

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG MASALAH

Serangan jantung merupakan salah satu penyakit paling mematikan di dunia[1]. Penyakit ini adalah salah satu penyakit yang paling sering terjadi kasusnya pada kalangan masyarakat, dimana penyakit jantung ini dapat menimpa dan menyerang siapapun tanpa memandang usia, jenis kelamin dan gaya hidup[2]. Serangan jantung dapat terjadi di mana saja dan kapan saja. Seringkali masyarakat terlambat dalam melakukan penanganan awal, mencari informasi dan pertolongan. Karena pada dasarnya serangan jantung dapat menyerang secara tiba-tiba dan dapat mengancam nyawa[3].

Menurut *World Health Organization*, penyakit kardiovaskuler merupakan penyebab utama kematian. Diperkirakan pada tahun 2015, lebih dari 80% kematian akibat penyakit kardiovaskuler dialami oleh negara dengan pendapatan perkapita rendah dan menengah, dan terjadi baik pada laki-laki maupun perempuan[4].

Berdasarkan data kejadian penyakit di Indonesia dari GLOBOCAN (IARC) tahun 2012, terdapat 43,30 kasus baru infark miokard dan 12,90% kematian. Jumlah penderita penyakit jantung pada tahun 2013 sebanyak 61.682 orang, dengan kejadian tertinggi per provinsi yaitu Jawa Tengah sebesar 11.511 orang. Beberapa faktor yang dapat menyebabkan penyakit jantung umumnya adalah kebiasaan hidup tidak sehat, merokok, begadang dan pola makan yang tidak baik[5].

Melihat banyaknya jumlah penderita penyakit serangan jantung, maka perlu dilakukan beberapa upaya penanganan untuk dapat mencegah penyakit serangan jantung, yakni dengan cara melakukan komparasi dalam memprediksi penyakit serangan jantung. Prediksi ini perlu dilakukan karena sering kali keputusan klinis atau ketidakakuratan yang dibuat oleh petugas medis dalam menangani pasien-pasien penyakit serangan jantung. Hal ini bisa dihindari dengan memanfaatkan data-data untuk dibuat suatu pola penentuan penyakit serangan jantung dengan teknik komputerisasi cerdas sehingga ketidaktepatan diagnosis dapat dihindari[6]. Beberapa metode yang dapat digunakan adalah data mining.

Beberapa penelitian yang menggunakan data mining untuk memprediksi penyakit serangan jantung diantaranya: penelitian yang dilakukan oleh Hendri Mahmud, Nawawi, dkk dengan judul komparasi algoritma *neural network* dan *naïve bayes* untuk memprediksi penyakit jantung akurasi 84,52% menggunakan algoritma *Neural Network backpropagation* dan 79,88%[5]. Pareja Alam Jusia menganalisis komparasi pemodelan algoritma *decision tree* menggunakan metode *particle swarm optimization* dan metode *adaboost* untuk prediksi awal penyakit jantung mendapatkan nilai algoritma *Decision Tree* nilai tingkat akurasi 79.26% dan nilai AUC 0.889. Setelah dilakukan modifikasi algoritma *decision tree* dengan metode *particle swarm optimization* nilai tingkat akurasi 82.59% dan untuk nilai AUC 0.916. Modifikasi algoritma *decision tree* dengan metode *adabost* nilai akurasi yang dihasilkan sebesar 79.26% dan nilai AUC sebesar 0.955[7]. Bakhtiar Rifai memprediksi algoritma *neural network* untuk penyakit jantung Model yang dihasilkan diuji untuk mendapatkan nilai *accuracy*, *precision*, *recall* dan AUC dari

algoritma sehingga didapat pengujian dengan menggunakan *neural network* didapat nilai *accuracy* adalah 91.45 % dengan nilai *precision* 92.79 % [8]. Srinivas, dkk melakukan deteksi penyakit jantung dengan menggunakan algoritma DT, *Multi- Layer Perceptron (MLP)*, *Baysian Model*, dan *Support Vevtor Machine (SVM)* dan parameter yang digunakan adalah *age, sex, chest pain type, fasting blood sugar, restecg, exercise, slope, CA, trest blood pressure, cholestrol, thalach, dan old peak*, dengan hasil akurasi DT 82%, MLP 89,2%, *Baysian Model* 82%, dan SVM 83,5% [9].

Berdasarkan dari beberapa penelitian data mining dalam memprediksi penyakit serangan jantung, penelitian itu menggunakan metode dan dataset yang berbeda-beda dengan tingkat akurasi yang berbeda. Maka dari itu penulis akan melakukan komparasi dengan menggunakan dataset yang sama untuk memperoleh tingkat akurasi berbeda. Metode yang digunakan yakni *naïve bayes* dan *adaboost*. Penggunaan metode *naïve bayes* dan *adaboost* dapat menjadi tolak ukur yang tepat, apalagi sampai saat ini belum diketahui secara pasti metode apa yang paling akurat dalam memprediksi penyakit serangan jantung. *Naïve bayes* sendiri memiliki kelebihan yakni dapat memprediksi probabilitas dalam suatu *class*, sedangkan kelebihan dari metode *Adaboost* adalah dapat meningkatkan optimasi yang dapat digabungkan dengan metode *naïve bayes* sehingga dapat meningkatkan keakurasian dari data yang diolah [10].

