

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Kesehatan jantung merupakan faktor penting dalam kehidupan manusia. Sebagai organ vital, jantung adalah bagian penting dari tubuh yang memompa darah ke seluruh tubuh dan memastikan bahwa oksigen dan nutrisi sampai ke semua sel. Jika aliran oksigen terganggu, kesehatan jantung dapat terpengaruh dan berdampak pada kualitas hidup seseorang. Berdasarkan data dari Survei Kesehatan Indonesia tahun 2023, diperoleh akumulasi penderita sejumlah 877.531 kasus penyakit jantung berdasarkan diagnosis dokter di Indonesia yang ditunjukkan dengan rata-rata nilai 0,85% (Kesehatan, 2023).

Salah satu jenis gangguan jantung yang dapat menyebabkan penyempitan serta penyumbatan pembuluh darah koroner adalah Penyakit Jantung Koroner (PJK). Kondisi ini dapat menghambat proses transportasi energi dalam tubuh, sehingga menyebabkan ketidakseimbangan suplai oksigen. PJK tergolong penyakit kardiovaskular yang terjadi akibat penyumbatan arteri koroner akibat akumulasi plak, polutan, atau zat kimia dari lingkungan yang masuk melalui makanan, minuman, atau udara, kemudian menempel pada dinding arteri koroner (Erdania et al., 2023).

Mendiagnosis penyakit jantung koroner merupakan keputusan diagnosa yang cukup mahal. *Machine Learning* menjadi salah satu pilihan yang digunakan dalam diagnosa medis. Mengungkap suatu informasi dapat dengan memproses

analisa data yang sudah dikumpulkan melalui basis data atau bisa disebut *data mining*. Proses *data mining* menerapkan hubungan yang bermanfaat yang ada di antara data dan aturan ini dapat diterapkan untuk mengambil suatu keputusan yang tepat. Penerapan metode *data mining* khususnya klasifikasi, merupakan salah satu bentuk analisis data yang dapat membantu dalam menetapkan kelas label sampel yang akan diklasifikasi. Tujuannya adalah untuk memperkuat keandalan hasil yang dihasilkan dari kumpulan data (Hendrian, 2018).

Naive Bayes adalah metode klasifikasi berbasis probabilitas yang menghitung kemungkinan berdasarkan frekuensi dan kombinasi nilai dalam *dataset*. Ini adalah salah satu metode klasifikasi yang paling umum digunakan. Namun, Naïve Bayes hanya mendukung data dengan atribut diskrit dan memiliki keterbatasan dalam menangani atribut kontinu (Mukaromah & Wasilah, 2024). Selain itu, algoritma ini dikenal memiliki kecepatan dan akurasi tinggi, terutama ketika digunakan untuk data yang sangat besar (Riany & Testiana, 2023).

*Bagging* adalah salah satu metode dalam *ensemble learning* yang beroperasi dengan cara memodifikasi data pelatihan. Metode ini menduplikasi data pelatihan sebanyak  $n$  kali menggunakan teknik *sampling with replacement*, sehingga menghasilkan  $n$  kumpulan data pelatihan baru. Dari kumpulan data tersebut akan dibangun sejumlah *classifier* yang disebut sebagai *bagged classifier* (Ridwan, 2020).

Terdapat penelitian terdahulu memprediksi penyakit stroke yang menggunakan algoritma C 4.5 menerapkan metode *bagging* dan *adaboost* dalam meningkatkan performa klasifikasi. Hasil penelitian penggunaan algoritma C 4.5

mendapatkan nilai akurasi sebesar 92,87, sedangkan penggunaan metode *bagging* adalah 95,02%, dan penggunaan metode *adaboost* adalah 94,63%. Dapat disimpulkan penggunaan *bagging* dan *adaboost* masing-masing mengalami peningkatan yang signifikan sejumlah 3% dan 2% (Saputri et al., 2022). Pada penelitian lain menggunakan metode *bagging* pada *stunting* balita yang menggunakan penggabungan algoritma K-Nearest Neighbor, Support Vector Machine, dan Naive Bayes menggunakan *Bagging* diperoleh akurasi senilai 89,77% (Masacgi & Rohman, 2023).

