

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Penyakit jantung koroner (PJK) adalah salah satu penyebab utama kematian secara global, dan dampaknya sangat terasa di Indonesia. Penyakit ini secara konsisten menyumbang angka kematian tertinggi, menjadikannya masalah kesehatan masyarakat yang serius dan mendesak untuk ditangani. Berdasarkan studi *Global Burden of Disease* (GBD), Kematian akibat penyakit jantung dan pembuluh darah di Indonesia mengalami peningkatan dari 292.000 kasus pada tahun 1990 menjadi lebih dari 659.000 kasus pada tahun 2019 [1]. Penyakit Jantung Koroner sebagai bentuk utama dari penyakit *kardiovaskular*, memiliki *prevalensi* yang tinggi dan menjadi tantangan serius bagi sistem pelayanan kesehatan nasional. Faktor risiko seperti *hipertensi*, *diabetes*, *dislipidemia*, dan gaya hidup sedentari turut memperburuk tren ini [2]. Selain itu, keterlambatan dalam deteksi dini dan pengobatan optimal turut menyebabkan angka mortalitas dan morbiditas tetap tinggi. Oleh karena itu, penting untuk mengembangkan solusi yang dapat meningkatkan efektivitas diagnosis dan intervensi terhadap Penyakit jantung koroner.

RSUD Dr. Soeselo Slawi sebagai rumah sakit rujukan regional menghadapi tantangan dalam melakukan diagnosis penyakit jantung koroner secara cepat dan akurat. Metode konvensional yang bergantung pada interpretasi klinis dan pengalaman dokter sering kali memerlukan waktu yang lama dan berisiko menimbulkan subjektivitas. Rumah sakit ini menangani banyak pasien dengan beragam gejala dan kondisi komorbid, sehingga membutuhkan alat bantu diagnosis yang dapat mempercepat proses pengambilan keputusan. Data rekam medis pasien yang tersimpan di RSUD memiliki potensi besar untuk dimanfaatkan dalam pengembangan sistem pendukung keputusan berbasis teknologi. Oleh karena itu, dibutuhkan pendekatan berbasis teknologi yang mampu mendukung proses diagnosis secara objektif dan efisien [3]. Selain itu, metode ini mampu mengolah data medis tabular secara efisien dan memberikan output prediktif yang berguna

bagi tenaga medis. Penggunaan *machine learning* juga dapat mengurangi kesalahan diagnosis dan memperkuat akuntabilitas keputusan klinis. Oleh karena itu, pemanfaatan algoritma yang tepat sangat penting dalam mempercepat dan meningkatkan ketepatan diagnosis penyakit jantung koroner.

Perkembangan teknologi *machine learning* memberikan peluang untuk meningkatkan akurasi diagnosis medis [4]. Algoritma machine learning dapat menganalisis pola kompleks dalam data medis yang sulit diidentifikasi secara manual, serta memberikan prediksi berbasis data historis pasien. Dalam konteks diagnosis penyakit jantung koroner, *Algoritma K-Nearest Neighbors (KNN)* serta *Random Forest* menjadi pilihan yang relevan karena telah terbukti efektif dalam klasifikasi penyakit berbasis data klinis [5]. Selain itu, metode ini mampu mengolah data medis tabular secara efisien dan memberikan output prediktif yang berguna bagi tenaga medis. Penggunaan *machine learning* juga dapat mengurangi kesalahan diagnosis dan memperkuat akuntabilitas keputusan klinis. Oleh karena itu, pemanfaatan algoritma yang tepat sangat penting dalam mempercepat dan meningkatkan ketepatan diagnosis penyakit jantung koroner.

Studi terbaru menunjukkan bahwa *Random Forest* sangat baik dalam mengelola data medis yang rumit menghasilkan akurasi tinggi dalam diagnosis penyakit [6]. Sementara itu, KNN dikenal sebagai algoritma yang sederhana namun cukup efektif dalam klasifikasi berbasis kemiripan antar data pasien [7]. Kedua algoritma ini telah banyak diterapkan dalam studi diagnosis penyakit, termasuk diabetes, hipertensi, dan gangguan kardiovaskular lainnya. Penelitian yang dilakukan Susi Eva Maria Purba bahwa *Random Forest* mencapai akurasi 100% dalam klasifikasi obat, sementara KNN hanya mencapai 70% [8]. Selain itu, menurut Shivahare et al., kedua algoritma efektif dalam diagnosis berbasis citra medis, namun *Random Forest* lebih tahan terhadap variasi data [9].

