

BAB I

PENDAHULUAN

1.1 Latar Belakang

Penyakit jantung merujuk pada sejumlah yang mempengaruhi kinerja jantung yang dapat menyebabkan kematian dan penanganannya harus diatasi paling awal karena bisa terjadi secara tiba-tiba pada penderita. Beberapa faktor penyakit jantung berdasarkan kondisi tubuh pada penderita sehingga perlu diketahui sejak awal untuk memperkecil kemungkinan terjadinya serangan jantung. Serangan jantung dapat diatasi dengan beberapa cara seperti mengatur pola hidup yang sehat dan olahraga teratur[1].

Dalam penelitian ini, akan membandingkan performa tiga algoritma dalam memprediksi penyakit jantung. Algoritma yang akan dianalisis meliputi Support Vector Machine (SVM), Logistic Regression, dan Naive Bayes. Pada setiap algoritma ini memiliki kelebihan dan kekurangan dalam hal performa. Tujuan penelitian ini diharapkan dapat memberikan performa terbaik dalam memprediksi penyakit jantung berdasarkan dataset yang digunakan. Ketiga algoritma ini dibandingkan karena memiliki karakteristik dan pendekatan yang berbeda dalam menangani data dan membuat prediksi, selain itu juga untuk mencari faktor dari rendahnya akurasi pada algoritma yang digunakan[2].

Support Vector Machine dikenal karena kemampuannya dalam menemukan hyperplane yang digunakan untuk memisahkan kelas-kelas data dalam kasus data yang tidak linier. Logistic Regression model dasar untuk membandingkan kinerja model yang lebih kompleks. Naive Bayes efektif saat jumlah data terbatas, namun dengan fitur yang cukup untuk mendistribusikan kelas. Pemilihan ketiga algoritma tersebut didasarkan kegunaan dan keefektifan performa masing-masing untuk menganalisis data[3].

Tujuan penelitian ini yaitu untuk mencari algoritma mana yang memiliki performa terbaik, mengevaluasi dan membandingkan kinerja Support Vector Machine (SVM), Logistic Regression, dan Naive Bayes.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah, rumusan masalah dalam penelitian disusun sebagai berikut:

1. Dari ketiga algoritma tersebut, algoritma mana yang memberikan performa paling tinggi untuk memprediksi penyakit jantung?
2. Faktor apa saja yang menyebabkan hasil akurasi untuk algoritma yang mendapatkan performa paling rendah?

1.3 Batasan Masalah

Berdasarkan rumusan masalah, Batasan masalah disusun sebagai berikut:

1. Fokus pada perbandingan performa antara algoritma Support Vector Machine(SVM), Logistic Regression, dan Naive Bayes untuk memprediksi penyakit jantung. Tidak membahas secara keseluruhan tentang pengobatan dan sebagainya didunia medis.
2. Fokus menggunakan dataset yang didapatkan dari Kaggle <https://www.kaggle.com/datasets/rashikrahmanpritom/heart-attack-analysis-prediction-dataset> yang pernah digunakan untuk penelitian sebelumnya. Penelitian ini tidak mencantumkan Kumpulan data yang baru.
3. Fokus pada evaluasi yang digunakan untuk mengukur performa algoritma, seperti akurasi, presisi, recall, dan F1-score. Penelitian ini tidak mempertimbangkan metode lain yang mungkin digunakan juga ranah ini.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah diatas, tujuan penelitian disusun sebagai berikut:

1. Untuk mengetahui performa paling tinggi dari ketiga algoritma tersebut.
2. Untuk menemukan faktor penyebab hasil akurasi algoritma yang paling rendah.

1.5 Manfaat Penelitian

Berdasarkan tujuan penelitian, berikut manfaat penelitian yang disusun sebagai berikut:

1. Penelitian ini memberikan pemahaman yang lebih baik mendapatkan performa algoritma dari ketiga algoritma yaitu Logistic regression, Support Vector Machine (SVM), dan Naïve Bayes, dan dapat membantu pengembangan dan pemilihan metode klasifikasi di masa depan.
2. Penelitian ini dapat memberikan wawasan tentang kaitan dengan penyakit jantung, sehingga dapat melakukan upaya pencegahan dan melakukan gaya hidup sehat.
3. Bagi peneliti selanjutnya dapat ikut serta mengembangkan ketiga algoritma tersebut untuk menjadi pembandingan dengan metode yang lainnya.

1.6 Sistematika Penulisan

BAB I PENDAHULUAN, Pada bagian ini terdiri dari latar belakang, rumusan masalah, tujuan penelitian, batasan masalah, manfaat penelitian, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA, Bagian penting dalam sistematika penulisan sebuah skripsi atau karya ilmiah. pada bagian ini berisi kajian teori dan penelitian terdahulu yang relevan dengan topik penelitian.

BAB III METODE PENELITIAN, pada bagian ini untuk mendapatkan, mengumpulkan, menganalisis, dan menginterpretasikan data yang diperlukan untuk menjawab pertanyaan penelitian atau tujuan penelitian yang telah ditetapkan.

BAB IV HASIL DAN PEMBAHASAN, pada bagian ini merupakan temuan penelitian serta analisis dan interpretasi data.

BAB V PENUTUP, berisi tentang Kesimpulan dan penutup.