
APLIKASI WEB PENGELOMPOKAN DATA PENERIMAAN TANDAN BUAH SEGAR (TBS) KELAPA SAWIT MENGUNAKAN ALGORITMA K-MEANS CLUSTERING

Stefen Manalu¹, Parasian D. P. Silitonga², Andy Paul Harianja, ST, M.Kom³, Wasit Ginting, S.Kom, M.Kom⁴

Universitas Katolik Santo Thomas

e-mail: ¹manalus018@gmail.com, ²parasianirene@gmail.com, ³apharianja@gmail.com,
⁴wasitginting74@gmail.com

Abstract: *This study aims to design and develop a web-based information system that applies the K-Means Clustering algorithm to group data from palm oil Fresh Fruit Bunch (FFB) suppliers. The problem faced by CV Saroha Maju Jaya is the management of massive historical FFB receipt data that is not yet analytically structured, complicating the supplier performance evaluation process. This research proposes a novelty through the application of Data Preprocessing techniques, specifically Logarithmic Transformation (Log10) and Min-Max Normalization, before the data is fed into the K-Means model. This approach is implemented to prevent computational bias caused by the dominance of large-scale variables (such as rupiah values) over small-scale variables (such as kilogram deductions). The system was developed using the Waterfall method and the Model-View-Controller (MVC) architecture with the PHP programming language and MySQL database. Computational results show that the algorithm successfully and accurately grouped suppliers into two main clusters, namely Small Volume (Cluster 1) and Large Volume (Cluster 2), reaching a point of convergence in the second iteration. The User Acceptance Test (UAT) yielded a satisfaction rate of 92.4%, which confirms that the system is highly feasible and effective in facilitating the company's strategic decision-making.*

Keyword: *Data Mining, K-Means Clustering, Fresh Fruit Bunches, Web Application, Preprocessing.*

Abstrak: Penelitian ini bertujuan untuk merancang dan membangun sistem informasi berbasis web yang menerapkan algoritma K-Means Clustering untuk mengelompokkan data pemasok (supplier) Tandan Buah Segar (TBS) kelapa sawit [cite: 1]. Permasalahan yang dihadapi oleh CV Saroha Maju Jaya adalah pengelolaan data historis penerimaan TBS yang masif dan belum terstruktur secara analitik, sehingga menyulitkan proses evaluasi performa pemasok [cite: 1]. Penelitian ini mengusulkan kebaruan (novelty) berupa penerapan teknik Data Preprocessing, yaitu Transformasi Logaritma (Log10) dan Normalisasi Min-Max, sebelum data dimasukkan ke dalam model K-Means [cite: 1]. Pendekatan ini dilakukan untuk mencegah bias komputasi yang diakibatkan oleh dominasi variabel berskala besar (seperti nilai rupiah) terhadap variabel berskala kecil (seperti potongan kilogram) [cite: 1]. Sistem dikembangkan menggunakan metode Waterfall dan arsitektur Model-View-Controller (MVC) dengan bahasa pemrograman PHP dan basis data MySQL [cite: 1]. Hasil komputasi menunjukkan algoritma berhasil mengelompokkan supplier ke dalam dua klaster utama, yaitu Volume Kecil (Klaster 1) dan Volume Besar (Klaster 2), secara akurat hingga mencapai titik konvergen pada iterasi kedua [cite: 1]. Pengujian penerimaan pengguna (User Acceptance Test/UAT) menghasilkan tingkat kepuasan sebesar 92,4%, yang menegaskan bahwa sistem sangat layak dan efektif dalam memfasilitasi pengambilan keputusan strategis perusahaan [cite: 1].

Kata kunci: Data Mining, K-Means Clustering, Tandan Buah Segar, Aplikasi Web, Preprocessing.

PENDAHULUAN

Industri kelapa sawit merupakan salah satu sektor agribisnis strategis yang menopang perekonomian nasional di Indonesia [cite: 1]. Tingginya tingkat produksi komoditas ini menuntut pengelolaan rantai pasok yang efektif, khususnya dalam proses penerimaan Tandan Buah Segar (TBS) dari petani atau supplier ke pihak pengepul (RAM) atau pabrik kelapa sawit [cite: 1]. Pada operasional kesehariannya, entitas usaha seperti CV Saroha Maju Jaya menerima puluhan hingga ratusan ton TBS yang diiringi dengan pencatatan data transaksi yang sangat besar, mencakup berat bersih (netto), potongan (kualitas sortiran), harga beli harian, dan total harga [cite: 1].

