

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

*Data Mining* merupakan disiplin ilmu yang mempelajari metode untuk mengekstrak pengetahuan atau menemukan pola dari suatu data (Han dan Kamber, 2006). Penerapan *Data Mining* sudah dilakukan pada beberapa bidang di antaranya bisnis, ilmu pengetahuan dan kesehatan. Pada bidang kesehatan, *Data Mining* dapat dimanfaatkan untuk memprediksi suatu penyakit dari data rekam medis pasien. Dengan metode klasifikasi pada *Data Mining*, data seperti umur, jenis kelamin, tekanan darah dan atribut lainnya, dapat digunakan untuk memprediksi kemungkinan pasien terkena suatu penyakit.

Beberapa penelitian sudah dilakukan terkait penerapan *Data Mining* untuk memprediksi suatu penyakit. Salah satunya yaitu mengembangkan suatu sistem pendukung keputusan untuk memprediksi penyakit jantung (Subbalakshmi, dkk. 2011). Pada penelitian tersebut, data yang digunakan adalah data penyakit jantung *Cleveland Clinic Foundation* yang berasal dari *University of California Irvine (UCI) repository* di mana seluruh atribut pada data digunakan untuk memprediksi penyakit jantung.

Adapun metode yang dapat digunakan dalam bidang data yang dapat melakukan proses prediksi yakni metode *Data Mining* menggunakan teknik klasifikasi. Klasifikasi adalah teknik pada *Data Mining* yang melakukan proses analisis data untuk menghasilkan model untuk menggambarkan kelas-kelas yang ada dalam data (Han, Kamber & Pei, 2012). Model-model ini disebut *classifier*. Oleh karena itu, *classifier* ini akan digunakan untuk menyusun kelas-kelas yang terkandung dari data, misalnya untuk *Decision Tree* maka kelas-kelas tersebut digambarkan dalam bentuk pohon.

Algoritma *Decision Tree* digunakan untuk mempelajari klasifikasi dan prediksi model dari data dan menggambarkan hubungan antara variabel atribut  $x$  dan variabel target  $y$  sebagai pohon (Ye, 2014). Sebuah *Decision Tree* adalah struktur

seperti *flowchart* di mana setiap *node* dalam (*node non-leaf* atau bukan *node* terluar) adalah pengujian variabel atribut, setiap cabang adalah hasil pengujian, sedangkan *node* terluar yaitu daun, adalah labelnya.

Terdapat penelitian terbaru dari Johns Hopkins Bloomberg School of Public Health pada 14 Maret 2022 menyatakan bahwa orang yang terinfeksi COVID-19 memiliki risiko lebih tinggi untuk masalah jantung. Risiko yang tetap ada bahkan pada orang yang relatif sehat lama setelah penyakit itu berlalu. Temuan utamanya adalah bahwa orang dengan COVID-19 memiliki risiko lebih tinggi dari semua jenis masalah jantung dalam satu tahun. Itu termasuk aritmia (detak jantung tidak teratur atau jantung berdetak terlalu cepat atau terlalu lambat) dan fibrilasi atrium (irama jantung cepat dalam pola tertentu). Dari penemuan tersebut ada kemungkinan bahwa virus itu sendiri dan respons imun terhadapnya menyebabkan peradangan hebat yang kemudian mengenai jantung. Ada kemungkinan bahwa COVID-19 dapat menyerang sel-sel endotel yang melapisi pembuluh darah jantung. Beberapa dari sel-sel ini mungkin mati dan akhirnya memfasilitasi pembentukan bekuan darah dan penyumbatan arteri atau pembuluh jantung. Ada beberapa mekanisme lain yang berputar di sekitar sesuatu yang disebut reseptor ACE. Virus memiliki sesuatu yang disebut protein lonjakan, yang seperti kunci yang menggunakan kunci reseptor ACE. Itu memungkinkan virus masuk ke sel, termasuk sel jantung.

Berdasarkan faktor-faktor yang belum diketahui secara pasti faktor yang mana yang menjadi pengaruh terbesar dari meningkatnya risiko penyakit jantung pada penderita COVID-19, penelitian ini bertujuan untuk melakukan identifikasi dan ekstraksi dari informasi yang sudah ada untuk mendapatkan pengetahuan untuk mengetahui faktor manakah yang paling mempengaruhi risiko penyakit jantung pada penderita COVID-19 untuk dilakukan prediksi risiko penyakit jantung pada penderita COVID-19. Oleh karena itu peneliti mengambil judul “ANALISIS *DATA MINING* KLASIFIKASI DENGAN ALGORITMA *DECISION TREE* UNTUK MEMPREDIKSI RISIKO PENYAKIT JANTUNG PADA PENDERITA COVID-19”.

