

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Jantung adalah organ vital dalam tubuh manusia yang berperan penting dalam sistem peredaran darah karena berfungsi mengalirkan darah yang mengandung oksigen dan nutrisi ke seluruh bagian tubuh [1, 2]. Penyakit jantung merupakan suatu kondisi yang disebabkan oleh adanya masalah pada pembuluh darah koroner, yang mengakibatkan penyempitan serta penyumbatan sehingga dapat mengganggu proses transportasi energi dalam tubuh, serta menyebabkan ketidakseimbangan antara kebutuhan dan suplai oksigen. Penyakit jantung merupakan penyakit kardiovaskular yang terjadi akibat penyumbatan pada arteri koroner oleh tumpukan plak, polutan, atau zat-zat kimia dari lingkungan yang masuk ke dalam tubuh melalui makanan, minuman, atau berbentuk gas yang kemudian menumpuk pada dinding arteri koronaria. Penyebab lainnya juga bisa berasal dari faktor keturunan [3, 4].

Menurut laporan Organisasi Kesehatan Dunia (WHO), pada tahun 2019, sekitar 17,9 juta orang meninggal akibat penyakit kardiovaskular (CVD). Angka ini mencakup 32% dari total kematian di seluruh dunia. Setiap tahunnya, lebih dari 17,7 juta orang kehilangan nyawa akibat penyakit ini [5, 6]. Deteksi dini penyakit jantung belum tertangani dengan optimal oleh metode diagnostik saat ini karena berbagai alasan, termasuk waktu yang dibutuhkan dan tingkat akurasi yang masih terbatas. Oleh karena itu, para ilmuwan terus berupaya mengembangkan metode yang lebih efektif untuk mengidentifikasi penyakit kardiovaskular sejak tahap awal. Di era modern ini, mendiagnosis dan mengobati penyakit jantung menjadi tantangan besar, terutama karena keterbatasan teknologi dan kurangnya tenaga medis yang ahli di bidang ini [7].

Pada dasarnya, penyakit jantung dapat dicegah melalui berbagai faktor, salah satunya adalah penerapan pola hidup sehat. Selain itu, deteksi dini juga sangat diperlukan untuk mencegah kematian pada penderitanya. Salah satu cara untuk melakukan deteksi dini adalah dengan memanfaatkan teknologi informasi seperti Machine Learning [8]. Penelitian sebelumnya menunjukkan bahwa penerapan Machine Learning telah terbukti efektif dalam membantu penilaian dini terhadap indikasi penyakit jantung, sehingga memungkinkan deteksi yang lebih cepat dan

akurat bagi pasien yang memerlukan perhatian medis [9].

Meskipun machine learning bukanlah teknologi dengan tingkat akurasi 100% dalam menganalisis data dan menghasilkan kesimpulan, keakuratan yang dihasilkan cukup efektif sehingga teknologi ini terbukti membantu di bidang kesehatan [10, 11]. Dalam dunia medis, peran machine learning sangat berguna dalam memberikan prediksi atau keputusan terhadap kondisi pasien [12]. Model Machine Learning (ML) memiliki kemampuan yang efisien dalam mengklasifikasikan jantung pasien sebagai normal atau tidak normal. Selain itu, ML juga dapat memprediksi kondisi pasien berdasarkan data yang sudah ada dalam waktu yang relatif singkat, sehingga membantu dalam mendeteksi penyakit jantung lebih cepat dan akurat [13].

Banyak pendekatan Machine Learning (ML) yang dikembangkan untuk mendeteksi penyakit jantung dengan menggunakan dataset pelatihan yang berisi data input dan output. Pendekatan ini memungkinkan sistem untuk belajar dari data sebelumnya dan menghasilkan prediksi yang lebih akurat [14]. Sebagai contoh, algoritma Machine Learning seperti Naïve Bayes (NB), Random Forest (RF) [15], Logistic Regression (LR) [16], K-Nearest Neighbor (K-NN) [17], Support Vector Machine (SVM), dan Artificial Neural Networks (ANN) [18] merupakan model-model canggih yang telah banyak digunakan untuk memprediksi pwnyakit jantung. Model-model ini tidak hanya dikembangkan di kalangan akademisi, tetapi juga terus disempurnakan oleh para profesional di industri [13]. Hal ini menjadi tantangan tersendiri di dunia medis, di mana akurasi yang sangat tinggi sangat dibutuhkan dalam mendeteksi penyakit jantung.