Permasalahan utama yang sering dihadapi adalah data penerimaan tersebut hanya diarsipkan sebagai catatan administratif (spreadsheet manual) tanpa dianalisis lebih lanjut untuk mengekstraksi informasi strategis [cite: 1]. Ketiadaan sistem analitik membuat perusahaan kesulitan dalam memetakan dan mengevaluasi kinerja masing-masing supplier secara objektif, yang pada akhirnya dapat memengaruhi ketepatan penentuan strategi operasional dan pembinaan mitra [cite: 1]. Oleh karena itu, diperlukan penerapan teknologi Data Mining untuk menemukan pola tersembunyi dari sekumpulan data masif tersebut [cite: 1].

Metode clustering, khususnya algoritma K-Means, telah dikenal luas kemampuannya dalam

mengelompokkan data berdasarkan kemiripan karakteristik [cite: 1]. Namun, terdapat tantangan teknis ketika algoritma ini diterapkan pada data dengan variansi skala yang sangat ekstrem, misalnya menggabungkan variabel "Total Harga" (jutaan rupiah) dengan "Potongan" (puluhan kilogram). Jika dihitung

langsung menggunakan Euclidean Distance, variabel berskala besar akan mendominasi hasil perhitungan (bias) [cite:1]. Oleh karena itu, penelitian ini memperkenalkan kebaruan dengan mengintegrasikan tahapan Data Preprocessing berupa Transformasi Log10 dan Normalisasi Min-Max sebelum algoritma K-Means dieksekusi [cite: 1]. Sistem ini diimplementasikan dalam bentuk aplikasi berbasis web berarsitektur MVC untuk memfasilitasi integrasi data yang tersentralisasi dan pelaporan yang interaktif [cite: 1].

METODE

Metode Pengumpulan Data

Pengumpulan data empiris dalam penelitian ini dilakukan melalui tiga pendekatan utama [cite: 1]. Pertama, observasi sistematis terhadap alur penimbangan dan penyortiran TBS di CV Saroha Maju Jaya. Kedua, wawancara mendalam dengan pihak manajemen dan operator lapangan terkait kendala pencatatan konvensional. Ketiga, studi literatur dari berbagai publikasi akademik terkait Data Mining dan metode Clustering [cite: 1].

Preprocessing Data

Sebelum masuk ke dalam perhitungan inti K-Means, data kuantitatif diproses terlebih dahulu agar setara dan terhindar dari bias [cite: 1].

1. Transformasi Logaritma Basis 10 (Log10): Digunakan untuk memperkecil rentang skala nilai yang ekstrem, seperti total harga transaksi [cite: 1].

$$X' = \text{Log}_{10}(X)$$

2. Normalisasi Min-Max: Mengonversi seluruh nilai data yang telah ditransformasikan ke dalam rentang rasio [0, 1] [cite: 1].

$$X_{norm} = (X - X_{min}) / (X_{max} - X_{min})$$

Algoritma K-Means Clustering

K-Means adalah algoritma unsupervised learning yang mempartisi data ke dalam K kluster. Dalam konteks ini, ditetapkan nilai $K = 2$ (Kluster 1 untuk Volume Kecil dan Kluster 2 untuk Volume Besar) [cite: 1]. Proses iteratif K-Means menggunakan persamaan Euclidean Distance untuk mengukur kedekatan setiap titik data ke pusat kluster (centroid) [cite: 1]:

$$D(X, Y) = \sqrt{\sum (X_i - Y_i)^2}$$

Pengembangan Sistem

Perangkat lunak dibangun menggunakan metodologi Waterfall, yang mencakup tahapan analisis, perancangan (berbasis Unified Modeling Language), pengkodean, dan pengujian [cite: 1]. Aplikasi dikembangkan menggunakan arsitektur Model-View-Controller (MVC) berbasis PHP, framework Tailwind CSS untuk antarmuka, dan sistem manajemen basis data MySQL [cite: 1].

HASIL DAN PEMBAHASAN

Analisis dan Simulasi K-Means

Sampel data yang digunakan dalam simulasi perhitungan manual melibatkan 5 transaksi supplier (P1 hingga P5) yang memuat empat atribut komputasional: Netto-2 (kg), Potongan (kg), Harga/kg (Rp), dan Total Harga (Rp) [cite: 1]. Data tersebut kemudian diproses menggunakan Transformasi Log10 dan dinormalisasi menggunakan Min-Max [cite: 1].

Pusat awal (centroid) ditentukan dengan pendekatan stabilitas data, dimana P1 ditetapkan sebagai pusat C1 dan P3 sebagai pusat C2 [cite: 1]. Setelah dilakukan perhitungan jarak Euclidean pada Iterasi 1, didapatkan pemetaan anggota kluster sementara. Centroid kemudian diperbarui berdasarkan nilai rata-rata dari keanggotaan kluster tersebut [cite: 1].

Pada Iterasi ke-2, perhitungan jarak dilakukan kembali menggunakan centroid yang baru. Hasilnya menunjukkan bahwa keanggotaan kluster tidak lagi mengalami perubahan (stabil), yang menandakan bahwa algoritma telah mencapai titik konvergen [cite: 1] naskah akan diberitahukan secara tertulis melalui email.

