

## BAB I PENDAHULUAN

### 1.1 Latar Belakang

Salah satu organ tubuh yang bekerja pada sistem peredaran darah manusia adalah Jantung. Penyakit jantung yakni suatu situasi di mana jantung tidak mampu berfungsi dengan normal. Penyebab kematian di seluruh dunia pertahun adalah penyakit kardiovaskular, ialah penyakit yang ditimbulkan oleh kegagalan fungsi jantung serta pembuluh darah (Kementrian Kesehatan RI, 2014). Sekitar 17,3 juta orang meninggal dunia ditimbulkan penyakit kardiovaskular pada tahun 2008. Lebih dari 3 juta kematian terjadi sebelum usia 60 tahun dan mampu dicegah. Jumlah kematian “dini” akibat penyakit jantung berkisar dari 4% di negara-negara pendapatan tinggi dan untuk 42% di negara-negara pendapatan rendah. Tekanan darah tinggi menyebabkan hingga 9,4% kematian di seluruh dunia pertahun, kemudian setidaknya 45% kematian penyakit jantung dan 51% kematian stroke. Jumlah kematian akibat penyakit kardiovaskular, khususnya penyakit jantung koroner dan stroke akan terus berkembang pesat sebesar 23,3 juta kematian di tahun 2030 (Kemenkes RI, 2014).

Ilmu komputer memiliki banyak disiplin ilmu salah satunya adalah ilmu tentang *data mining*. *Data mining* adalah proses mengekstraksi data di dalam database dengan menggunakan teknik tertentu untuk mendapatkan pengetahuan atau informasi baru. Teknik yang dipakai untuk mengekstrak informasi baru dalam penambangan data meliputi estimasi, *clustering*, asosiasi, prediksi dan klasifikasi [1]. Teknik klasifikasi dapat dilakukan dengan menggunakan berbagai metode termasuk metode klasifikasi *Naive bayes classifier* (NBC). Metode klasifikasi ini memperkirakan nilai probabilitas setiap kelas dan mengklasifikasikan secara linier dengan cara yang efektif [2].

Teknik klasifikasi ini dapat digunakan untuk membantu dalam mengidentifikasi penyakit jantung berdasarkan kondisi manusia terhadap atribut klasifikasi untuk penyakit jantung. Teknik klasifikasi digunakan untuk mengidentifikasi apakah seorang ini mengidap penyakit jantung atau tidak megidap penyakit jantung. Metode yang dipakai untuk mengidentifikasi yakni metode *Naïve*

*Bayes Classifier* karena memiliki hasil kinerja yang relatif kuat, mudah diimplementasikan, cepat dan akurat [2].

Sebuah penelitian tahun 2017 bertajuk “Diagnosa Prediksi Penyakit Jantung Dengan Model Algoritma Naive Bayes dan Algoritma C4.5”, membandingkan algoritma Naive Bayes dengan Algoritma C4.5 untuk menentukan algoitma yang sangat akurat untuk memprediksi penyakit jantung. Kesimpulan dari pengujian penelitian ini, algoritma *Naïve Bayes* mempunyai nilai akurasi tertinggi sebesar 86.67% dan algoritma C4.5 mempunyai nilai akurasi sebesar 83.70% [3]. Penelitian kedua yang bertajuk “*The Prediction Of Heart Disease Using Naïve Bayes Classifier*” tahun 2019, berkesimpulan bahwa metode *naïve bayes* dari penambahan data dapat mengklasifikasikan penyakit jantung atau kardiovaskular. Dari observasi ini didapatkan tingkat akurasi dari prediksi *naïve bayes* adalah 98,8636%. Penelitian diselesaikan mengimplementasikan algoritma *Naïve Bayes* dan aplikasi WEKA yang diterapkan pada dataset penyakit jantung. Tantangan utama dalam data *mining* adalah membangun pengklasifikasi yang akurat dan efisien secara komputasi untuk aplikasi medis [4].

Proses untuk mengimplementasikan metode *Naive Bayes Classifier* pada sistem identifikasi penyakit jantung, dilakukan dengan mengumpulkan beberapa data latih (*training*) yang berisikan pengalaman kondisi tubuh orang-orang dalam jumlah yang banyak untuk melakukan klasifikasi. Oleh karena itu penulis bermaksud menerapkan klasifikasi penyakit jantung dengan judul “Identifikasi Penyakit Jantung Menggunakan Metode *Naïve Bayes Classifier*” Sistem yang dibuat sudah dapat mengklasifikasi penderita penyakit jantung atau non jantung dari gejala klinis penyakit jantung, yang mana telah menghasilkan performa yang rata-rata pengujian 5 *k-fold cross validation* diperoleh nilai akurasi sebesar 79.24%, presisi 78.77%, recall 80.74% dan f-measure 79.72%.

