

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang

Jantung merupakan organ manusia yang berperan dalam sistem peredaran darah. Penyakit jantung adalah sebuah kondisi dimana jantung tidak dapat melaksanakan tugasnya dengan baik. Data WHO menyatakan bahwa sebanyak 7,3 juta penduduk dunia meninggal dikarenakan penyakit jantung. Meskipun penyakit jantung merupakan penyakit tidak menular, penyakit ini merupakan jenis penyakit yang mematikan nomor satu di dunia. Salah satu tipe penyakit jantung adalah *heart disease* yang menyebabkan serangan jantung. Jenis penyakit jantung yang lain termasuk katup jantung atau jantung yang tidak terpompa dengan baik dan menyebabkan gangguan jantung (H.Promotion, 2009).

Dalam bidang kesehatan banyak digunakan *Data Mining* misalnya mendiagnosis penyakit kanker payudara, penyakit jantung, penyakit diabetes dan lain-lain (Larose, 2005). Metode yang digunakan dalam mendiagnosis penyakit jantung misalnya *K-Nearest Neighbor (KNN)*, Neural Network dan lain-lain (Baraani, 2009).

Salah satu metode klasifikasi yang banyak digunakan adalah *K Nearest Neighbor* merupakan algoritma klasifikasi berdasarkan kedekatan jarak suatu data dengan data yang lain. Pada algoritma KNN kelas data diitung dari K jarak terdekat. Nilai jarak inilah yang digunakan sebagai kedekatan atau kemiripan antara data uji dengan data latih. Nilai K pada KNN artinya K adalah data terdekat dari data uji. (Baraani, 2009).

Dalam banyak kasus yang berkaitan dengan pengenalan pola maupun klasifikasi, *missing data* merupakan permasalahan yang dapat mempengaruhi hasil klasifikasi. *Missing data* menjadi kelemahan umum dalam klasifikasi dimana hampir semua metode klasifikasi hanya dapat bekerja pada data yang lengkap. Hal ini yang mendorong perlunya dicari metode khusus untuk penanganan terhadap permasalahan yang berkaitan dengan *missing data*. Metode yang umum digunakan adalah dengan cara membuang data yang mengandung missing (case

deletion). Beberapa metode telah dikembangkan khusus untuk menangani missing data. Salah satu metode tersebut adalah teknik imputasi. Teknik imputasi sendiri dibedakan menjadi beberapa metode, yang paling populer adalah mean, median, modus dan klusterisasi.

Dari latar belakang tersebut penelitian ini bertujuan untuk mendiagnosis penyakit jantung koroner menggunakan algoritma KNN sebagai algoritma klasifikasi untuk mendapatkan hasil diagnosis penyakit jantung dengan melakukan imputasi pada missing value. Maka dari itu, penelitian ini menggunakan perbandingan algoritma klasifikasi berbasis imputasi missing data dengan mean, median, modus dan KNN untuk mendiagnosis penyakit jantung koroner agar mengetahui kinerja dari metode tersebut dengan menganalisa klasifikasi data yang dihasilkan.

### **1.2 Rumusan Masalah**

Dari latar belakang dapat dirumuskan permasalahan yaitu, Bagaimana akurasi hasil dari klasifikasi KNN dengan metode imputasi mean, median, modus dan 1-NN pada data penyakit jantung koroner?

### **1.3 Maksud dan Tujuan Penulisan**

Maksud dan tujuan pada penulisan ini adalah membandingkan performa metode imputasi mean, median, modus dan 1-NN pada penyakit jantung koroner dengan KNN.

### **1.4 Batasan Masalah**

1. Menggunakan metode mean, median, modus dan 1-NN pada saat mengimputasi data.
2. Menggunakan data penyakit jantung koroner.

## 1.5 Sistematika Pembahasan

Secara garis besar proposal ini terdiri dari 4 (Empat) bab dengan beberapa sub bab. Agar mendapat arah dan gambaran yang jelas mengenai hal yang tertulis, berikut ini sistematika penulisannya secara lengkap:

### **BAB I : PENDAHULUAN**

Pada bab ini membahas latar belakang, perumusan masalah, maksud dan tujuan penulisan, batasan masalah, dan sistematika penulisan.

### **BAB II : TINJAUAN PUSTAKA**

Bab ini berisi uraian teori-teori yang mendasari penelitian secara detail, berupa metode, algoritma, aplikasi yang digunakan, atau definisi yang berkaitan dengan topik penelitian.

### **BAB III : METODELOGI PENELITIAN**

Bab ini menjelaskan tahapan penelitian, deskripsi pendekatan teoritis, desain eksperimen, deliniasi wilayah kajian, sumber data, teknik pengumpulan data, teknik pengolahan data, dan teknik analisis data, yang digunakan untuk mencapai tujuan penelitian. Untuk setiap proses yang dijalankan, harus dijelaskan dasarnya.

### **BAB IV : Hasil dan Pembahasan**

Bab ini berisi hasil dan pembahasan dari penelitian berupa data-data dan tabel hasil dari perhitungan itu sendiri. Pada bab ini juga menjelaskan data dari penelitian. Bab ini berisi persiapan data, tahap praproses, pembagian data, tahap klasifikasi K-NN, dan evaluasi.

### **Bab V : Penutup**

Pada bab ini menjelaskasn kesimpulan dan saran dari penelitian yang sudah di buat.

