

BAB I

PENDAHULUAN

1.1 Latar Belakang

Serangan jantung adalah penyakit manusia yang paling mematikan di seluruh dunia [1]. Penyakit jantung adalah suatu kondisi medis yang melibatkan gangguan pada jantung, organ sentral dalam sirkulasi darah. Penyakit jantung meliputi kondisi seperti penyumbatan pembuluh darah (aterosklerosis), gangguan irama jantung (aritmia), gagal jantung, dan penyakit katup jantung. Penyakit-penyakit ini dapat mengganggu aliran darah yang memadai ke seluruh tubuh, menyebabkan gejala-gejala seperti nyeri dada, sesak napas, kelelahan, dan bahkan serangan jantung. Menurut analisis data CDCWONDER, terdapat 141.137 kematian akibat penyakit jantung di dunia [2]. Perkembangan data medis membuka peluang baru bagi para dokter untuk mendeteksi penyakit jantung pada pasien sejak dini.

Penggunaan machine learning dalam dunia medis dapat membantu dalam pengolahan citra medis, data kesehatan elektronik, prediksi risiko penyakit jantung, dan pemantauan pasien. Algoritma machine learning dapat diterapkan untuk menganalisis data, melakukan prediksi dini, dan memungkinkan pengambilan keputusan yang tepat dalam mencegah penyakit jantung. Prediksi dini dapat meminimalkan keterlambatan penanganan penyakit jantung karena kurangnya pemahaman tentang gejala utama penyakit jantung. Dengan mengumpulkan data pasien yang positif terkena penyakit jantung berupa beberapa karakteristik dan kebiasaan, algoritma dapat membuat model prediksi untuk mengklasifikasikan pasien yang memiliki risiko penyakit jantung atau tidak. Kemampuan machine learning untuk memproses data medis dalam jumlah besar secara real time meningkatkan akurasi dan efisiensi diagnosa.

Mengintegrasikan algoritma machine learning di bidang medis memiliki potensi yang sangat besar untuk mencegah deteksi dini penyakit jantung, meningkatkan hasil pengobatan, mengurangi risiko komplikasi penyakit serius

lainnya, dan meningkatkan kualitas hidup. Seiring kemajuan teknologi, penyempurnaan dan pengembangan lebih lanjut dalam algoritma ini mengarah pada sistem yang belajar dari masa lalu, terus meningkatkan akurasi dan efektivitas di masa depan [3]. Para praktisi telah menggunakan teknologi komputer untuk mendukung pengambilan keputusan, dan machine learning menjadi solusi penting untuk membantu mendiagnosis pasien. Machine learning adalah alat analisis yang kuat yang dapat digunakan untuk tugas-tugas besar dan kompleks, seperti menganalisis rekam medis, memprediksi penyakit, dan menganalisis data.

Kemajuan dalam machine learning telah meningkatkan kesadaran untuk melakukan deteksi dini apakah seseorang beresiko menderita penyakit jantung. Algoritma machine learning dapat melatih dan mengevaluasi model untuk memprediksi penyakit jantung berdasarkan faktor-faktor yang memengaruhi tingkat risiko. Akurasi dipengaruhi oleh kualitas data, kinerja model, dan pemilihan parameter pelatihan. Dataset yang dilatih akan menghasilkan prediksi dengan tingkat akurasi yang berbeda berdasarkan kualitas data dan performa model[4]. Evaluasi terbaik memiliki fungsi penting dalam pengambilan keputusan, mengukur kinerja model, memonitor pengembangan dan pengujian model, dan memahami tingkat kegagalan. [5] Evaluasi dari berbagai algoritma dapat digunakan secara medis untuk menyimpulkan mana yang terbaik dan mana yang lebih tepat dan dapat digunakan sebagai acuan untuk proses pengambilan keputusan.

Pada penelitian ini, peneliti mengusulkan prediksi penyakit jantung dengan menggunakan tiga algoritma klasifikasi yaitu Logistic Regression, Random Forest, dan Decision Tree. Algoritma klasifikasi adalah metode machine learning yang dapat memprediksi kelas data. Penelitian ini bertujuan untuk mendapatkan klasifikasi yang paling akurat dalam mendiagnosis penyakit jantung.

1.2 Rumusan Masalah

Berdasarkan uraian yang dijelaskan didalam latar belakang, maka disusunlah sebuah rumusan masalah yaitu "Bagaimanakah metode machine learning untuk memperoleh prediksi penyakit jantung yang lebih optimal?"

1.3 Batasan Masalah

Untuk mencegah penyimpangan dari fokus penelitian, maka perlu ditentukan batasan masalah yang akan dibahas dan diteliti. Berikut ini adalah batasan masalah yang digunakan dalam penelitian ini:

1. Penelitian ini terbatas pada evaluasi performa model dalam melakukan klasifikasi.
2. Penerapan machine learning pada penelitian ini terbatas pada tiga algoritma yaitu Decision Tree, Random Forest dan Logistic Regression.
3. Metrik evaluasi terbatas pada beberapa penjabaran turunan dari confusion metric seperti akurasi, presisi, recall, f1-score dan Area Under Curve (AUC).

1.4 Tujuan Penelitian

Berdasarkan pernyataan penelitian kali ini, hasil yang akan dicapai adalah:

1. Mengaplikasikan bidang ilmu machine learning agar bermanfaat bagi ilmu pengetahuan dan pengembangan penelitian.
2. Memberikan kontribusi teknologi terhadap dunia medis, utamanya tindakan preventif berupa prediksi awal resiko penyakit jantung.
3. Membantu masyarakat dalam mengetahui prediksi dini resiko penyakit jantung, sehingga masyarakat dapat melakukan pencegahan dini terhadap bahaya resiko penyakit jantung.

1.5 Manfaat Penelitian

Adapun manfaat penelitian kali ini yaitu :

1. Dapat menjadi referensi bagi pengembangan ilmu pengetahuan dan keberlanjutan penelitian terutama di bidang machine learning.
2. Dapat diterapkan dalam dunia medis sebagai tindakan preventif berupa deteksi dini, sehingga terbatasnya jumlah tenaga medis tidak menghalangi peran medis dalam melakukan pelayanan.

3. Dapat meningkatkan kewaspadaan masyarakat dan kesadaran akan hidup sehat demi mengurangi resiko penyakit jantung.

1.6 Sistematika Penulisan

Pada penelitian kali ini penulis memuat uraian secara garis besar isi Skripsi untuk tiap-tiap bab. Pada bagian ini mendeskripsikan apa saja isi masing-masing bab yang akan disusun dan akan dijelaskan secara singkat dari Bab I, Bab II, Bab III, Bab IV Bab V, sebagai berikut :

Bab 1 PENDAHULUAN: Latar Belakang, Rumusan Masalah, Batasan masalah, Tujuan Penelitian, Manfaat Penelitian, Sistematika Penulisan

Bab 2 TINJAUAN PUSTAKA : Studi Literature, Tabel Keashian Penelitian, Machine Learning, Algoritma Decision Tree, Algoritma Random Forest, Algoritma Logistic Regression

Bab 3 METODE PENELITIAN : Alur Penelitian, Exploratory Data Analysis (EDA), Data Preprocessing, Anlysis (Train The Model), Evaluasi Kinerja

Bab 4 HASIL DAN PEMBAHASAN: Klasifikasi Algoritma Decision Tree, Klasifikasi Algoritma Random Forest, ,Klasifikasi Logistic Regression, Kurva ROC

Bab 5 PENUTUP : Kesimpulan, Saran, Referensi