## **1.2 Perumusan Masalah**

### **1.2.1 Permasalahan**

Penyakit yang ditimbulkan akibat kegagalan fungsi jantung dan pembuluh darah disebut penyakit jantung. Penyakit tersebut dipengaruhi oleh beberapa faktor. Kurangnya akses untuk mencari informasi tentang penyakit jantung ini dapat meningkatkan angka kematian setiap tahunnya. Kerena itu perlu dibutuhkan sebuah

sistem klasifikasi yang dapat memberikan informasi tentang penyakit jantung serta dapat melakukan pengecekan klasifikasi secara dini tentang penyakit jantung.

### **1.2.2 Solusi Permasalahan**

Berdasarkan permasalahan tersebut maka di bangun sistem identifikasi penyakit Jantung menggunakan metode *Naïve Bayes Classifier* untuk mengurangi angka kematian dari penyakit Jantung. Metode ini dapat melakukan klasifikasi dengan mengimplementasikan data latih (training) tentang Riwayat penyakit jantung serta gejala-gejalanya.

### **1.2.3 Pertanyaan Penelitian**

Dalam penelitian ini, terdapat pertanyaan penelitian yang diperlukan selama proses analisis hasil pengujian. Berikut pertanyaan pada penelitian ini :

Berapa akurasi dari metode *Naïve Bayes Classifier* dalam penelitian ini?

## **1.3 Tujuan dan Manfaat**

### **1.3.1 Tujuan**

Observasi ini bertujuan untuk mengetahui presentase akurasi metode *Naive bayes classifier* pada Identifikasi Penyakit Jantung.

### **1.3.2 Manfaat**

Peneliti dapat mengetahui metode *Naive bayes classifier* sebagai metode yang tepat untuk membangun sebuah sistem identifikasi penyakit jantung , dan menjadikan penerapan metode *Naïve Bayes Classifier* ini sebagai bahan literatur untuk penelitian-penelitian teknik klasifikasi selanjutnya.

## **1.4 Batasan-Batasan**

Sesuai rumusan masalah yang ada, maka penulis membatasi masalah sebagai berikut.

1. Observasi ini menggunakan metode *Naïve Bayes Classifier*.
2. Dataset yang dipakai ialah data yang di unduh dari kaggle, oleh Manu Siddhartha di Allahabad Bank Kolkata, West Bengal, India. Dimana terdapat data test dan data training sebanyak 11 atribut. Data yang dipergunakan adalah data tahun 2019 berjumlah 1190 data.
3. Penelitian ini dibangun berbasis website, bahasa pemrograman yang dipakai ialah PHP.
4. Observasi ini menggunakan database MySQL.

5. Atribut “Identifikasi Penyakit Jantung menggunakan *Naïve Bayes Classifier*” terdapat 11 atribut : *age* (umur), *sex* (jenis kelamin), *chest pain type* (jenis nyeri dada), *resting blood pressure* (tekanan darah pasien), *cholesterol* (kolesterol), *fasting blood sugar* (kadar gula darah saat puasa), *resting ecg* (status elektrokardiogram), *max heart* (denyut jantung maksimum), *exercise angina*, *oldpeak* (penurunan ST karena olahraga), *ST slope* (kemiringan segmen ST).
6. Kelas dari “Identifikasi Penyakit Jantung menggunakan *Naïve Bayes Classifier*” adalah *Non* jantung dan jantung.

### **1.5 Sistematika Proposal**

Sistematika penyusunan proposal ini dibuat untuk mengetahui deskripsi umum dari isi proposal

#### **BAB I PENDAHULUAN**

Pada Bab I menjelaskan latar belakang, rumusan masalah, tujuan dan manfaat, batasan-batasan masalah, dan sistematika skripsi pada akhir bab ini.

#### **BAB II KAJIAN PUSTAKA**

Pada Bab II menjelaskan tentang landasan teori yang berisi mengenai materi-materi terkait yang akan dibahas pada penelitian ini serta referensi dari penelitian sebelumnya menjadi acuan pada penelitian ini.

#### **BAB III METODE USULAN**

Dalam bab ini, mendiskusikan tentang jenis dan sumber data, tahapan penelitian, arsitektur sistem dan rancangan pengujian.

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

Pada bab ini menjelaskan mengenai implementasi pembuatan sistem berdasarkan rancangan arsitektur yang dibuat sebelumnya.

#### **BAB V PENUTUP**

Dalam bab ini, berisikan kesimpulan penelitian yang disesuaikan dengan pertanyaan penelitian, dan saran mengenai metode dan hal-hal apa yang perlu dikembangkan kedepannya.