

SKRIPSI
IMPLEMENTASI METODE *K-MEANS CLUSTERING* UNTUK
PENGELOMPOKAN LAHAN SAWIT PRODUKTIF PADA PT.
ANDALAN ALAM SUMATRA



Oleh :
YULIANA SULISTIAWATI
20552010039

PROGRAM STUDI ILMU KOMPUTER
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SUMATRA SELATAN
2024

HALAMAN PENGESAHAN LAPORAN SKRIPSI

Nama : Delta Sari
NIM : 20552010044
Program Studi : Ilmu Komputer
Judul Skripsi : **Penerapan Data Mining Untuk Menganalisis Penjualan Kerupuk Menggunakan Metode Apriori di Restoran Beringin Anwar Sastro**

Telah dipertahankan dihadapan dewan penguji skripsi Program Studi Ilmu Komputer Fakultas Ilmu Komputer Universitas Sumatera Selatan dan dinyatakan **LULUS** pada 16 Juli 2024 di Ruang Lab Lantai 2.

Palembang, 16 Juli 2024

Ketua Penguji



Ubaidillah, S.Kom., M.Kom

NIP/NIDN. 0227127402

Penguji I



(Usep Teisnajaya S.Kom. M.Kom)

NIP/NIDN. 0221028101

Penguji II



(Riswa Dwipa S.Kom.M.Kom)

NIP/NIDN. 0215067403

Mengetahui,

Dekan



Dr. M. Taufik Roseno, S.T

NIP/NIDN. 0023037705

Ketua Program Studi



Serly Oktarina, S.Kom., M.Kom

NIP/NIDN. 0215109003

ABSTRAK

Penelitian ini berjudul implementasi metode k-means clustering untuk pengelompokan lahan sawit produktif pada PT Andalan Alam Sumatra. Tujuan dilakukan penelitian ini untuk dapat mengetahui lahan sawit mana yang dapat menghasilkan buah kelapa sawit lebih banyak. Metode penelitian yang digunakan adalah metode penelitian kuantitatif. Hasil penelitian ini menggunakan teknik pengumpulan data adapun data yang digunakan yaitu, observasi, wawancara, dan study literatur. Tools yang digunakan RapidMiner Studio. Analisis data dilakukan dengan teknik knowledge discovery database (KDD). KDD memiliki 5 tahapan yaitu: *Data Selection*, di tahap ini informasi tentang data panen dikumpulkan sebelum proses KDD dimulai. Atribut dipilih dan beberapa data dikonversi untuk mempermudah proses penambangan data. Data yang tidak lengkap atau berubah-ubah disaring agar hanya data dengan atribut lengkap yang digunakan. *Data cleaning*, Langkah pertama setelah seleksi data adalah membersihkan data. Ini termasuk menghapus duplikat, menangani nilai yang hilang (dalam hal ini, mengganti nilai kosong dengan 0 untuk bulan November), dan menghapus baris yang tidak lengkap. *Data Transformation*, data yang dipilih kemudian diubah ke format yang sesuai untuk kebutuhan analisis. Misalnya data luas blok (Ha) dikonversi menjadi rata-rata produksi perbulan setiap blok (A, B, C, dst). *Data Mining*, ini adalah tahapan penemuan pengetahuan utama, dimana teknik data mining diterapkan untuk mengidentifikasi pola dan tren dari data set besar. Dalam konteks ini, fokusnya adalah pada hubungan antara luas blok tanah, jumlah pohon, dan hasil produksi bulanan kelapa sawit. *Evaluation*, pada tahap ini, model yang telah dibuat dievaluasi. Dalam kasus ini algoritma *clustering* (khususnya *K-means*) digunakan untuk membentuk cluster berdasarkan data produksi kelapa sawit. Hasilnya, menunjukkan cluster-cluster yang berbeda berdasarkan karakteristik produksi tahunan dari setiap blok. Berdasarkan hasil pengujian tersebut dapat disimpulkan bahwa dari pengelompokan lahan sawit dengan k-means mengidentifikasi produktivitas, tinggi (cluster 3), sedang (cluster 2), rendah (cluster 1) untuk pemeliharaan efektif.

Kata Kunci : *K-Means Clustering*, Produktivitas Sawit, RapidMiner, Data Mining

ABSTRACT

This research is titled "Implementation of K-Means Clustering Method for Grouping Productive Palm Oil Land at PT Andalan Alam Sumatra." The purpose of this research is to determine which palm oil land can produce more palm oil fruit. The research method used is a quantitative research method. The results of this research use data collection techniques, including observation, interviews, and literature study. The tool used is RapidMiner Studio. Data analysis is conducted using the Knowledge Discovery Database (KDD) technique. KDD has five stages: Data Selection: In this stage, information about harvest data is collected before the KDD process begins. Attributes are selected, and some data is converted to facilitate the data mining process. Incomplete or inconsistent data is filtered so that only data with complete attributes is used. Data Cleaning: The first step after data selection is cleaning the data. This includes removing duplicates, handling missing values (in this case, replacing missing values with 0 for the month of November), and deleting incomplete rows. Data Transformation: The selected data is then transformed into a format suitable for analysis needs. For example, block area data (Ha) is converted into the average monthly production per block (A, B, C, etc.). Data Mining: This is the main knowledge discovery stage, where data mining techniques are applied to identify patterns and trends from large data sets. In this context, the focus is on the relationship between the land block area, the number of trees, and the monthly production yield of palm oil. Evaluation: In this stage, the model that has been created is evaluated. In this case, the clustering algorithm (specifically K-means) is used to form clusters based on palm oil production data. The results show different clusters based on the annual production characteristics of each block. Based on the test results, it can be concluded that the grouping of palm oil land with K-means identifies productivity as high (cluster 3), medium (cluster 2), and low (cluster 1) for effective maintenance.

Keywords: *K-Means Clustering, Palm Oil Productivity, RapidMiner, Data Mining*