

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Penelitian

Penyakit jantung, khususnya gagal jantung (*heart failure*), merupakan salah satu penyebab utama kematian di dunia. Menurut data dari Organisasi Kesehatan Dunia (WHO), penyakit jantung terus menjadi penyebab utama kematian global, dengan estimasi sekitar 17 juta orang meninggal setiap tahunnya karena penyakit kardiovaskular. Salah satu tantangan terbesar dalam pengelolaan gagal jantung adalah diagnosis yang terlambat, yang sering kali disebabkan oleh gejala yang mirip dengan penyakit lain, serta kurangnya kesadaran akan gejala awal penyakit ini. [2]

Gagal jantung adalah kondisi medis di mana jantung tidak mampu memompa darah secara efisien ke seluruh tubuh. Diagnosis gagal jantung yang cepat dan akurat sangat penting untuk memulai pengobatan yang tepat guna mengurangi komplikasi lebih lanjut. Oleh karena itu, pendekatan berbasis data dan analitik yang dapat membantu dalam mengidentifikasi pasien berisiko tinggi mengalami gagal jantung menjadi sangat relevan. [3]

Dengan kemajuan teknologi, penggunaan data klinis dalam bidang medis semakin berkembang. Data medis yang beragam, seperti hasil pemeriksaan laboratorium, riwayat medis pasien, dan hasil pemeriksaan fisik, memberikan potensi besar untuk dianalisis menggunakan teknik-teknik kecerdasan buatan (AI). Salah satu teknik yang telah terbukti efektif dalam mengolah data medis dan memberikan prediksi yang akurat adalah *machine learning* (ML) [4]. Algoritma *machine learning* dapat mempelajari pola dari data historis untuk membuat prediksi tentang kemungkinan suatu kejadian di masa depan, seperti risiko gagal jantung.

Di antara berbagai algoritma *machine learning*, Random Forest adalah salah satu yang paling banyak digunakan dalam prediksi medis [1]. Random Forest adalah algoritma *ensemble learning* yang menggabungkan banyak pohon keputusan

(decision trees) untuk memberikan hasil yang lebih stabil dan akurat. Algoritma ini sangat cocok untuk menangani data medis yang kompleks dengan banyak fitur, karena mampu mengurangi overfitting dan memberikan interpretasi yang baik mengenai pentingnya fitur dalam model prediksi.

Penelitian ini bertujuan untuk mengimplementasikan algoritma Random Forest dalam memprediksi gagal jantung berdasarkan data klinis pasien. Dengan memanfaatkan data medis yang mencakup faktor-faktor seperti usia, jenis kelamin, tekanan darah, kadar kolesterol, dan hasil pemeriksaan lainnya, diharapkan algoritma ini dapat memberikan prediksi yang lebih akurat dan lebih cepat dibandingkan dengan metode tradisional. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi signifikan terhadap pengembangan sistem prediksi medis yang dapat digunakan untuk membantu dokter dalam memberikan diagnosis yang lebih tepat dan cepat, serta memberikan peringatan dini kepada pasien yang berisiko tinggi.

## 1.2 Rumusan Masalah Penelitian

Berdasarkan latar belakang yang telah dipaparkan, terdapat beberapa permasalahan yang akan dibahas dalam penelitian ini:

1. Bagaimana menerapkan algoritma Random Forest dan XGBoost dalam prediksi gagal jantung.
2. Bagaimana perbandingan kinerja algoritma Random Forest dan XGBoost dalam prediksi gagal jantung.

## 1.3 Batasan Masalah Penelitian

Agar penelitian ini lebih terfokus dan terarah, beberapa batasan masalah yang diterapkan adalah sebagai berikut:

1. Fitur atau variabel yang digunakan dalam model prediksi hanya terbatas pada data klinis yang tersedia dalam dataset.
2. Jumlah dataset sebanyak 919 data yang didapatkan dari Kaggle.

3. Variable yang digunakan adalah Usia, jenis kelamin, kolesterol, tekanan darah, dan kadar natrium.
4. Hasil prediksi berbentuk “ya” atau “tidak”

