

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Pertama kali karet di temukan daerah Subang Jawa Barat Indonesia yang mulai di tanam pada tahun 1862. Setelah 2 tahun berkembang karet mulai di tanam oleh pengurus kebun raya bogor untuk di koleksi. Perkebunan karet merupakan salah satu pemasok ekspor terbesar untuk negara,terlebih lagi untuk di wilayah Sumatra Selatan perkebunan karet merupakan mata pencarian masyarakat di daerah tertentu. Banyak di daerah Sumsel masyarakatnya bergantung hidup dengan menjadi petani karet. Akan tetapi seiring bejalannya waktu harga karet kian mengalami naik turun setiap tahun. Di bidang perkebunan karet ada suatu organisasi yang bernama Gapkindo ( gabungan perusahaan karet indonesia ) di bentuk pada tanggal 25 Mei 1971 bertujuan agar membangun produksi karet tertentu di wilayah Indonesia.

Dalam proses pengolahan statistik karet pada GAPKINDO mesih menggunakan manual dan belum ada grafik untuk membandingkan produksi pertahunnya.Semakin besarnya organisasi dan semakin banyaknya data produksi dalam pertahun untuk membuat laporan kenaikan produksi atau mengalami penurunan jumlah produksi karet.Menerapkan *data mining* yang nantinya dapat digunakan untuk memonitoring atau sebagai perbandingan oleh pihak GAPKINDO dalam proses rekap penjualan.

*Data mining* adalah bidang interdisipliner yang telah terintegrasi berbagai bidang seperti database, statistik, pembelajaran mesin dan bidang terkait lainnya, sehingga informasi berharga dan pengetahuan yang tersembunyi dalam jumlah besar data dapat diekstraksi (Panjaitan, 2018). *Data mining* banyak digunakan untuk mencari dan memperoleh informasi data yang akan digunakan,selain itu ada juga teknik-teknik dalam proses

pengolahannya seperti Klasterisasi (*clustering*), Regresi (*regression*), Asosiasi (*association*), dan Klasifikasi (*classification*).

*K-NN* merupakan salah satu algoritma klasifikasi yang cukup terkenal dan di terapkan untuk memprediksi kelas dari catatan atau sampel dengan kelas yang tidak di tentukan berdasarkan kelas dari data tetangganya (Panjaitan et al., 2018)

KNN terdiri atas tiga tahapan berikut ini :

1. Menghitung jarak rekam masukan terhadap semua catatan pelatihan.
2. Menembuat catatan sesuai jarak dan pemilihan tetangga terdekat.
3. Menenttukan kelas yang memiliki mayoritas diantara k-tetangga terdekat.

## **1.2 Rumusan Masalah**

Adapun rumusan masalah dalam penelitian ini berdasarkan uraian latar belakang ialah Bagaimana mengaplikasikan data mining dalam sistem penjualan karet di Sumsel ?

## **1.3 Batasan Masalah**

Agar penelitian ini dapat mudah dipahami penulis hanya berfokus pada:

1. Data yang diambil berupa penjualan karet tahun 2016-2019
2. Pengolahan data menggunakan metode *k-nearest neighbor*

## **1.4 Tujuan dan Manfaat Penelitian**

### **1.4.1 Tujuan Penelitian**

Adapun tujuan dalam penelitian ini adalah mengaplikasikan konsep *data mining* dalam sistem penjualan karet di Sumatera Selatan.

### **1.4.2 Manfaat Penelitian**

Manfaat dari penelitian ini berupa :

1. Membuat grafik produksi karet di Sumatera Selatan.

2. Dapat digunakan untuk memprediksi produksi karet yang akan datang.

## **1.5 Metode Penelitian**

### **1.5.1 Waktu dan Tempat Penelitian**

#### 1.5.1.1 Waktu Penelitian

Waktu penelitian yang akan dilakukan pada bulan Desember.

#### 1.5.1.2 Tempat Penelitian

Lokasi penelitian ini di GAPKINDO ( gabungan perusahaan karet indonesia ) Cabang Sumsel yang beralamat di jalan Hang Tuah 10/2375 Palembang,20135 Sumatera Selatan,Indonesia

### **1.5.2 Alat dan Bahan**

#### 1.5.2.1 Alat

Alat yang digunakan dalam penelitian ini :

1. Laptop
2. Microsoft Excel
3. Microsoft Word
4. Microsoft Visio
5. Software Rapid miner

#### 1.5.2.2 Bahan

Bahan yang digunakan untuk melakukan penelitian ini ialah laporan penjualan karet tahun 2016 sampai 2019.

### **1.5.3 Metode Penelitian**

Metode digunakan dalam penelitian ini adalah Metode *deskriptif kuantitatif* bertujuan menjelaskan gambaran yang menggunakan angka bertujuan menunjukkan karakteristik sendiri atau suatu kelompok (Syamsudin & Damiyanti: 2011). Metode ini dapat menilai berdasarkan sifat yang terlihat. Penelitian ini bertujuan untuk menggambarkan karakteristik sesuatu sebagaimana adanya.