Di antara berbagai model Machine Learning, Support Vector Machine (SVM) telah terbukti memberikan performa yang efektif dalam berbagai permasalahan klasifikasi dan prediksi di berbagai bidang. Beberapa contoh penerapan SVM dalam kehidupan nyata antara lain adalah untuk mendeteksi kerusakan sistem, klasifikasi gambar, dan juga di bidang biomedis [13].

Support Vector Machine (SVM) adalah salah satu algoritma machine learning yang termasuk dalam kategori supervised learning. Algoritma ini umumnya digunakan untuk menyelesaikan masalah klasifikasi maupun regresi, di mana model dilatih menggunakan data yang sudah diberi label untuk mempelajari pola dan membuat prediksi di masa depan [19]. Pada dasarnya, SVM merepresentasikan data dari berbagai kelas dalam ruang multidimensi. Tujuan utamanya adalah memisahkan data ke dalam kelas-kelas yang berbeda dengan mencari hyperplane terbaik yang memiliki margin pemisah maksimal. Maka

dari itu, diperlukan optimasi pada SVM untuk menemukan jarak maksimum hyperplane dengan kedua kelas tersebut. Salah satu keunggulan dari metode ini adalah kemampuannya dalam melakukan komputasi secara cepat dan efisien [20]. Untuk meningkatkan tingkat optimal sebuah model, kini sudah tersedia metode yang bernama *Hyperparameter Tuning*, agar model tersebut memiliki *output* yang maksimal.

Penelitian sebelumnya mengenai prediksi penyakit jantung telah dilakukan dengan menggunakan metode Support Vector Machine (SVM) [6], dan dataset yang digunakan terdiri dari 302 baris dan 14 kolom. Dalam eksperimen tersebut, dilakukan penyetelan hyperparameter menggunakan GridSearchCV dan mendapatkan parameter C 1000, gamma 0.0001, dan menggunakan kernel RBF. Hasil yang diperoleh menunjukkan performa yang cukup baik, dengan nilai akurasi sebesar 86,0%, precision sebesar 84,0%, recall sebesar 91,0%, dan f1-score sebesar 87,0%.

Penelitian serupa yang menggunakan dataset yang sama telah dilakukan dengan menguji seluruh jenis fungsi kernel yang tersedia pada algoritma SVM. Hasil yang diperoleh menunjukkan bahwa kernel linear memberikan performa terbaik dengan accuracy sebesar 90,11%, precision sebesar 90,38%, dan recall sebesar 92,15%. Sementara itu, kernel RBF (Radial Basis Function) menghasilkan akurasi sebesar 69,23%, precision sebesar 86,27%, dan recall sebesar 67,69%. Kernel polynomial menunjukkan hasil yang sedikit lebih rendah, dengan akurasi sebesar 67,03%, precision sebesar 68,42%, dan recall sebesar 76,47%. Adapun kernel sigmoid memberikan hasil yang paling rendah, dengan akurasi sebesar 56,04%, precision sebesar 56,04%, dan recall hanya sebesar 1% [21]. Perbedaan pada penelitian ini adalah penggunaan standarisasi menggunakan standard scaler, handling outlier menggunakan z-score dan kernel yang dipakai untuk meneliti adalah kernel polynomial.

Penelitian serupa berikutnya menggunakan penggabungan 2 dataset, kedua dataset tersebut berasal dari UCI Machine Learning Repository, salah satu dataset tersebut dipakai pada penelitian ini. Hasil prediksi penggabungan 2 dataset yang diperoleh dari metode Random Forest dan AdaBoost menunjukan performa accuracy sebesar 96.16% dan precision sebesar 96.20% [22].

Penelitian yang lebih lanjut dan pengembangan teknologi ini diharapkan dapat memberikan solusi lebih optimal dalam mendiagnosis dan menangani penyakit jantung, sehingga dapat menurunkan angka kematian yang disebabkan oleh kondisi ini. Oleh karena itu, kolaborasi antara ilmuwan, tenaga medis, dan

teknologi sangat penting untuk mewujudkan sistem kesehatan yang lebih baik di masa depan.

1.2 Rumusan Masalah

Dari latar belakang penelitian ini didapatkan beberapa rumusan masalah sebagai berikut.

