boy, A. F. (2020). Implementasi Data Mining Dalam Memprediksi Harga Crude Palm Oil (Cpo) Pasar Domestik Menggunakan Algoritma Regresi Linier Berganda (Studi Kasus Dinas Perkebunan Provinsi Sumatera Utara). *Journal Of Science And Social Research*, 3(2), 78–85.
- Chaira, N., Napitulu, D., & Ulma, R. O. (2022). Analisis Efisiensi Teknis Penggunaan Faktor Produksi Pada Usahatani Kelapa Sawit Rakyat Di Kecamatan Kumpeh Ulu Kabupaten Muaro Jambi. *Journal Of Agribusiness And Local Wisdom*, 5(1), 15–29.
- Fernando, E., Surjandy, M., & Siagian, P. (2020). Desain Sistem Pengenalan Varietas Bibit Tanaman Kelapa Sawit Dengan Pendekatan Design Science Research Methodology (Dsr) Design System Of Palm Oil Plant Variety Recognition Using Design Science Research Methodology (Dsr) Approach. *Jurnal Teknologi Informasi Dan Ilmu Komputer (Jtiik)*, 7(2), 249–258.
- Harahap, F. (2021). Perbandingan Algoritma K Means Dan K Medoids Untuk Clustering Kelas Siswa Tunagrahita. *Tin: Terapan Informatika Nusantara*, 2(4), 191–197.
- Hasibuan, B. R., Rahayu, E., & Astuti, Y. T. M. (2019). Kajian Pengaruh Topografi Terhadap Produksi Kelapa Sawit Di Pt. Gunung Sejahtera Yoli Makmur (Gsym) Kecamatan Arut Utara, Kabupaten Kotawaringin Barat,

Kalimantan Tengah. *Jurnal Agromast*, 3(1).

Kurniawan, R., Martadinata, A. T., & Cahyo, S. D. (2023). Klasifikasi Tingkat Kematangan Buah Sawit Berbasis Deep Learning dengan Menggunakan Arsitektur Yolov5. *Journal of Information System Research (JOSH)*, 5(1), 302–309.
<https://doi.org/10.47065/josh.v5i1.4408>

Lesmana, B. (2021). Pengelompokan Pengiriman Hasil Kelapa Sawit Berdasarkan Tonase Dan Kualitas Menggunakan Metode Clustering (Studi Kasus: Kud Bumi Pusaka). *Seminar Nasional Informatika (Senatika)*, 154–165.

Nabila, Z., Isnain, A. R., Permata, P., & Abidin, Z. (2021). Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means. *Jurnal Teknologi Dan Sistem Informasi*, 2(2), 100–108.

Nurmalita, V., & Bowo, P. A. (2019). Analisis Faktor-Faktor Yang Mempengaruhi Ekspor Minyak Kelapa Sawit Indonesia Ke India. *Economic Education Analysis Journal*, 8(2), 605–619.

Oktarian, S., Sarjon Defit, S., & Padang, P. I. Y. (2020). Klasterisasi Penentuan Minat Siswa Dalam Pemilihan Sekolah Menggunakan Metode Algoritma K-Means Clustering. *Jurnal Informasi Dan Teknologi Vol*, 2(3), 68–75.

Okta et al. (2019)Agustini, L., Kirana, I. O., & Others. (2022). Pengelompokan Data Janjang Panen Kelapa Sawit Menggunakan Algoritma K-Medoids Pada Pt Sir Mandau. *Journal Of Machine Learning And Data Analytics*, 1(1), 36–44.

Saputra, N., Abdinagoro, S. B., Kuncoro, E. A., & Others. (2020). *Sustainable Growth Formula: Keterlekatan Kerja Dan Ketangkasan Belajar Dari Managerial Resources Industri Minyak Sawit Indonesia*. Scopindo Media

Pustaka.

