

BAB I

PENDAHULUAN

1.1 Latar Belakang

Seiring kemajuan teknologi yang membawa perubahan besar dalam berbagai aspek kehidupan, masyarakat semakin banyak menghadapi tantangan kesehatan yang kompleks. Salah satu dampak yang mencolok adalah meningkatnya jumlah penderita berbagai jenis penyakit yang beragam. Hal ini sangat dipengaruhi oleh perubahan gaya hidup, seperti pola makan yang kurang sehat, minimnya aktivitas fisik, serta kebiasaan yang tidak mendukung kesehatan secara menyeluruh (Rahmadayanti et al., 2023). Faktor-faktor tersebut menjadi salah satu peningkatan resiko seseorang terjangkit penyakit jantung.

Penyakit jantung merupakan salah satu penyebab kematian utama di Indonesia (Saputra Hs et al., 2024). Gaya hidup tidak sehat menjadi faktor utama penyebab adanya risiko penyakit jantung yang menimbulkan (Scandea et al., 2023). Beberapa aspek yang digunakan dalam mendiagnosis penyakit jantung meliputi kadar kolesterol, tekanan darah tinggi, kurangnya aktivitas fisik, dan obesitas. Faktor yang tidak dapat diubah mencakup jenis kelamin, riwayat keluarga, dan usia. Sementara itu, faktor yang dapat diubah meliputi kebiasaan merokok (aktif maupun pasif), tekanan darah tinggi, kadar kolesterol yang tinggi, dan kurangnya olahraga. Selain itu, analisis dokter menjadi salah satu tantangan dalam menentukan tingkat risiko pasien terhadap serangan jantung, apakah risikonya lebih rendah atau lebih tinggi (Arif et al., 2024).

Pada penelitian Samosir et al., (2021) tentang penyakit jantung, dengan penelitian menggunakan algoritma *Naïve Bayes*, algoritma *K-Nearest Neighbor*, dan algoritma *Random Forest* menunjukkan *Naïve Bayes* memiliki akurasi terbaik (91%) dalam klasifikasi data. Wardhana et al., (2022) pernah meneliti kacang kering dengan menggunakan algoritma *Random Forest* (RF), *Gradient Boosting Machine* (GBM), *light GBM* serta model evaluasi *repeated k-folds*. Selanjutnya, (Purnomo et al., 2024) menggunakan algoritma *Random Forest* dan algoritma *SVM* untuk memprediksi banjir, menghasilkan akurasi tertinggi 99,6% pada algoritma *Random Forest*.



Penerapan algoritma *machine learning* menawarkan solusi yang menjanjikan untuk prediksi penyakit, termasuk penyakit jantung. Algoritma Random Forest telah mendapat hasil terbaik dalam penelitian Apriliah et al., (2021) dengan akurasi 97.88%, dimana algoritma ini akan menjadi algoritma pertama dalam penelitian ini. Sedangkan, algoritma kedua yang akan digunakan adalah Gradient Boosting, dimana pada penelitian Andryan & Fajri., (2022) mendapatkan hasil terbaik 95.12%. Kedua algoritma ini telah menunjukkan potensi yang tinggi dalam berbagai studi kasus, namun perbandingan kinerjanya dalam konteks prediksi penyakit jantung masih terbatas dan memerlukan penelitian lebih lanjut.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan maka terdapat hal yang perlu dikaji lebih lanjut, antara lain:

1. Bagaimana kinerja algoritma *Random Forest* dan *Gradient Boosting* dalam memprediksi penyakit jantung pada data multivariable.
2. Algoritma manakah yang memberikan hasil terbaik berdasarkan metrik evaluasi seperti akurasi, *precision*, *recall*, dan *F1-score*.
3. Apa saja *variabel* atau fitur yang paling berpengaruh dalam klasifikasi penyakit jantung menggunakan kedua algoritma tersebut.

1.3 Tujuan Penelitian

Adapun tujuan pada penelitian ini yaitu:

1. Mengetahui kinerja algoritma *Random Forest* dan *Gradient Boosting* dalam memprediksi penyakit jantung berdasarkan data *multivariabel*.
2. Mengetahui algoritma yang memberikan performa terbaik dalam prediksi penyakit jantung.
3. Menentukan fitur-fitur utama yang berkontribusi terhadap akurasi prediksi penyakit jantung.

1.4 Manfaat

Manfaat dari penelitian ini dibagi menjadi tiga bagian antara lain :

1. Bagi Dunia Kesehatan: Memberikan model prediksi yang akurat untuk membantu deteksi dini penyakit jantung, sehingga pengobatan dapat dilakukan lebih cepat dan efektif.

2. Bagi Peneliti: Menyediakan analisis komparatif antara algoritma *Random Forest* dan *Gradient Boosting*, sebagai referensi dalam studi *data mining* pada bidang kesehatan.
3. Bagi Masyarakat: Meningkatkan kesadaran akan pentingnya deteksi dini penyakit jantung melalui penerapan teknologi *machine learning*.