Adapun ciri dari penelitian ini :

1. Kebanyakan menggunakan satu variabel untuk kegiatan operasionalnya.
2. Memungkinkan untuk menerapkan dua variabel atau lebih tetapi tidak untuk dihubungkan, dibandingkan, atau dicari akibatnya.
3. Analisis data yang di arahkan terhadap pencarian mean, persentase, atau modus.
4. Kegiatan data yang memungkinkan dapat di wakikan.
5. Analisis data dilakukan sesudah semua data cukup.

#### **1.5.4 Metode Pengumpulan Data**

Metode pengumpulan data pada penelitian ini :

1. Wawancara  
Merupakan interaksi antara dua orang untuk bertukar informasi dan ide melalui tanya jawab, agar mendapat informasi yang di perlukan dalam suatu pembahasan.
2. Dokumentasi  
Merupakan langkah mencari dan mengumpulkan data berbentuk catatan, gambar, video ataupun yang lainnya.
3. Studi pustaka  
Penelitian ini memerlukan teknik pengumpulan data dengan mengadakan studi penelaah terhadap buku, jurnal, dan laporan-laporan yang ada hubungannya dengan masalah yang dipecahkan pada penelitian ini.
4. Pengumpulan data penjualan
5. Analisa data transaksi penjualan dengan algoritma *K-Nearest Neighbor*.
6. Implementasi  
Pada penelitian ini peneliti mengimplementasikan proses hasil analisa data menggunakan algoritma *K-Nearest Neighbor* dan teknik klasifikasi, *Tools* yang digunakan adalah *rapidminer*.

### 1.5.5 Metode Analisis Data

Algoritma *K-Nearest Neighbor* (Pencarian tetangga terdekat) merupakan teknik klasifikasi yang sangat populer yang diperkenalkan oleh Fix dan Hodges, yang telah terbukti menjadi algoritma sederhana yang baik. KNN merupakan salah satu metode yang digunakan dalam pengklasifikasian dengan menggunakan algoritma supervised (Amalia, n.d.).

Algoritma KNN adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek yang diuji. Pengklasifikasian kerabat terdekat (nearest neighbor) didasarkan pada pembelajaran dengan analogi, yaitu dengan membandingkan data yang dites dengan kumpulan data pelatihan yang mirip dengan itu. Data pelatihan dijelaskan oleh atribut  $n$ . Setiap data merupakan titik dalam ruang- $n$  dimensi. Dengan cara ini, semua data pelatihan disimpan di ruang pola- $n$  dimensi.

Ketika diberi data yang tidak diketahui, pengklasifikasian  $k$ -nearest neighbor akan mencari ruang pola untuk data  $k$  pelatihan yang paling dekat dengan data yang tidak diketahui. Data  $k$  pelatihan ini adalah  $k$  "nearest neighbor" dari data yang tidak diketahui dekat atau jauhnya tetangga jarak biasanya dihitung berdasarkan jarak euclidean dengan rumus pada persamaan.

$$dist(x_1, x_2) = \sqrt{\sum_{i=1}^n (x_{1i} - x_{2i})^2}$$

Keterangan :

Dist	=	Jarak
X1	=	Nilai data latih
X2	=	Nilai data uji
n	=	Dimensi data
i	=	Variabel data

## **1.6 Sistematika Penulisan**

Sistematika penulisan yang digunakan dalam penulisan tugas akhir ini akan menjelaskan mengenai uraian secara garis besar isi dari setiap bab adalah sebagai mana sebagai berikut :

### **BAB I PENDAHULUAN**

Berisikan tentang latar belakang, perumusan masalah, batasan masalah, tujuan, manfaat, metode penelitian dan sistematika penulis.

### **BAB II LANDASAN TEORI**

Berisikan tentang teori secara umum tentang *data mining*, *knowledge Discovery in Database* (KDD), *K-Nearest Neighbor* (K-NN), *Rapidminer*, *Use Case*, dan penelitian sebelumnya.

### **BAB III METODELOGI PENELITIAN**

Berisikan tentang metode penelitian, lokasi penelitian, bahan penelitian, metode pengumpulan data, tahapan penelitian, pengolahan data mining,

### **BAB IV HASIL DAN PEMBAHASAN**

Berisikan tentang hasil observasi lapangan, pembahasan, implementasi rapid miner, grafik perbandingan.

### **BAB V PERANCANGAN APLIKASI**

Berisikan tentang desain, use case, desain tampilan admin, desain tampilan karyawan, desain tampilan masyarakat

### **BAB VI PENUTUP**

Berisikan tentang kesimpulan dan saran.