

BAB I PENDAHULUAN

1.1 Latar Belakang

Berdasarkan pengamatan pada berbagai fasilitas kesehatan di Indonesia, ditemukan bahwa proses diagnosis penyakit jantung masih menghadapi kendala signifikan dalam hal kecepatan dan akurasi deteksi dini. Keterbatasan jumlah tenaga medis spesialis jantung, terutama di daerah-daerah yang jauh dari pusat kesehatan, menyebabkan tertundanya proses diagnosis yang berpotensi membahayakan nyawa pasien. Data dari *World Health Organization* menunjukkan bahwa penyakit jantung masih menjadi penyebab kematian tertinggi secara global, dengan 17,9 juta kematian setiap tahunnya [1].

Hasil observasi di beberapa rumah sakit menunjukkan bahwa proses *screening* awal pasien penyakit jantung masih dilakukan secara manual dan membutuhkan waktu yang relatif lama. Kondisi ini diperparah dengan meningkatnya jumlah pasien yang membutuhkan pemeriksaan jantung, sementara kapasitas pelayanan kesehatan tetap terbatas. Riset Kesehatan Dasar mencatat peningkatan prevalensi penyakit jantung dari 1,5% menjadi 1,8% dalam kurun waktu lima tahun [2].

Penerapan teknologi *machine learning*, khususnya algoritma *Random Forest*, telah menunjukkan potensi yang menjanjikan dalam bidang kesehatan [3]. Ketersediaan *University of California Irvine (UCI) Heart Disease Dataset* yang telah tervalidasi, dikombinasikan dengan kemampuan algoritma *Random Forest* dalam menganalisis data kompleks, membuka peluang untuk mengembangkan sistem prediksi yang dapat membantu proses *screening* awal [4]. Teknologi ini dapat menganalisis 13 parameter medis secara simultan untuk menghasilkan prediksi yang lebih cepat dan akurat.

Pengembangan sistem prediksi berbasis *machine learning* ini menjadi relevan mengingat tren peningkatan kasus penyakit jantung dan keterbatasan sumber daya medis yang ada. Pada penelitian [5] menunjukkan bahwa

implementasi sistem prediksi penyakit jantung berbasis *machine learning* dapat meningkatkan efisiensi diagnosis hingga 85%. Penelitian ini fokus pada penggunaan bahasa pemrograman Python yang memiliki berbagai *library machine learning* yang komprehensif untuk mendukung pengembangan model prediksi yang efektif [3].

1.2 Rumusan Masalah

Penelitian ini bertujuan untuk mengevaluasi tingkat akurasi algoritma *Random Forest* dalam memprediksi penyakit jantung menggunakan *UCI Heart Disease Dataset* dan menganalisis performa model prediksi yang dikembangkan dalam mendukung proses screening awal pasien penyakit jantung.

1.3 Batasan Masalah

Untuk memfokuskan penelitian ini, ditetapkan batasan-batasan masalah sebagai berikut:

1. Data yang digunakan adalah *UCI Heart Disease Dataset* dengan 1026 records dan 13 parameter medis.
2. Algoritma yang digunakan terbatas pada *Random Forest* dengan implementasi menggunakan bahasa pemrograman Python dan *library scikit-learn*.
3. Proses evaluasi model menggunakan *metrics accuracy, precision, recall, dan F1-score* dengan pembagian data *training-testing ratio 80:20*.
4. Penelitian tidak mencakup pengembangan antarmuka pengguna (*user interface*) dan implementasi sistem berbasis web atau *mobile*.

1.4 Tujuan Penelitian

Tujuan yang akan dicapai oleh peneliti dalam penelitiannya adalah:

1. Menghasilkan model prediksi penyakit jantung menggunakan algoritma *Random Forest* dengan tingkat akurasi yang dapat diukur.
2. Mengidentifikasi parameter medis yang memiliki pengaruh paling signifikan dalam prediksi penyakit jantung berdasarkan hasil analisis.

3. Menghasilkan prediksi yang dapat membantu proses screening awal pasien penyakit jantung dengan mengukur performa model melalui *metrics accuracy, precision, recall, dan F1-score*.

1.5 Manfaat Penelitian

Manfaat penelitian ini dapat ditinjau dari dua aspek sebagai berikut:

1. Memberikan kontribusi dalam pengembangan model prediksi penyakit jantung menggunakan algoritma *Random Forest*.
2. Menambah pengetahuan tentang pengaruh parameter medis terhadap prediksi penyakit jantung.
3. Menjadi referensi bagi peneliti selanjutnya dalam pengembangan sistem prediksi penyakit menggunakan *machine learning*.
4. Membantu tenaga medis melakukan screening awal pasien penyakit jantung dengan lebih efisien.
5. Mempercepat proses identifikasi pasien yang memerlukan pemeriksaan lebih lanjut.
6. Mengoptimalkan penggunaan sumber daya medis, terutama di daerah dengan keterbatasan akses ke spesialis jantung.
7. Meningkatkan kesempatan deteksi dini penyakit jantung sehingga pasien dapat ditangani lebih cepat.

1.6 Sistematika Penulisan

Penulisan skripsi ini disusun dalam lima bab dengan sistematika sebagai berikut:

BAB I PENDAHULUAN

Berisi latar belakang masalah yang mendasari penelitian tentang prediksi penyakit jantung, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Berisi landasan teori tentang penyakit jantung, *machine learning*, algoritma

Random Forest, UCI Heart Disease Dataset, penjelasan tentang *metrics* evaluasi model, dan kajian penelitian terdahulu yang relevan dengan topik penelitian.

BAB III METODE PENELITIAN

Membahas tentang tahapan penelitian yang dilakukan, meliputi pengumpulan data, preprocessing data, implementasi algoritma *Random Forest* menggunakan Python, pengembangan model prediksi, dan metode evaluasi yang digunakan untuk mengukur performa model.

BAB IV HASIL DAN PEMBAHASAN

Menyajikan hasil implementasi sistem, analisis performa model prediksi, pengujian akurasi, analisis *feature importance*, interpretasi hasil prediksi, serta pembahasan mengenai efektivitas sistem dalam konteks screening awal penyakit jantung.

BAB V PENUTUP

Berisi kesimpulan dari hasil penelitian yang telah dilakukan, mencakup tingkat akurasi model, dan parameter medis yang paling berpengaruh serta saran untuk pengembangan penelitian selanjutnya.