Oleh karena itu dilakukan penelitian yang berjudul “Penerapan Bagging Ensemble Learning Pada Algoritma Naïve Bayes Untuk Memprediksi Penyakit Jantung Koroner”. Penelitian ini bertujuan untuk mengevaluasi hasil klasifikasi penyakit jantung koroner dengan menerapkan algoritma Naïve Bayes serta kombinasi Naïve Bayes dengan teknik *bagging* dalam *ensemble learning*.

## 1.2. Identifikasi Masalah

Dari latar belakang permasalahan yang diuraikan, maka dapat diidentifikasi masalah yang akan diteliti. Berdasarkan beberapa penelitian sebelumnya, algoritma *bagging ensemble learning* menghasilkan nilai akurasi yang meningkat pada penggunaan satu metode klasifikasi maupun lebih dalam klasifikasi penyakit terkait kesehatan. Maka penulis ingin melakukan penelitian untuk membandingkan penggunaan algoritma Naïve Bayes dan kombinasi *bagging ensemble learning* secara spesifik untuk prediksi penyakit jantung koroner.

### 1.3. Batasan Masalah Penelitian

Batasan yang ditetapkan pada penelitian ini antara lain:

1. Penerapan algoritma *bagging ensemble learning* pada klasifikasi algoritma Naïve Bayes dilakukan untuk mengetahui nilai komparasi kinerja dari metode tersebut.
2. *Dataset* yang digunakan berdasar pada gejala penyakit jantung koroner yang disesuaikan dengan sumber data *Kaggle* dengan judul *dataset “Coronary Heart Disease”*.
3. Parameter yang digunakan adalah *sbp, tobacco, ldl, adiposity, famhist, typea, obesity, alcohol, age*.
4. Evaluasi menggunakan *Cross Validation* dengan membandingkan nilai *accuracy, recall, precision*, serta *f1-score* dari algoritma Naïve Bayes dan kombinasi algoritma Naïve Bayes dengan *bagging ensemble learning*.
5. Penelitian ini mencari hasil akurasi algoritma *bagging ensemble learning* pada Naïve Bayes untuk kasus klasifikasi penyakit jantung koroner.

### 1.4. Rumusan Masalah Penelitian

Berdasarkan latar belakang, rumusan masalah yang bisa diidentifikasi antara lain:

1. Bagaimana penerapan *bagging ensemble learning* pada klasifikasi algoritma Naïve Bayes pada klasifikasi penyakit jantung koroner?

2. Bagaimana perbandingan klasifikasi algoritma Naïve Bayes dan kombinasi algoritma Naïve Bayes dengan *bagging ensemble learning* pada klasifikasi penyakit jantung koroner?

### 1.5. Tujuan Penelitian

Berdasarkan rumusan masalah, tujuan yang ditetapkan pada penelitian ini antara lain:

1. Mengetahui penerapan *bagging ensemble learning* pada klasifikasi algoritma Naïve Bayes pada klasifikasi penyakit jantung koroner
2. Mengetahui perbandingan klasifikasi algoritma Naïve Bayes dan kombinasi algoritma Naïve Bayes dengan *bagging ensemble learning* pada klasifikasi penyakit jantung koroner

### 1.6. Manfaat Penelitian

Beberapa manfaat yang didapatkan dari penelitian ini diantara lainnya adalah teoritis serta praktis:

#### 1.6.1 Manfaat Teoritis

Penelitian ini meliputi pengembangan pengetahuan pada penelitian dalam ilmu komputer untuk *data mining*, prediksi penyakit, klasifikasi Naïve Bayes, dan *Bagging ensemble learning* dalam memprediksi penyakit jantung koroner.

### 1.6.2 Manfaat Praktis

a. Bagi peneliti

Mampu menyumbangkan dan mengembangkan wawasan atau pemahaman dalam penerapan *data mining* pada prediksi suatu penyakit. Sehingga dapat menjadi acuan terhadap pengembangan penelitian sejenis ke depan.

b. Bagi masyarakat

Mampu memberikan karya penelitian yang dapat mendorong kemajuan penelitian yang berhubungan dengan *data mining* dan prediksi suatu penyakit.

