

1. Pendahuluan

Penyakit jantung atau penyakit kardiovaskular adalah salah satu penyakit kompleks dan paling mematikan di dunia. Pada penyakit ini biasanya jantung tidak mampu mendorong jumlah darah yang dibutuhkan ke bagian tubuh lain untuk memenuhi fungsi normal tubuh dan terjadinya gagal jantung[1]. Dengan kurangnya aktifitas fisik karena perubahan gaya hidup, penyakit ini menjadi sangat umum bahkan pada kelompok usia muda. Merokok, kurang olahraga dan makanan berkolesterol tinggi adalah penyebab utama penyakit jantung[2]. Menurut Organisasi Kesehatan Dunia (WHO) diperkirakan 17 juta orang meninggal setiap tahun akibat penyakit kardiovaskular terutama serangan jantung dan stroke[3]. Pada saat ini kurangnya keahlian staff medis menghasilkan prediksi yang salah dalam melakukan tes untuk prediksi penyakit jantung[4]. Tantangan terhadap penanganan bedah penyakit jantung semakin sulit, terutama di negara berkembang yang kekurangan ahli tenaga medis terlatih serta sumber daya lain yang diperlukan untuk diagnosis dan perawatan yang tepat bagi pasien dengan masalah jantung[5].

Salah satu alternatif dalam mengidentifikasi dan memprediksi penyakit jantung adalah dengan memanfaatkan algoritma machine learning[6]. Pembelajaran mesin memungkinkan untuk memperoleh pengetahuan dari sejumlah data yang besar yang sangat berat bagi manusia bahkan terkadang tidak mungkin [3]. Para peneliti telah menerapkan teknik yang berbeda untuk mengembangkan algoritma klasifikasi. Terdapat beberapa masalah terkait dengan implementasi pembelajaran mesin untuk mendeteksi serangan jantung, seperti terbatasnya dataset medis, pemilihan fitur, aplikasi algoritma machine learning, dan kurangnya analisis yang mendalam[6].

Telah dilakukan beberapa penelitian terkait implementasi pembelajaran mesin untuk mendeteksi penyakit jantung. Pada tahun 2018, M. Bahaj melakukan studi dengan mengimplementasikan metode Fast Correlation Based Feature Selection (FCBC), Particle Swarm Optimization, dan Ant Colony Optimization (ACO) untuk deteksi serangan jantung. Mereka menemukan bahwa akurasi klasifikasi maksimal 99,65% dioptimalkan menggunakan model tersebut[3]. Pada tahun 2020, Spencer R melakukan studi Seleksi Fitur dan Metode Klasifikasi untuk Memprediksi Penyakit Jantung. Mereka menemukan bahwa model terbaik menggunakan kombinasi pemilihan fitur Chi square dengan algoritma BayesNet mencapai akurasi 85,00% pada kumpulan datanya[7]. Pada tahun 2021, M. Kavita melakukan studi dengan mengimplementasikan metode Random Forest, Decision tree dan Hybrid Model untuk deteksi serangan jantung. Mereka menemukan bahwa tingkat akurasi 88,7% melalui model prediksi serangan jantung dengan hybrid model[2]. Pada tahun 2021, Ghost P melakukan studi prediksi efisien penyakit kardiovaskular menggunakan algoritma pembelajaran mesin dengan Relief dan LASSO Teknik pemilihan fitur. Mereka menemukan bahwa model yang diusulkan menghasilkan akurasi tertinggi saat menggunakan metode pemilihan fitur Random Forest bagging Method (RFBM) dan Relief 99,05%[6]. Pada tahun 2022, Elescano Avendano melakukan studi Aplikasi Pembelajaran Mesin untuk Memprediksi Serangan Jantung pada Pasien dari Eropa. Mereka menemukan bahwa dengan model logistic regression diperoleh hasil akurasi 87% untuk orang yang menderita serangan jantung dan akurasi 81% untuk orang yang tidak akan menderita penyakit ini[8].

Berdasarkan survey literature tersebut, implementasi seleksi fitur pada pengembangan model prediksi serangan jantung belum banyak dilakukan. Salah satu metode yang dapat digunakan untuk seleksi fitur tersebut adalah pendekatan statistik, seperti Mutual Information. Terkait hal tersebut, penelitian ini bertujuan untuk mengimplementasi kombinasi seleksi fitur berbasis statistik dan support vector machine pada prediksi penyakit serangan jantung. Penelitian ini menggunakan metode seleksi fitur yaitu Mutual Information. Selanjutnya model prediksi dibangun menggunakan metode Support Vector Machine (SVM). SVM adalah metode yang kuat untuk membangun classifier. Hal ini bertujuan untuk membuat batas keputusan antara dua kelas yang memungkinkan prediksi label dari satu atau lebih vector fitur[9].

Rumusan Masalah

- Bagaimana mengimplementasikan Mutual Information (MI) dalam seleksi fitur
- Bagaimana prediksi serangan jantung menggunakan Support Vector Machine (SVM)
- Bagaimana performa dan validasi model Support Vector Machine (SVM)

Tujuan penelitian

- Mengimplementasikan Mutual Information (MI) dalam seleksi fitur
- Memprediksi serangan jantung menggunakan Support Vector Machine (SVM)
- Mengukur performa dan validasi model Support Vector Machine (SVM)