- Sari, W. E., Muslimin, M., Franz, A., & Sugiartawan, P. (2022). Deteksi Tingkat Kematangan Tandan Buah Segar Kelapa Sawit Dengan Algoritme K-Means. *Sintech (Science And Information Technology) Journal*, 5(2), 154–164.
- Septiani, I. W., Fauzan, A. C., & Huda, M. M. (2022). Implementasi Algoritma K-Medoids Dengan Evaluasi Davies-Bouldin-Index Untuk Klasterisasi Harapan Hidup Pasca Operasi Pada Pasien Penderita Kanker Paru-Paru. *Jurnal Sistem Komputer Dan Informatika (Json)*, 3(4), 556–566.
- Setio, P. B. N., Saputro, D. R. S., & Winarno, B. (2020). Klasifikasi Dengan Pohon Keputusan Berbasis Algoritme C4. 5. *Prisma, Prosiding Seminar Nasional Matematika*, 3, 64–71.
- Sugiono, E., Efendi, S., & Al-Afgani, J. (2021). Pengaruh Motivasi Kerja, Kompetensi, Dan Kompensasi Terhadap Kinerja Karyawan Melalui Kepuasan Kerja Pt. Wibee Indoedu Nusantara (Pustaka Lebah) Di Jakarta. *Jurnal Ilmiah Mea (Manajemen, Ekonomi, & Akuntansi)*, 5(1), 718–734.
- Thoharudin, M., Huda, F. A., & Suryadi, T. (2020). Penanganan Hasil Panentandan Buah Segar Kelapa Sawit Oleh Koperasi Harapan Jaya. *Papatung: Jurnal Ilmu Administrasi Publik, Pemerintahan Dan Politik*, 3(1), 150–158.
- Waskitha, R. P., Santoso, T. N. B., & Astuti, Y. T. M. (2018). Pengaruh Usia Pemanen Terhadap Hasil Panen Dengan Topografi Berbeda. *Jurnal Agromast*, 3(2).
- Yansa, A., & Others. (2023). Perbandingan Implementasi Ct Pro Dan

AlgoritmaC45 Menentukan Pola Penyakit Pasien Stroke (Studi Kasus Rs. Siti Fatimah Palembang). *Jupiter (Jurnal Penelitian Ilmu Dan Teknik Komputer)*, 15(1c), 549–558.



Lampiran SK Dosen Pembimbing



UNIVERSITAS MEDAN AREA FAKULTAS TEKNIK

Kampus I : Jalan Kolam Nomor 1 Medan Estata/Jalan PBSI Nomor 1 ☎ (061) 7366878, 7360168, 7364348, 7366781, Fax. (061) 7366908 Medan 20223
Kampus II : Jalan Seliabudi Nomor 79 / Jalan Sei Semayu Nomor 70 A, ☎ (061) 8225602, Fax. (061) 8226331 Medan 20122
Website: www.teknik.uma.ac.id E-mail: univ_medanarea@uma.ac.id

Nomor : 649/FT.6/01.10/VIII/2023
Lamp : -
Hal : **Perubahan Judul Tugas Akhir**

22 Agustus 2023

Yth, Pembimbing Tugas Akhir
Dr. Dian Noviandri, S.T, M.Kom
di
Tempat

Dengan hormat, Sehubungan dengan adanya perubahan judul tugas akhir maka perlu diterbitkan kembali SK Pembimbing Skripsi baru atas nama mahasiswa tersebut :

Nama : Yusril Izza Haholongan Siregar
N P M : 198160028
Jurusan : Teknik Informatika

Maka dengan hormat kami mengharapkan kesediaan saudara :

Dr. Dian Noviandri, S.T, M.Kom (Sebagai Pembimbing)

Adapun Tugas Akhir Skripsi berjudul :

“Analisis Algoritma K-Medoids untuk Pengelompokkan Hasil Panen Sawit Berdasarkan Berat dan Kualitas”.

SK Pembimbing ini berlaku selama enam bulan terhitung sejak SK ini diterbitkan. Jika proses pembimbing melebihi batas waktu yang telah ditetapkan, SK ini dapat ditinjau ulang.

