

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

*Data mining* yaitu serangkaian proses untuk mengekstraksi pengetahuan dan pola dari kumpulan data. *Data mining* bertujuan untuk memecahkan masalah dan berkontribusi pada pengambilan Keputusan dengan menganalisis data yang tersedia dalam *database*. Hasil dari proses *data mining* ini dapat digunakan untuk meningkatkan kualitas pengambilan Keputusan di berbagai bidang. Terdapat beberapa teknik yang dapat kita gunakan dalam dalam *data mining* untuk menemukan sebuah informasi baru, salah satu teknik dari *data mining* adalah klasifikasi.

Klasifikasi adalah teknik mengelompokkan data secara terstruktur sesuai dengan aturan dan kaidah yang telah ditentukan sebelumnya. Dalam konteks diagnosis penyakit jantung, klasifikasi merupakan pendekatan yang digunakan dengan menerapkan algoritma klasifikasi. Tujuannya adalah untuk meningkatkan akurasi dan efektivitas dalam mengklasifikasikan kondisi pasien. Penggunaan klasifikasi dalam diagnosis penyakit jantung menjadi bagian integral dalam upaya untuk menyempurnakan proses pengklasifikasiannya.

Salah satu metode yang sering digunakan dalam klasifikasi data, antara lain ialah metode *K-Nearest Neighbor* (KNN), *Decision Tree*, *NeighborWeighted K-Nearest Neighbor* (NWKNN), *Modified K-Nearest Neighbor* (MKNN) dan algoritma genetika. Banyak peneliti yang menggunakan algoritma *Modified K-Nearest Neighbor* dalam pengklasifikasian, diantaranya yang dilakukan oleh (Putra, Murdiansyah, & Aditsania, 2019) Berdasarkan hasil pengujian, ditemukan bahwa semakin tinggi nilai K, akurasi sistem cenderung menurun. Akurasi tertinggi tercapai pada K=1, dengan Tingkat akurasi mencapai 95.61%. Selain itu, penggunaan jumlah data latih yang lebih besar juga berdampak positif pada akurasi sistem. Semakin banyak data latih yang digunakan, semakin tinggi akurasi yang dicapai. Hal ini disebabkan oleh peningkatan jumlah data yang dapat digunakan untuk proses pembelajaran, sehingga sistem dapat membuat prediksi yang lebih tepat. Pada pengujian ini, akurasi tertinggi terjadi saat menggunakan K=1 dengan

90% data latih, mencapai 97.61%. oleh karena itu, pengaturan nilai K dan jumlah data latih memainkan peran penting dalam meningkatkan akurasi sistem klasifikasi dalam penelitian ini. Selain itu penelitian dari (Fernanda, Ratnawati, & Adikara, 2017) dari hasil penelitian ini menggunakan lima belas parameter, yang merupakan gejala penyakit yang dialami pasien diabetes mellitus tipe 1 dan tipe 2. Hasil pengujian menunjukkan akurasi sebesar 93,33% dan Tingkat kesalahan sebesar 6,67%. Nilai K terbaik adalah 3, nilai alfa adalah 0,5, dan jumlah data latih dan data uji adalah 100 dan 30 data. Sehingga dari dua penelitian yang telah dibuat sebagai acuan penulis, dapat disimpulkan bahwa metode *Modified K-Nearest Neighbor* termasuk metode dengan tingkat akurasi yang tinggi dengan akurasi hampir mencapai 100%. Dengan alasan tersebut penulis mengambil metode *Modified K-Nearest Neighbor* sebagai metode penelitian tugas akhir ini.

Dalam penelitian ini, diaplikasikan dengan metode *Modified K-Nearest Neighbor* (MKNN) untuk mengklasifikasikan data. MKNN bekerja dengan melihat kelompok objek dalam data latih yang paling dekat atau serupa dengan objek dalam data baru atau data uji. Dalam konteks ilmu Kesehatan, MKNN juga dapat digunakan untuk menemukan Solusi bagi pasien baru dengan merujuk pada kemiripan dengan kasus pasien yang sudah ada sebelumnya.

Jantung berperan sebagai motor utama dalam sistem peredaran darah manusia. Penyakit jantung, pada dasarnya, adalah kondisi di mana jantung tidak dapat menjalankan fungsi-fungsinya secara optimal. Penyakit jantung adalah penyebab kematian yang paling umum.