#### 1.4 Tujuan Penelitian

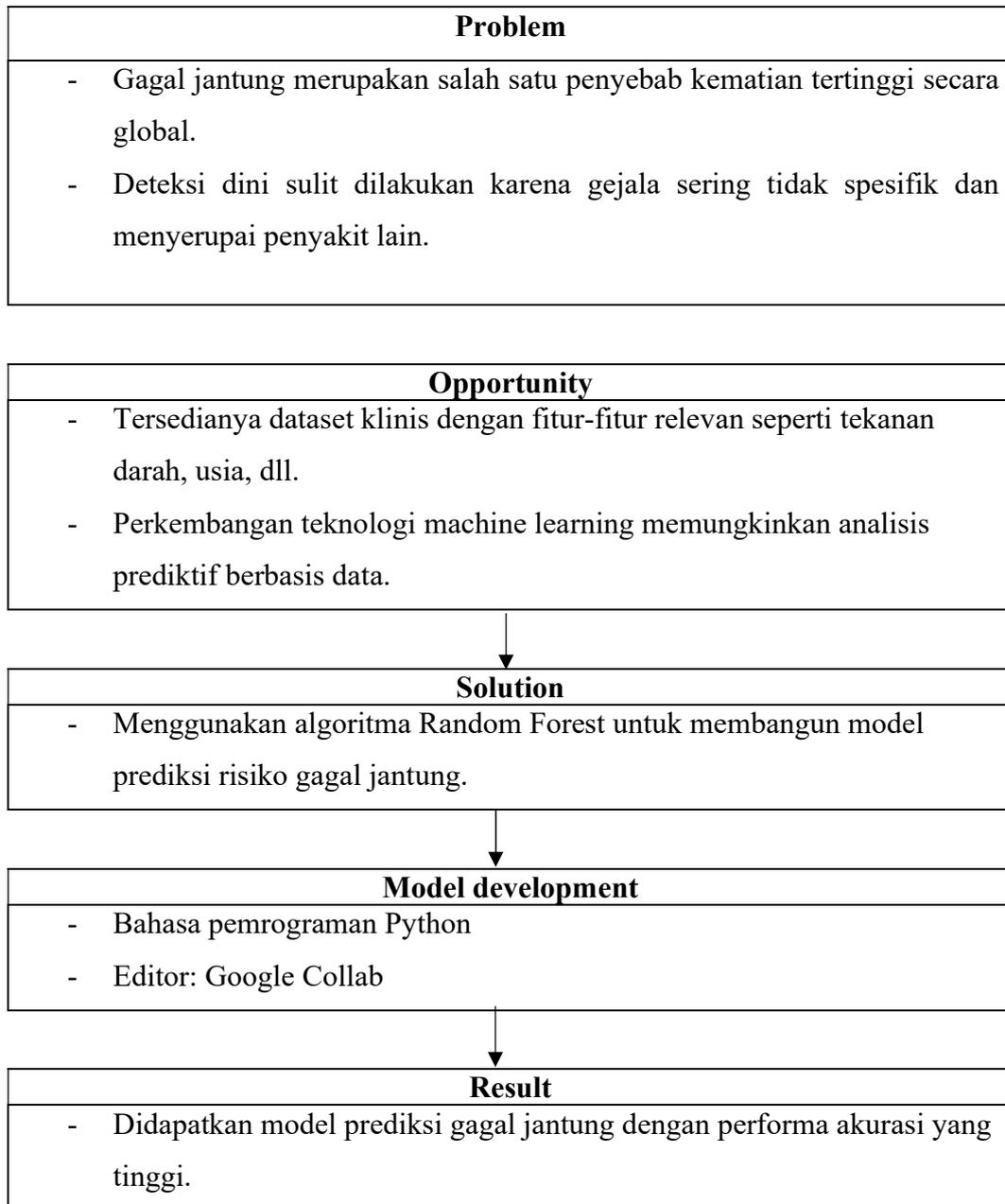
Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Menerapkan algoritma Random Forest dan XGBoost dalam model prediksi gagal jantung.
2. Mengetahui perbandingan kinerja Random Forest dan XGBoost pada model prediksi gagal jantung.

#### 1.5 Manfaat Penelitian

Hasil dari penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Pasien dengan prediksi risiko tinggi dapat diberikan peringatan dini, memungkinkan mereka untuk mendapatkan perhatian medis yang lebih baik dan mencegah komplikasi lebih lanjut.
2. Bagi Pengembang dapat memanfaatkan hasil penelitian ini untuk mengimplementasikan dan mengoptimalkan algoritma *machine learning*, khususnya Random Forest, dalam aplikasi medis lainnya. Penelitian ini juga memberikan wawasan mengenai bagaimana cara menangani data medis yang kompleks dan beragam.
3. Penelitian ini memberikan contoh nyata penerapan algoritma *machine learning*, khususnya Random Forest, dalam bidang medis. Hal ini membuka peluang bagi peneliti lain untuk mengembangkan penelitian serupa dalam bidang lain, seperti prediksi penyakit atau analisis data medis lainnya.



Gambar 1.1 Kerangka Pemikiran