Demikian kami sampaikan, atas kesediaan saudara diucapkan terima kasih.

Dekan,


Dr. Rahmiad Syah, S. Kom, M. Kom

Lampiran Surat Pengantar Riset



UNIVERSITAS MEDAN AREA FAKULTAS TEKNIK

Kampus I : Jalan Kolam Nomor 1 Medan Estate/Jalan PBSI Nomor 1 (061) 7366878, 7360160, 7364340, 7366781, Fax. (061) 7366998 Medan 20223
Kampus II : Jalan Sotabudi Nomor 79 / Jalan Sei Semayu Nomor 70 A, (061) 8225602, Fax. (061) 8226331 Medan 20122
Website: www.teknik.uma.ac.id E-mail: univ_medan@uma.ac.id

Nomor : 572 /FT.6/01.10/VII/2023
Lamp : -
Hal : Penelitian Dan Pengambilan Data Tugas Akhir

27 Juli 2023

Yth. Pimpinan CV. Ram Bintang Motor
Desa Binanga
Di
Padang Lawas

Dengan hormat,

Kami mohon kesediaan Bapak/Ibu berkenan untuk memberikan izin dan kesempatan kepada mahasiswa kami tersebut dibawah ini :

NO	NAMA	NPM	PRODI
1	Yusril Izza Haholongan Siregar	198160028	Teknik Informatika

Untuk melaksanakan Penelitian dan Pengambilan Data Tugas Akhir pada perusahaan/Instansi yang Bapak/Ibu Pimpin.

Perlu kami jelaskan bahwa Pengambilan Data tersebut adalah semata-mata untuk tujuan ilmiah dan Skripsi yang merupakan salah satu syarat bagi mahasiswa tersebut untuk mengikuti ujian sarjana pada Fakultas Teknik Universitas Medan Area dan tidak untuk dipublikasikan, dengan judul penelitian :

Analisis Metode Clustering menggunakan Algoritma K-Medoids Pengelompokan Pengiriman Hasil Sawit Masyarakat Berdasarkan Berat dan Kualitas

Atas perhatian dan kerja sama yang baik diucapkan terima kasih.



Dr. Rahmat Syah S. Kom, M. Kom

Tembusan :
1. Ka. BAMAI
2. Mahasiswa
3. File

Lampiran Surat Selesai Riset



CV. RAM BINTANG MOTOR
Alamat : Jalan lintas Gunung Tua Sibuhuan/Desa Binanga ☎ 082168234857
Jual beli kelapa sawit (TBS) Kec Barumon Tengah Kab Padang Lawas

SURAT KETERANGAN SELESAI PENELITIAN

Yang bertanda tangan dibawah ini :

Nama : Meri Analisa Harahap
Jabatan : Pimpinan Ram Bintang Motor
Alamat : Desa Binanga Kabupaten Padang lawas

