

## **BAB I**

### **PENDAHULUAN**

#### **1.1 Latar Belakang**

Salah satu penyakit yang sering terjadi dan tidak menular (PTM) adalah penyakit jantung. Penyakit jantung tersebut merupakan penyakit merupakan kondisi yang terjadi ketika pembuluh darah utama yang menyuplai darah ke jantung mengalami kerusakan dan tidak dapat bekerja dengan semestinya hal tersebut dikarenakan banyak faktor. Tumpukan kolesterol pada pembuluh darah serta proses peradangan diduga salah satu menjadi faktor penyebab penyakit ini. Ketika terjadi penumpukan kolesterol (plak), pembuluh darah koroner akan menyempit sehingga aliran darah dan suplai oksigen menuju jantung pun akan terhambat sehingga mempengaruhi kestabilan proses bekerja jantung. Kurangnya aliran darah ini akan menyebabkan rasa nyeri pada dada (angina) dan sesak napas, hingga suatu saat terjadi hambatan total pada aliran darah menuju jantung atau yang disebut juga dengan serangan jantung. Menurut banyak penelitian penyakit jantung koroner termasuk salah satu penyebab kematian tertinggi di dunia.

Federasi Jantung Dunia memperkirakan angka kematian akibat penyakit jantung koroner di Asia Tenggara mencapai 1,8 juta kasus pada tahun 2014. Di Indonesia sendiri pada tahun 2013 tercatat ada setidaknya 883.447 orang yang terdiagnosis penyakit jantung koroner di Indonesia dengan mayoritas penderita berusia 55-64 tahun. Angka kematian akibat penyakit jantung pun menjadi cukup tinggi, yakni sekitar 45 persen dari seluruh angka kematian di Indonesia[1].

Pada acara peringatan hari jantung sedunia (*world heart day*) yang dicetuskan pertama kali oleh World Heart Federation pada tahun 2000 menginformasikan kepada orang-orang di seluruh dunia bahwa penyakit jantung dan stroke merupakan penyebab utama kematian di dunia yang saat ini di klaim mencapai 17,3 juta kematian setiap tahunnya. Angka kematian akibat penyakit jantung tersebut diperkirakan akan terus meningkat hingga mencapai 23,3 juta pada tahun 2030. Berdasarkan Hasil Riset Kesehatan Dasar (Riskesdas) Kemenkes RI Tahun

2013, prevalensi penyakit jantung koroner di Indonesia mencapai 0,5% dan gagal jantung sebesar 0,13% dari total penduduk berusia 18 tahun keatas sehingga harus penyakit tersebut harus selalu diwaspadai ([dinkes.inhukab.go.id](http://dinkes.inhukab.go.id), 2015). Banyak faktor yang dapat meningkatkan risiko terkena penyakit jantung. Faktor risiko tersebut terdiri dari faktor risiko yang tidak dapat dimodifikasi seperti riwayat keluarga, umur serta jenis kelamin dan faktor risiko yang dapat dimodifikasi seperti hipertensi, kebiasaan merokok, penyakit diabetes, dislipidemia, obesitas, kurang aktifitas fisik, pola makan serta stres. Penyakit pembuluh darah, atau disebut dengan penyakit kardiovaskular atau disebut juga penyakit jantung. Penyakit ini berhubungan dengan proses aterosklerosis, yaitu suatu kondisi pada organ tubuh yang terjadi ketika zat yang disebut plak menumpuk di dinding arteri. Penumpukan ini menyebabkan penyempitan pada arteri sehingga aliran darah terhenti atau tersumbat dan tidak dapat beredar dengan semestinya, hal ini dapat mengakibatkan serangan jantung atau stroke [2].

Salah satu bagian penting dari pengobatan atau tindakan medis adalah pengambilan keputusan dan proses klasifikasi atau prediksi pada suatu hal yang menjadi fokus seperti pendeteksian penyakit, namun klasifikasi medis atau prediksi biasanya merupakan proses yang sangat kompleks dan sulit dilakukan jika tak dapat mengetahui metode yang tepat dan terbaik dalam memberikan solusinya[3] Tantangan yang dihadapi oleh organisasi kesehatan adalah mendiagnosa pasien dengan benar, diagnosa atau prediksi yang buruk dapat menyebabkan konsekuensi yang mendatangkan malapetaka yang kemudian tidak dapat diterima. Untuk menjawab tantangan tersebut beberapa penelitian telah dilakukan dalam bidang [4]kesehatan untuk mendapatkan prediksi penyakit dengan lebih akurat, namun belum diketahui metode apa yang paling akurat dalam memprediksi penyakit pasien. Berbagai macam teknik analisa yang secara konvensional dan manual yang selama ini digunakan tidak lagi begitu efektif digunakan untuk hal mendiagnosa suatu penyakit. Seiring dengan perkembangan teknologi dan system. berbasis pengetahuan terutama medis tuntutan akan adanya penggunaan sistem pengetahuan berbasis komputer sebagai teknik analisa dalam mendiagnosa penyakit menjadi semakin penting dan harus selalu dikembangkan.

Oleh karenanya, saat inilah waktu yang tepat untuk mengembangkan sistem pengetahuan berbasis komputer yang modern, efektif dan efisien dalam mendiagnosa masalah penyakit[5].