## **1.2 Identifikasi Dan Rumusan Masalah**

### **1.2.1 Identifikasi Masalah**

Identifikasi masalah yang terdapat pada penelitian ini yakni:

1. Belum jelasnya dan belum pastinya atribut kesehatan yang mana yang menjadi faktor penentu paling dominan dan yang tidak dominan pada faktor penentu keputusan tingginya risiko penyakit jantung pada penderita COVID-19.
2. Belum adanya penelitian yang menggunakan model *Data Mining* klasifikasi menggunakan algoritma *Decision Tree* untuk melakukan pencarian atribut paling dominan dan kurang dominan yang menjadi faktor terhadap penentuan ada tidaknya risiko penyakit jantung pada penderita COVID-19.
3. Belum adanya penelitian yang menggunakan model *Data Mining* klasifikasi menggunakan algoritma *Decision Tree* untuk melakukan prediksi terhadap penentuan ada tidaknya risiko penyakit jantung pada penderita COVID-19.

### **1.2.2 Rumusan Masalah**

Rumusan masalah dalam penelitian ini di antaranya:

1. Bagaimana menerapkan metode Klasifikasi *Data Mining* dengan algoritma *Decision Tree* untuk melakukan prediksi risiko penyakit jantung pada penderita COVID-19.
2. Atribut atau variabel apa saja yang menjadi faktor paling berpengaruh dalam melakukan prediksi risiko penyakit jantung pada penderita COVID-19 menggunakan metode Klasifikasi *Data Mining* dengan algoritma *Decision Tree*.
3. Bagaimana tingkat akurasi metode Klasifikasi *Data Mining* dengan algoritma *Decision Tree* untuk melakukan prediksi risiko penyakit jantung pada penderita COVID-19 berdasarkan hasil *Data Mining* yang diperoleh.

### **1.3 Tujuan Penelitian**

1. Untuk merancang program yang menerapkan metode Klasifikasi *Data Mining* dengan algoritma *Decision Tree* untuk melakukan prediksi risiko penyakit jantung pada penderita COVID-19.
2. Untuk mengetahui atribut atau variabel apa saja yang menjadi faktor paling berpengaruh dalam melakukan prediksi risiko penyakit jantung pada penderita COVID-19 menggunakan metode Klasifikasi *Data Mining* dengan algoritma *Decision Tree*.
3. Untuk mengetahui tingkat akurasi metode Klasifikasi *Data Mining* dengan algoritma *Decision Tree* untuk melakukan prediksi risiko penyakit jantung pada penderita COVID-19.

### **1.4 Pembatasan Masalah**

1. Data yang diolah hanya data-data yang berhubungan langsung dengan data penyakit jantung dan data penderita COVID-19.
2. Hanya membahas pengolahan data yang didapatkan dari instansi penyedia data kesehatan dan Rumah Sakit Umum Daerah Kabupaten Sumedang.

### **1.5 Kegunaan Penelitian**

#### **1.5.1 Kegunaan Pengembangan Ilmu**

1. Bagi pengembangan ilmu pengetahuan, dapat memberikan suatu karya peneliti baru yang dapat mendukung dalam pengembangan sistem informasi.
2. Bagi dunia kesehatan, dapat memberikan pengetahuan baru berdasarkan objek penelitian yang digunakan.
3. Bagi Peneliti Dapat Menambah Wawasan Dengan Mengaplikasikan Ilmu Yang Telah Diperoleh Secara Teori Di Lapangan.
4. Bagi Peneliti Lain Dapat Dijadikan Sebagai Acuan Terhadap Pengembangan Atau pun Pembuatan Dalam Penelitian Yang Sama.

### **1.5.2 Kegunaan Operasional**

Bagi pihak yang terkait dengan penelitian ini, dapat mengembangkan sistem informasi yang menerapkan metode Klasifikasi *Data Mining* dengan algoritma *Decision Tree* untuk melakukan prediksi yang akurat dan relevan dan tepat waktu agar dapat membantu dalam proses pengambilan keputusan.

## **1.6 Sistematika Penulisan**

### **BAB I PENDAHULUAN**

Bab ini akan membahas mengenai latar belakang, identifikasi masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan.

### **BAB II KAJIAN PUSTAKA**

Bab ini membahas tentang teori-teori yang digunakan dalam penelitian.

### **BAB III OBJEK DAN METODOLOGI PENELITIAN**

Bab ini membahas tentang metode penelitian mencakup materi dan sasaran penelitian, teknik pengumpulan data, dan penjelasan metode yang digunakan dalam penelitian.

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

Bab ini membahas tentang hasil analisis *Data Mining* klasifikasi dengan algoritma *Decision Tree* untuk memprediksi risiko penyakit jantung pada penderita COVID-19.

### **BAB V KESIMPULAN**

Bab ini membahas mengenai kesimpulan akhir dari penelitian yang dilakukan.