Menurut Kementerian Kesehatan Republik Indonesia (Kemenkes) berdasarkan Global Burden of Disease dan Institute for Health Metrics and Evaluation (IHME) dari tahun 2014 hingga 2019, penyakit jantung menjadi penyebab kematian utama di Indonesia. Data Riset Kesehatan Dasar (Riskesdas) pada tahun 2013 dan 2018 menunjukkan adanya tren peningkatan kasus penyakit jantung. Prevalensi penyakit jantung meningkat dari 0,5% pada tahun 2013 menjadi 1,5% pada tahun 2018.

Bahkan penyakit jantung telah menjadi salah satu beban biaya terbesar dalam sektor kesehatan. Berdasarkan data BPJS Kesehatan pada 2021 pembiayaan kesehatan terbesar dikaitkan dengan penyakit jantung, mencapai jumlah Rp.7,7

triliun. Bahkan, hingga tahun 2022, penyakit jantung tetap menjadi penyebab utama kematian di Indonesia.

Menurut *World Health Organization* (WHO) tahun 2021 diperkirakan 17,8 juta kematian akibat penyakit jantung. Artinya satu dari tiga kematian di dunia setiap tahunnya disebabkan oleh penyakit jantung dan menjadi penyebab kematian tertinggi di Indonesia.

Mengingat bahwa banyak kematian yang disebabkan oleh penyakit jantung seperti data diatas. Maka membuktikan bahwa di Indonesia masih kurangnya pengetahuan dan informasi yang dapat menjaga kesehatan jantung. Oleh karena itu dibutuhkan analisis penyakit jantung yang melihat peluang penyakit jantung pada pasien berdasarkan atribut dan data sampel pasien yang akan dimasukkan kedalam sistem. Dan atribut ataupun sampel yang akan diambil mengacu pada Rumah Sakit Sari Mutiara Lubuk Pakam yang berada di Kabupaten Deli Serdang, Sumatera Utara.

Analisis ini bertujuan untuk membantu ahli jantung pada Rumah Sakit Sari Mutiara Lubuk Pakam menyusun atau mengkategorikan data ke dalam pola atau tema. Selanjutnya, analisis tersebut menjelaskan kategori atau pola tertentu, dan menemukan hubungan antara berbagai data. Salah satu pendekatan yang digunakan adalah dengan adanya sistem klasifikasi yang akan dibuat dalam penelitian tugas akhir ini. Dengan demikian, sistem klasifikasi ini akan mempermudah penyusunan data yang kompleks menjadi informasi yang lebih terstruktur dan berguna bagi para ahli jantung.

Berdasarkan latar belakang yang dijelaskan sebelumnya, penulis memutuskan melakukan penelitian untuk tugas akhir dengan judul **“Klasifikasi Penyakit Jantung Menggunakan Metode *Modified K-Nearest Neighbor*”**.

## **1.2 Rumusan Masalah**

Dengan mempertimbangkan latar belakang di atas, rumusan masalah penelitian ini adalah sebagai berikut :

1. Bagaimana merancang sistem klasifikasi penyakit jantung pada Rumah Sakit Sari Mutiara Lubuk Pakam.

2. Bagaimana menerapkan algoritma *Modified K-Nearest Neighbor* dalam mengklasifikasikan penyakit jantung.
3. Bagaimana hasil akurasi dari metode *Modified K-Nearest Neighbor* dalam mengklasifikasikan penyakit jantung.

### **1.3 Batasan Masalah**

Untuk menjaga fokus penelitian ini, batasan masalah pada penelitian tugas akhir ini adalah sebagai berikut :

1. Data yang digunakan adalah data hasil rekam medis pada Rumah Sakit Sari Mutiara Lubuk Pakam tahun 2021 dan 2023.
2. Bahasa pemrograman yang digunakan PHP dengan menggunakan *database* MySQL.
3. Metode yang digunakan algoritma *Modified K-Nearest Neighbor*
4. Hasil outputnya berupa jantung koroner, aritmia dan gagal jantung.

### **1.4 Tujuan Penelitian**

Berdasarkan latar belakang dan rumusan masalah di atas, berikut adalah tujuan penelitian yang dilakukan penulis :

1. Dapat mengetahui bagaimana rancangan sistem klasifikasi penyakit jantung pada Rumah Sakit Sari Mutiara Lubuk Pakam.
2. Dapat mengetahui bagaimana penerapan algoritma *Modified K-Nearest Neighbor* dalam mengklasifikasikan penyakit jantung.

### **1.5 Manfaat Penelitian**

Adapun manfaat dari penelitian tugas akhir ini adalah membantu masyarakat dalam mengetahui informasi seputaran kondisi kesehatan penyakit jantung.