Oleh karena itu, penelitian ini bertujuan untuk membandingkan performa *K-Nearest Neighbors (KNN)* dan *Random Forest*, sebagai dua algoritma *machine learning*., dalam klasifikasi diagnosis penyakit jantung koroner. Kedua algoritma dipilih karena mampu menangani data medis tabular dengan banyak *variabel*, seperti yang umum ditemukan dalam rekam medis pasien. Berdasarkan studi

terkini, KNN dan *Random Forest* menunjukkan performa tinggi dalam klasifikasi penyakit. Berdasarkan keunggulan tersebut, karena algoritma KNN dan *Random Forest* dinilai sangat menjanjikan untuk aplikasi di sistem pendukung keputusan klinis, penelitian ini dirancang untuk mengidentifikasi algoritma mana yang memberikan hasil klasifikasi paling akurat.

1.2. Rumusan Masalah

Perumusan masalah yang akan dibahas dalam penelitian ini adalah sebagai berikut:

1. Algoritma manakah yang memiliki akurasi lebih tinggi antara KNN dan *Random Forest* untuk diagnosis penyakit jantung koroner?
2. Bagaimana perbandingan kinerja model machine learning berdasarkan variasi pembagian data latih dan data uji (60:40, 70:30, 80:20, dan 90:10) dalam menghasilkan akurasi dan metrik evaluasi yang optimal?

1.3. Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Penelitian ini difokuskan pada penggunaan dua algoritma machine learning, yaitu *K-Nearest Neighbors* (KNN) dan *Random Forest* dalam mendiagnosa penyakit jantung koroner.
2. Penilaian performa algoritma hanya berdasarkan metrik evaluasi klasifikasi, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*. Evaluasi lanjutan seperti ROC-AUC, SHAP, atau LIME tidak digunakan.
3. Data yang digunakan terbatas pada rekam medis pasien RSUD Dr. Soeselo Slawi.
4. Proses pengujian dilakukan dengan variasi rasio pembagian data latih dan data uji, yaitu 60:40, 70:30, 80:20, dan 90:10. Rasio di luar ini tidak dianalisis.
5. Implementasi dilakukan menggunakan bahasa pemrograman *Python* dan pustaka pendukung seperti *scikit-learn*, *pandas*, dan *matplotlib*, melalui platform *Jupyter Notebook*.

1.4. Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut:

1. Mengevaluasi dan membandingkan performa algoritma *K-Nearest Neighbors* (KNN) dan *Random Forest* dalam diagnosis penyakit jantung koroner berdasarkan akurasi dan metrik evaluasi lainnya (seperti *precision*, *recall*, dan *F1-score*) berdasarkan data pasien RSUD Dr. Soeselo.
2. Menganalisis pengaruh variasi proporsi data latih dan data uji (60:40, 70:30, 80:20, dan 90:10) terhadap performa model *machine learning*, guna menemukan konfigurasi pembagian data yang menghasilkan hasil evaluasi paling optimal.

1.5. Manfaat Penelitian

Berikut adalah manfaat yang diharapkan dari penelitian ini:

A. Manfaat bagi medis adalah sebagai berikut:

1. Membantu tenaga medis di RSUD Dr. Soeselo dalam proses diagnosis penyakit jantung koroner yang lebih akurat dan cepat.
2. Mengurangi kesalahan diagnosis dan meningkatkan kualitas layanan kesehatan di rumah sakit.
3. Meningkatkan efisiensi pelayanan kesehatan dengan sistem pendukung keputusan berbasis *machine learning*.

B. Manfaat bagi akademisi adalah sebagai berikut:

1. Berperan dalam kemajuan ilmu pengetahuan di bidang *machine learning* khususnya dalam aplikasi diagnosis medis.
2. Memperkaya literatur ilmiah terkait penerapan algoritma KNN dan *Random Forest* di sektor kesehatan.
3. Berfungsi sebagai rujukan untuk studi berikutnya yang berkaitan dengan diagnosis penyakit jantung koroner menggunakan pendekatan *machine learning*.

C. Manfaat bagi masyarakat adalah sebagai berikut:

1. Mendorong peningkatan kualitas layanan kesehatan berbasis data sehingga masyarakat mendapatkan layanan medis yang lebih cepat dan tepat.
2. Secara tidak langsung membantu menurunkan angka kematian akibat penyakit jantung koroner melalui deteksi dini yang lebih akurat.

D. Manfaat bagi pengembangan teknologi adalah sebagai berikut:

1. Memberikan dasar untuk pengembangan sistem informasi kesehatan yang lebih canggih, terintegrasi, dan berbasis kecerdasan buatan.
2. Menjadi pijakan awal bagi pemanfaatan algoritma machine learning di rumah sakit dalam rangka transformasi digital layanan medis.