Oleh karena itu, penelitian ini dilakukan untuk membantu menyelesaikan permasalahan tersebut dengan data mining klasifikasi untuk prediksi penyakit jantung, diperlukan suatu metode atau teknik yang dapat mengolah data-data yang sudah ada. Salah satu metodenya menggunakan teknik data mining klasifikasi. Penggunaan data mining Algoritma C4.45, Naive Bayes dan F = Measure Nilai Tengah Reccal Aprecici dalam Klasifikasi Penderita Gagal Jantung sebagai pilihan untuk diagnosa penyakit jantung dapat menjadi alternatif pilihan yang tepat, tetapi sampai saat ini belum diketahui algoritma klasifikasi yang paling akurat dalam prediksi penyakit jantung.

Pada penelitian sebelumnya membandingkan algoritma klasifikasi data mining Naive Bayes Berbasis PSO untuk deteksi penyakit jantung. Pengukuran dengan Naives Bayes menghasilkan akurasi 82.14%, sementara dengan Naive Bayes Berbasis Particle Swarm Optimization akurasi meningkat menjadi 92.86%. [6]

Pada penelitian algoritma menggabungkan k-NN dengan algoritma genetika untuk klasifikasi yang lebih efektif. Algoritma genetika melakukan proses yang kompleks dan berikan solusi optimal. Hasil percobaan menunjukkan bahwa algoritma kami Meningkatkan akurasi diagnosis penyakit jantung[7].

Berdasarkan penelitian tersebut penulis melakukan penelitian dengan melakukan komparasi data mining dengan metode Komparasi Algoritma C4.5 dan Naive Bayes dalam Klasifikasi Penderita Penyakit gagal Jantung untuk mengetahui algoritma yang memiliki akurasi yang lebih tinggi dalam hal ini yaitu prediksi penyakit jantung. Karena kedua algoritma perlunya penanganan missing attribute value dengan sequential methods atau yang biasa disebut preprocessing karena pada sequential methods di dalamnya terdapat beberapa teknik penanganan missing value, sehingga teknik-teknik tersebut bisa dikomparasi dan diterapkankan pada algoritma C4.5 dan Naive Bayes sehingga pada akhirnya akan

*due to more intense and significant pressure because it goes beyond just changing the state of things.*

yang sebenarnya juga tetap memperhitungkan keterkaitan dengan atribut yang lain dan not be able to obtain critical information from the existing data set. Sehingga penulis mengambil judul “**Komparasi Algoritma C4.5 dan Naive Bayes Menggunakan Model Forward Selection dalam Prediksi Penyakit gagal Jantung**”.

## **1.2 Identifikasi Masalah**

Identifikasi Masalah Berdasarkan latar belakang tersebut, maka dapat diidentifikasi beberapa masalah diantaranya Diagnosa penyakit gagal jantung oleh dokter atau tim medis masih harus diperkuat dengan adanya metode klasifikasi data mining, Terdapat beberapa metode klasifikasi data mining yang mempunyai keakuratan yang berbeda untuk prediksi penyakit gagal jantung, Belum diketahui diantara metode klasifikasi untuk prediksi penyakit gagal jantung yang memiliki keakuratan yang terbaik.

## **1.3 Rumusan Masalah**

Berdasarkan identifikasi masalah yang telah diuraikan diatas, maka penulis merumuskan masalahnya sebagai berikut: Algoritma manakah yang lebih baik diantara Algoritma C4.5 dan Naive Bayes dalam prediksi gagal penyakit jantung

## **1.4 Batasan Penelitian**

Penelitian ini bertujuan untuk mengkomparasi beberapa algoritma klasifikasi dalam data mining dan memilih algoritma yang paling akurat untuk diterapkan pada proses prediksi penyakit gagal jantung.

## **1.5 Manfaat Penelitian**

Manfaat yang ingin dicapai dalam penelitian ini adalah:

- a) Diharapkan dapat memberikan kontribusi khususnya dalam bidang Kesehatan yaitu deteksi penyakit Gagal Jantung dengan membandingkan dua metode terkait.
- b) Dapat memberikan data bantu sebelum dilakukan pemeriksaan lebih lanjut terkait deteksi Gagal Jantung.
- c) Dapat menjadi rujukan atau penelitian selanjutnya yang dapat dikembangkan.

## **1.6 Ruang Lingkup Penelitian**

Pada penelitian ini akan membahas mengenai prediksi penyakit jantung dengan menggunakan metode klasifikasi data mining kemudian mengevaluasi hasil komparasi algoritma tersebut untuk mengetahui metode klasifikasi data mining yang paling akurat.

## **1.7 Sistematika Penulisan**

Untuk mempermudah penyusunan laporan penelitian ini, maka sistematika penulisan terbagi atas lima bab sebagai berikut:

Bab I Pendahuluan adalah membahas mengenai latar belakang, masalah penelitian yang terdiri dari identifikasi masalah, pembatasan masalah dan perumusan masalah, tujuan dan manfaat penelitian, serta sistematika penulisan.

Bab II Tinjauan Pustaka pada bagian ini berisi penelitian yang pernah dilakukan terkait topik penelitian; dan teori-teori pendukung yang digunakan yang berhubungan erat dengan pokok-pokok landasan berpikir, yang relevan dengan topik penelitian.

Bab III Metode Penelitian pada bagian ini berisi metode penelitian menggunakan Metode Algoritma C4.5 dan Naive Bayes serta tahapan yang

dilakukan dalam penelitian antara lain: langkah- Langkah penelitian dan pengumpulan data.

Bab IV Hasil dan Pembahasan Pada bab ini diuraikan mengenai garis besar tentang penelitian dari deskripsi data, hasil.

Bab V Simpulan dan Saran Pada bab ini diuraikan mengenai kesimpulan dari hasil penelitian. Isi dari bab ini meliputi kesimpulan dan saran.