Dengan ini menerangkan bahwa mahasiswa yang beridentitas

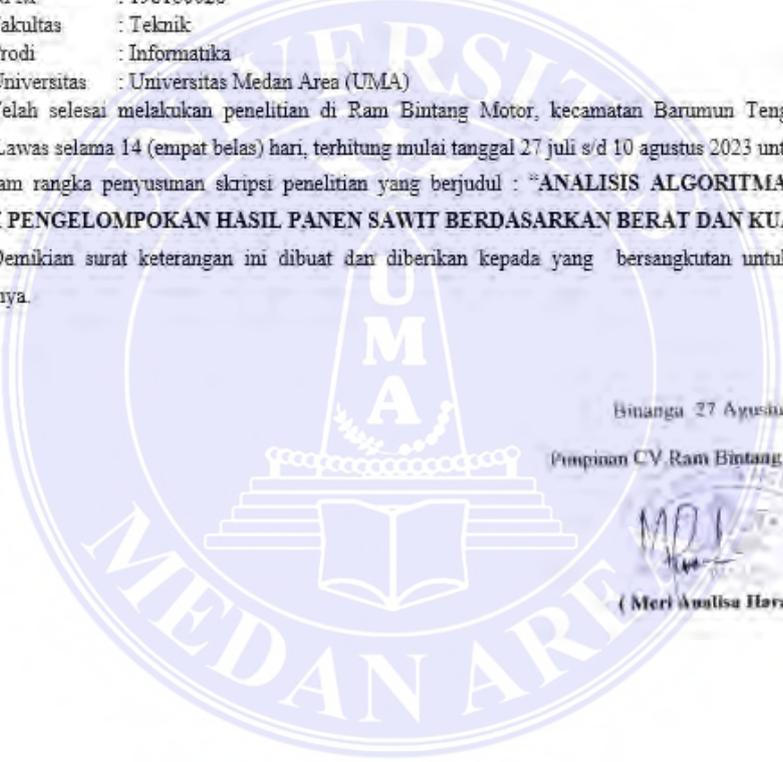
Nama : Yusril Izza Haholongan Siregar
NPM : 198160028
Fakultas : Teknik
Prodi : Informatika
Universitas : Universitas Medan Area (UMA)

Telah selesai melakukan penelitian di Ram Bintang Motor, kecamatan Barumon Tengah, Kabupaten Padang Lawas selama 14 (empat belas) hari, terhitung mulai tanggal 27 juli s/d 10 agustus 2023 untuk memperoleh data dalam rangka penyusunan skripsi penelitian yang berjudul : "ANALISIS ALGORITMA K-MEDOIDS UNTUK PENGELOMPOKAN HASIL PANEN SAWIT BERDASARKAN BERAT DAN KUALITAS".

Demikian surat keterangan ini dibuat dan diberikan kepada yang bersangkutan untuk dipergunakan sepenuhnya.

Binanga 27 Agustus 2023
Pimpinan CV Ram Bintang Motor

(Meri Analisa Harahap)



Lampiran Turnitin

 **Similarity Report ID:** oid:29477:44319302

PAPER NAME: **Revisi Similarity Yusril 1.docx** AUTHOR: **Yusril izza Haholongan Siregar**

WORD COUNT: **6772 Words** CHARACTER COUNT: **33882 Characters**

PAGE COUNT: **42 Pages** FILE SIZE: **525.0KB**

SUBMISSION DATE: **Oct 4, 2023 9:47 PM GMT+7** REPORT DATE: **Oct 4, 2023 9:48 PM GMT+7**

● **21% Overall Similarity**

The combined total of all matches, including overlapping sources, for each database.

- 20% Internet database
- 9% Publications database
- Crossref database
- Crossref Posted Content database
- 14% Submitted Works database

● **Excluded from Similarity Report**

- Small Matches (Less then 10 words)

Lampiran Source Code

```

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

from sklearn.metrics import silhouette_samples,
silhouette_score

def kmedoids(X, low_threshold, high_threshold, initial_medoids,
max_iterations=10):
    n_clusters = len(initial_medoids)
    medoids = np.array(initial_medoids)
    total_deviation = 0

    for iteration in range(max_iterations):
        distances = np.linalg.norm(X[:, np.newaxis, :] -
medoids, axis=2)
        labels = np.argmin(distances, axis=1)

        new_medoids = np.array([X[labels == i].mean(axis=0) for
i in range(n_clusters)])

        # Hitung total simpangan di setiap kluster

        total_deviation =
np.sum([np.sum(np.linalg.norm(X[labels == i] - new_medoids[i],
axis=1)) for i in range(n_clusters)])

        if np.all(medoids == new_medoids):

