

# BAB 1

## PENDAHULUAN

### 1.1. Latar Belakang

Gagal jantung adalah kondisi medis serius di mana jantung tidak dapat memompa darah dengan cukup efisien untuk memenuhi kebutuhan tubuh. Kondisi ini sering berkembang secara bertahap dan bisa disebabkan oleh berbagai faktor seperti penyakit jantung koroner, tekanan darah tinggi, dan diabetes. Menurut Organisasi Kesehatan Dunia (WHO), lebih dari 17,7 juta orang diperkirakan meninggal karena penyakit jantung, yang menyebabkan 31% dari seluruh kematian secara global. Dengan kata lain bahwa penyakit kardiovaskular khususnya penyakit jantung adalah salah satu penyakit paling mematikan baik di negara maju maupun berkembang, perhatian terhadap penyakit tersebut sangatlah penting dan sangat diperlukan (Alizadehsani dkk., 2019)

Terlepas dari fakta bahwa angka kematian akibat penyakit jantung tinggi, peluang untuk bertahan hidup lebih tinggi jika diagnosis dilakukan sejak awal. Oleh karena itu, peneliti membuat model klasifikasi untuk menemukan pasien yang memiliki tingkat risiko yang tinggi. Pendekatan pembelajaran mesin (ML) menjadi lebih populer baru-baru ini untuk membangun model untuk diagnosis awal penyakit jantung.

Klasifikasi penyakit jantung menggunakan algoritma Naive Bayes adalah salah satu upaya memanfaatkan teknologi pembelajaran mesin dalam bidang medis. Naive Bayes sebagai salah satu algoritma pembelajaran mesin yang populer, menawarkan keunggulan dalam hal kesederhanaan dan kecepatan. Tujuan dari model klasifikasi ini untuk membantu dokter dan tenaga medis dalam menentukan seorang pasien berisiko mengalami penyakit jantung.

Naive Bayes adalah salah satu algoritma pembelajaran mesin yang populer dan efektif untuk diagnosis penyakit jantung. Meskipun memiliki beberapa keterbatasan, kelebihanannya dalam hal kesederhanaan dan kecepatan menjadikannya pilihan yang baik untuk aplikasi medis di mana keputusan cepat sangat penting. Dengan pengumpulan data yang tepat dan pra-pemrosesan yang baik, Naive Bayes dapat membantu dalam memberikan diagnosis awal yang akurat dan menyelamatkan banyak nyawa.

Pada penelitian terdahulu komparasi algoritma knn dan naïve bayes untuk klasifikasi diagnosis penyakit diabetes melitus, dengan menggunakan lima pembagian data nilai akurasi dari Naive Bayes sebesar 80%, sedangkan KNN memperoleh akurasi sebesar 75%.(Putry, 2022).

Pada penelitian terdahulu komparasi algoritma Naive bayes dan svm untuk memklasifikasi keberhasilan imunoterapi pada penyakit kutil. Dataset yang digunakan bersumber dari situs UCI Machine learning Repository yaitu Immunotherapy Dataset terdiri dari 90 record dan 7 variabel. Memperoleh hasil akurasi 1 untuk Naive bayes dan 0,8 untuk Support Vector Machine (SVM) (Supriyatna & Mustika, 2018).

Perbedaan penelitian terdahulu dengan penelitian ini terdapat pada objek penelitian yang mengangkat penyakit gagal jantung menurut organisasi kesehatan dunia (who), lebih dari 17,7 juta orang diperkirakan meninggal karena penyakit jantung, yang menyebabkan 31% dari seluruh kematian secara global, dengan dataset yang berjumlah 12 variabel yang mencakup faktor resiko penyakit gagal jantung. Algoritma yang digunakan dalam penelitian ini adalah Naive Bayes dengan fokus pada klasifikasi penyakit jantung.

Menggunakan dataset dengan 12 variabel utama yang mencakup berbagai faktor risiko penyakit jantung seperti usia, jenis kelamin, tipe nyeri dada, tekanan darah istirahat, kolesterol, gula darah puasa, EKG istirahat, detak jantung maksimum, angina yang disebabkan oleh olahraga, oldpeak, kemiringan ST, dan lain-lain. Algoritma yang digunakan dalam penelitian ini adalah Naive Bayes dengan fokus pada klasifikasi penyakit jantung.

Penelitian ini bertujuan untuk mengembangkan model klasifikasi gagal jantung yang menggunakan algoritma Naive Bayes. Tujuan utama dari penelitian ini adalah mengidentifikasi faktorresiko penyakit gagal jantung, mengembangkan model klasifikasi berbasis algoritma Naive Bayes, penelitian ini diharapkan dapat memberikan kontribusi untuk mengurangi angka kematian akibat penyakit jantung.

## **1.2. Rumusan Masalah**

1. Bagaimana mengembangkan model klasifikasi risiko gagal jantung menggunakan algoritma Naive Bayes dengan bantuan Scikit-learn?

2. Bagaimana evaluasi performa model klasifikasi gagal jantung dalam hal akurasi, presisi, recall, dan F1-score berdasarkan dataset yang tersedia?

### **1.3. Tujuan Penelitian**

1. Mengembangkan model klasifikasi risiko gagal jantung yang akurat menggunakan algoritma Naive Bayes dengan bantuan Scikit learn.
2. Mengevaluasi performa model klasifikasi gagal jantung berdasarkan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score.

### **1.4. Manfaat Penelitian**

1. Model klasifikasi yang dikembangkan dapat membantu dalam diagnosis dini penyakit gagal jantung, sehingga memungkinkan intervensi dan perawatan yang lebih tepat waktu.
2. Dengan adanya model klasifikasi, praktisi kesehatan dapat mengelola pasien dengan lebih efektif, termasuk merencanakan perawatan yang sesuai dengan kondisi pasien secara individual.
3. Penelitian ini dapat membuka jalan untuk pengembangan teknologi kesehatan yang lebih canggih dan inovatif dalam mendukung diagnosis dan pengelolaan penyakit kardiovaskular.

### **1.5. Batasan Penelitian**

1. Penelitian ini akan menggunakan dataset yang relevan untuk klasifikasi penyakit gagal jantung. Jumlah data yang digunakan sebanyak 1190.
2. Data yang digunakan dalam penelitian ini akan mencakup rentang waktu tertentu, sesuai dengan ketersediaan data yang dapat diakses di kaggle <https://www.kaggle.com/datasets/sid321axn/heart-statlog-cleveland-hungary-final/data>.
3. Atribut data yang akan digunakan untuk klasifikasi penyakit gagal jantung antara lain : age, sex, chest pain type, resting bp s, cholesterol, fasting blood sugar, resting ecg, max heart rate, exercise angina, oldpeak, st slope.
4. Penelitian ini akan menggunakan Scikit Learn sebagai library machine learning utama untuk mengembangkan model klasifikasi penyakit gagal jantung. Selain itu, pengembangan website dan antarmuka pengguna akan menggunakan visual studio code, github dan streamlit share.

5. Atribut kelas yang akan digunakan untuk klasifikasi penyakit gagal jantung antara lain : resiko tinggi dan resiko rendah