Tabel 1 Contoh Tabel

Kluster	Kategori	Anggota (Kode Supplier)
C1	Volume Kecil	P1, P2, P4
C2	Volume Besar	P3, P5

Penerapan tahapan preprocessing terbukti sangat signifikan. Pemasok dengan volume di bawah 1.500 Kg dapat secara presisi diklasifikasikan ke dalam Kluster 1, sedangkan yang berada di atas 1.500 Kg diklasifikasikan ke Kluster 2 tanpa adanya distorsi perhitungan akibat nominal total harga [cite: 1].

Implementasi Antarmuka Sistem

Logika matematis yang telah dianalisis kemudian ditransformasikan ke dalam basis kode PHP. Halaman Dasbor Analitik (Dashboard) menyajikan ringkasan eksekutif, meliputi total metrik, grafik Scatter Plot (untuk memvisualisasikan sebaran titik jarak Euclidean), dan proporsi kluster [cite: 1]. Aplikasi ini juga dibekali kemampuan "Import Excel" yang memungkinkan administrator untuk mengunggah ratusan data transaksi sekaligus, sehingga meminimalisir kesalahan human error pada pencatatan manual [cite: 1].

Pengujian Sistem

Pengujian fungsionalitas menggunakan metode Black Box menunjukkan bahwa seluruh modul (otentikasi, upload data, kalkulasi K-Means, dan pencetakan PDF) beroperasi secara valid [cite: 1]. Selain itu, dilakukan User Acceptance Test (UAT) menggunakan kuesioner skala Likert kepada perwakilan manajemen CV Saroha Maju Jaya [cite: 1]. Hasilnya

memperoleh skor rata-rata kepuasan sebesar 92,4% [cite: 1]. Angka ini membuktikan bahwa aplikasi dinilai "Sangat Layak" dan diakui mempercepat proses analitik kinerja pemasok secara transparan [cite: 1].

SIMPULAN

Pengintegrasian algoritma K-Means Clustering dengan teknik Data Preprocessing (Log10 dan Min-Max Normalization) terbukti menjadi solusi matematis yang tangguh untuk mengatasi bias dominasi nilai pada dataset penerimaan kelapa sawit [cite: 1]. Sistem berbasis web yang dikembangkan berhasil memetakan profil supplier ke dalam kluster Volume Kecil dan Volume Besar secara otomatis [cite: 1]. Implementasi ini berfungsi sebagai Sistem Pendukung Keputusan (SPK) yang memberdayakan CV Saroha Maju Jaya untuk menetapkan kebijakan pembinaan khusus bagi supplier bervolume rendah serta mempertahankan loyalitas pemasok bervolume tinggi [cite: 1].

Dengan tingkat penerimaan pengguna (UAT) mencapai 92,4%, aplikasi ini dinyatakan sangat sukses dalam meningkatkan efisiensi operasional manajemen data [cite: 1].

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada program Studi Teknik Informatika Universitas Katolik Santo Thomas Medan atas dukungan yang diberikan, serta kepada pihak CV Saroha Maju Jaya atas izin dan bantuan selama proses

pengumpulan data.

DAFTAR PUSTAKA

- S. Manalu, "Aplikasi Web Pengelompokan Data Penerimaan Tandan Buah Segar (TBS) Kelapa Sawit Menggunakan Algoritma K-Means Clustering," Skripsi, Universitas Katolik Santo Thomas, Medan, 2026.
- A. A. Rahma and Z. Gusmanely, "Analisis Pengelompokan Kebun Kelapa Sawit PTPN IV Regional 4 Berdasarkan Produktivitas Tandan Buah Segar Menggunakan Metode K-Means," STATMAT: Jurnal Statistika Dan Matematika, vol. 7, no. 3, 2025.
- Y. P. Anggriani, A. Arif, and F. Febriansyah, "Implementasi Algoritma K-Means Clustering Dalam Menentukan Blok Tanaman Sawit Produktif Pada PT Arta Prigel," Jurnal Komtika (Komputasi Dan Informatika), vol. 8, no. 1, pp. 22–32, 2024.
- S. Ekie, N. Putra, and M. Furqan, "Penerapan Data Mining dalam Pengelompokan Kualitas Produk Kelapa Sawit Menggunakan Algoritma K-Means Clustering," Journal of Computing Engineering, System and Science, vol. 9, no. 2, 2024.
- W. J. Mawaddah, I. Gunawan, and I. P. Sari, "Implementation of Data Mining Algorithm for Clustering of Palm Oil Harvested Data," JOMLAI: Journal of Machine Learning and Artificial Intelligence, vol. 1, no. 1, pp. 43–54, 2022.