Naïve Bayes merupakan suatu pendekatan yang cukup sederhana dan baik dalam melakukan pelatihan data untuk klasifikasi. Sedangkan *Naïve Bayes Classifier* juga dapat dikatakan menghitung kemungkinan kelas atau kategori

klasifikasi data atau dapat dijelaskan seperti atribut kelas atau kategori yang diberi label[1]. *Naïve bayes* juga dapat di definisikan sebagai pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan asal inggris yakni Thomas Bayes[5].

Metode *Adaptive Boosting* atau *Adaboost* adalah variasi dari metode *boosting* yang memiliki hasil prediksi akurat dan teori dasar yang kuat. Selain itu metode *adaboost* dapat dilatih secara efektif Secara umum, metode *adaboost* memang lebih unggul dari metode *bagging*, metode ini dapat menetapkan bobot yang berbeda untuk data latih pada setiap iterasi dengan menambahkan dan mengurangi bobot pada pengklasifikasi sehingga dapat mengubah distribusi dari data latih secara efektif[11]. Klasifikasi menggunakan metode *Adaboost* mulai dari pengukuran F. dan area di bawah kurva karakteristik operasi penerima (AUC) digunakan untuk mengevaluasi kinerja metode pembelajaran mesin[10].

Berdasarkan penelitian tersebut penulis dapat melakukan penelitian dengan melakukan komparasi menggunakan metode *naïve bayes* dan *adaboost* untuk mengetahui metode mana yang lebih akurat dalam memprediksi penyakit serangan jantung terhadap dataset *Heart Analysis & Prediction Dataset* yang di peroleh dari *kaggle.com*. Sehingga penulis mengambil judul “**KOMPARASI METODE NAÏVE BAYES DAN ADABOOST UNTUK MEMPREDIKSI PENYAKIT SERANGAN JANTUNG**”.

1.2 RUMUSAN MASALAH

Berdasarkan uraian diatas, maka dapat dirumuskan masalah yang akan dikaji dalam penelitian ini adalah bagaimana perbandingan performa dari metode *naïve bayes* dan *adaboost* dalam memprediksi penyakit serangan jantung?

1.3 BATASAN MASALAH

Beberapa masalah yang dijadikan fokus dalam penelitian ini antara lain:

1. Penelitian hanya terfokus pada penyakit jantung yang umum dimasyarakat.
2. Algoritma *naïve bayes* dan *adaboost* untuk memprediksi menggunakan *weka*.
3. Input berupa nilai kriteria dari hasil pemeriksaan pasien penyakit jantung.
4. Output berupa hasil, apakah pasien menderita penyakit jantung atau tidak.
5. Dataset yang digunakan *Heart Analysis & Prediction Dataset* yang di peroleh dari *kaggle.com*.
6. Atribut yang digunakan adalah usia, jenis kelamin, nyeri dada, tekanan darah, kolesterol, gula darah, hasil elektrokardiogram, denyut per menit, angina, depresi, kemiringan segmen, pembuluh darah, kondisi jantung, dan status.
7. Alat bantu dalam memprediksi penyakit serangan jantung menggunakan *tools WEKA* versi 3.8.6.

1.4 TUJUAN PENELITIAN

Adapun tujuan dari penulis melakukan penelitian ini yaitu sebagai berikut:

1. Membandingkan performa metode *naïve bayes* dan *adaboost* dalam memprediksi penyakit serangan jantung.
2. Mengetahui tingkat prediksi penyakit serangan jantung dengan menggunakan metode *naïve bayes* dan *adaboost*.

1.5 MANFAAT PENELITIAN

Manfaat yang diperoleh dari penelitian ini adalah:

1. Membantu masyarakat untuk dapat mengetahui gejala dini penyakit jantung.
2. Memberikan hasil prediksi penyakit serangan jantung yang di derita oleh pasien.

1.6 SISTEMATIKA PENULISAN

BAB I : PENDAHULUAN

Pada bab ini memuat tentang latar belakang dari masalah penelitian, perumusan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

BAB II : LANDASAN TEORI

Dalam bab ini berisi tentang teori-teori dan pendapat para ahli sebagai penunjang yang digunakan oleh penulis yang dikutip dari jurnal, buku, dan lain-lain yang berhubungan dengan permasalahan yang akan di analisis sebagai pedoman penelitian.

BAB III : METODOLOGI PENELITIAN

Pembahasan bab ini menjelaskan tentang perangkat kerja penelitian, metode yang akan di gunakan, serta alat bantu yang digunakan dalam penelitian ini.

BAB IV : ANALISIS DAN HASIL

Pada bab ini dilakukan perhitungan dan visualisasi dari tools WEKA untuk menjelaskan hasil yang didapat dengan menggunakan metode *naïve bayes* dan *adaboost*.

BAB V : PENUTUP

Pada bab ini berisi tentang penutup dari penelitian ilmiah ini yang berisi dari kesimpulan dari pembahasan pada bab-bab sebelumnya dan saran-saran yang dapat berguna pada semua pihak yang terlibat dalam penelitian ini.