1. Bagaimana implementasi algoritma *Support Vector Machine (SVM)* untuk deteksi dini penyakit jantung?
2. Berapa akurasi, *precision*, *recall*, dan *f1-score* pada algoritma *Support Vector Machine (SVM)* setelah dilakukan hyperparametertuning untuk deteksi dini penyakit jantung?

1.3 Batasan Permasalahan

Untuk menjaga fokus serta cakupan penelitian tetap terarah, penelitian yang dilakukan memiliki batasan masalah sebagai berikut.

1. Penggunaan dataset “Statlog(Heart)” berupa data mentah yang berasal dari ”UCI Machine Learning Repository” dan dibuat pada tahun 1988.
2. Dataset yang digunakan memiliki 270 data dan 14 *features*.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini sebagai berikut.

1. Mengimplementasikan algoritma *SVM (Support Vector Machine)* untuk deteksi dini penyakit jantung.
2. Mengukur akurasi, *precision*, *recall* dan *f1-score* pada algoritma *SVM (Support Vector Machine)* setelah dilakukan hyperparamater tuning untuk deteksi dini penyakit jantung.

1.5 Manfaat Penelitian

Manfaat atas penelitian yang dilakukan adalah diharapkan jika model diterapkan pada aplikasi, membantu pihak medis membuat keputusan deteksi dini penyakit jantung sehingga pasien mendapatkan penanganan yang tepat.

Dari sisi keilmuan, penelitian ini memberikan kontribusi terhadap pengembangan ilmu pengetahuan di bidang ilmu komputer dan data sains, terutama dalam penerapan algoritma Support Vector Machine (SVM) pada kasus medis. Penelitian ini juga menyajikan analisis komparatif terhadap beberapa jenis kernel SVM (linear, polynomial, RBF, dan sigmoid) yang dapat menjadi referensi untuk penelitian selanjutnya. Selain itu, penelitian ini menunjukkan bahwa kombinasi preprocessing data menggunakan IQR untuk deteksi outlier, StandardScaler untuk standarisasi, serta SMOTE untuk penyeimbangan kelas, dapat meningkatkan performa model dalam klasifikasi penyakit jantung. Hasil dari penelitian ini diharapkan dapat menjadi dasar dan inspirasi bagi pengembangan lebih lanjut dalam sistem deteksi penyakit berbasis kecerdasan buatan.

1.6 Sistematika Penulisan

Berisikan uraian singkat mengenai struktur isi penulisan laporan penelitian, dimulai dari Pendahuluan hingga Simpulan dan Saran.

Sistematika penulisan laporan adalah sebagai berikut:

- **Bab 1 PENDAHULUAN**
Bab pendahuluan memuat unsur-unsur penting seperti latar belakang, rumusan masalah, batasan masalah, tujuan, serta manfaat dari penelitian. Bagian ini memberikan gambaran umum mengenai pokok permasalahan yang diangkat serta alasan mendasar dilaksanakannya penelitian tersebut.
- **Bab 2 LANDASAN TEORI**
Dalam penelitian ini, terdapat sejumlah teori dan metode yang digunakan, antara lain mengenai penyakit jantung, Machine Learning, Supervised Learning, Standard Scaler, teknik oversampling SMOTE, algoritma Support Vector Machine (SVM), serta evaluasi model menggunakan Confusion Matrix.
- **Bab 3 METODOLOGI PENELITIAN**
Metodologi penelitian berisi penjelasan mengenai tahapan-tahapan yang ditempuh selama proses penelitian. Pada bagian ini disajikan flowchart yang menggambarkan jalannya penelitian serta algoritma yang diterapkan dalam studi ini.
- **Bab 4 HASIL DAN DISKUSI**

Bab Hasil dan Diskusi memaparkan temuan dari penelitian beserta analisis yang dilakukan terhadapnya. Penyajian hasil disertai dengan visualisasi berupa gambar, tabel, serta uraian tertulis yang menjelaskan secara rinci pencapaian dan interpretasi dari penelitian ini.

- Bab 5 KESIMPULAN DAN SARAN

Bagian ini berisi kesimpulan dari hasil penelitian yang telah dilaksanakan serta saran yang dapat diberikan untuk pengembangan atau penelitian selanjutnya.