```

```

        print(f"Iterasi berhenti pada iterasi ke-{'iteration
+ 1}')

        break # Berhenti jika medoids tidak berubah

        medoids = new_medoids.copy() # Memperbarui medoids
dengan nilai baru

        low_cluster = X[labels == 0]

        medium_cluster = X[(labels == 1) & (low_threshold <= X[:,
1]) & (X[:, 1] <= high_threshold)]

        high_cluster = X[labels == 2]

        # Hitung total simpangan untuk seluruh klaster
        s = np.sum([np.sum(np.linalg.norm(X[labels == i] -
medoids[i], axis=1)) for i in range(n_clusters)])

        # Menampilkan medoids baru
        print("Medoids Baru:")
        print(medoids)

        # Perhitungan Silhouette

        silhouette_avg = silhouette_score(X, labels)
        print(f"Silhouette Score: {silhouette_avg}")

        # Grafik Silhouette

        plot_silhouette(X, labels)

```

```

    return low_cluster, medium_cluster, high_cluster, labels,
medoids, total_deviation, s

def plot_silhouette(X, labels):
    silhouette_vals = silhouette_samples(X, labels)
    y_lower = 10

    for i in range(len(set(labels))):
        cluster_silhouette_vals = silhouette_vals[labels == i]
        cluster_silhouette_vals.sort()

        y_upper = y_lower + len(cluster_silhouette_vals)

        color = plt.cm.nipy_spectral(float(i) /
len(set(labels)))
        plt.fill_betweenx(np.arange(y_lower, y_upper),
                        0, cluster_silhouette_vals,
                        facecolor=color, edgecolor=color,
alpha=0.7)

        # Label silhouette plot with cluster numbers at the
middle
        plt.text(-0.05, y_lower + 0.5 *
len(cluster_silhouette_vals), str(i))

        # Compute the new y_lower for next plot
        y_lower = y_upper + 10

```

```
plt.xlabel("Silhouette Coefficient Values")

plt.ylabel("Cluster Label")

plt.show()

# Baca data dari file Excel
data = pd.read_excel('//content/drive/MyDrive/data.ram.xlsx')

# Ambil dua kolom sebagai variabel X
X = data[['SORT_PERSEN', 'NETT 1']].values

# Inisialisasi medoid awal
initial_medoids = [
    [0.4, 0.083087],
    [0.4, 0.202205],
    [0.3, 0.464630]
]

# Threshold untuk klaster sedang
low_threshold = 0.7
high_threshold = 0.3

# Lakukan k-medoids clustering dengan medoid yang diperbarui
setiap iterasi
low_cluster, medium_cluster, high_cluster, labels, medoids,
total_deviation, s = kmedoids(X, low_threshold, high_threshold,
initial_medoids, max_iterations=10)
```

```

# Menampilkan nomor data dan label klaster
for i, label in enumerate(labels):
    print(f>Data ke-{i+1}: Klaster {label + 1} ({'rendah' if
label == 0 else ('sedang' if label == 1 else 'tinggi')})")

# Menampilkan klaster
print("Cluster rendah:")
print(low_cluster)
print("Cluster sedang:")
print(medium_cluster)
print("Cluster tinggi:")
print(high_cluster)

# Print total deviation for the entire clustering
print(f>Total Deviation (s) untuk seluruh klaster: {s}")

# Scatter plot 3D untuk hasil klaster
fig = plt.figure()
ax = fig.add_subplot(111, projection='3d')

# Plot data untuk setiap klaster
for i in range(len(medoids)):
    cluster_points = X[labels == i]
    ax.scatter(cluster_points[:, 0], cluster_points[:, 1],
np.ones_like(cluster_points[:, 1]) * i, label=f'Cluster {i +
1}')

```

```
# Plot medoids
ax.scatter(medoids[:, 0], medoids[:, 1],
np.arange(len(medoids)), marker='X', c='red', s=200,
label='Medoids')

# Label sumbu dan judul plot
ax.set_xlabel('SORT_PERSEN')
ax.set_ylabel('NETT 1')
ax.set_zlabel('Cluster')
plt.title('Hasil Klaster sawit')

# Tampilkan legenda
plt.legend()

# Tampilkan plot
plt.show()
```