

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Salah satu penyakit yang sangat mematikan di dunia yaitu serangan jantung, penyakit tersebut merupakan peristiwa terjadinya hambatan pada aliran darah *arteri coroner* yang menyebabkan otot pada jantung memiliki kadar oksigen yang menurun sampai terjadi infark (D. Kurniawan et al., 2015). Faktor yang mempengaruhi terjadinya penyakit serangan jantung yaitu kebiasaan dalam merokok, kualitas tidur yang tidak seimbang dan mengkonsumsi makanan dengan kandungan kolestrol tinggi yang sudah menjadi kebiasaan bagi masyarakat seiring dengan perubahan pola hidup sehari-hari (Saktiningtyastuti & Astuti, 2017). Menurut penelitian yang dilakukan oleh (Bianto et al., 2020) menyatakan bahwa di Indonesia pada tahun 2012, terdapat sejumlah kasus berkisar 43,30% orang terkena serangan jantung dan menyebabkan jumlah kematian sebanyak 12,90%. Pada tahun 2013, tercatat sebanyak 61.682 orang terkena penyakit serangan jantung, dan penderita terkena penyakit tersebut dominan berada di Jawa Tengah dengan kasus sebanyak 11.511 kasus. Dan data-data tersebut juga diperkuat dengan adanya hasil pada penelitian yang telah dilakukan oleh (Baransyah L, 2014) menyatakan bahwa diperkirakan pada tahun 2020, penyakit serangan jantung menjadi penyakit pembunuh tersering didunia dengan sebesar 36% dari seluruh kematian, dan angka tersebut dinilai dua kali lebih besar dibandingkan kematian oleh kanker.

Dari hasil data diatas membuktikan bahwa penyakit serangan jantung sangat berbahaya, maka pada penelitian ini digunakan data pasien penyakit serangan jantung yang diambil lewat *website datasets* yaitu Kaggle.com. Data dari serangan jantung terlalu banyak untuk dilakukan pengolahan, sehingga data tersebut akan diproses lewat *data mining* menggunakan *machine learning* (Sulastri & Gufroni, 2017). Teknik *data mining* memiliki kaitan relasi dengan bidang ilmu pengetahuan dan teknologi seperti statistik data, kecerdasan buatan, dan *machine learning* (Nasir, 2021). Dalam *data mining* terdapat beberapa teknik pengelolaan data yang digunakan seperti salah satunya yaitu klasifikasi. Menurut (Septiani et al., 2017) para peneliti sering menggunakan klasifikasi untuk dijadikan sebagai proses menempatkan objek tertentu ke dalam satu pengelompokkan dengan dibagi menjadi beberapa kategori. Pada penelitian saat ini, metode klasifikasi tersebut dilakukan guna memprediksi seseorang terkena serangan jantung agar serangan jantung dapat dilakukan penanganan lebih awal. Berdasarkan penelitian yang dilakukan oleh (Ryfai et al., 2022) dalam melakukan klasifikasi penyakit serangan jantung dengan algoritma *K-Nearest Neighbor* (KNN) menghasilkan tingkat akurasi yang pertama sebanyak 88.333%, dan pada percobaan kedua menghasilkan tingkat akurasi sebanyak 96.667%. Berdasarkan penelitian yang telah dilakukan (Pradana et al., 2022) dalam melakukan klasifikasi penyakit serangan jantung dengan algoritma *Support Vector Machine* (SVM) menghasilkan tingkat akurasi percobaan pertama sebanyak 87%, dan pada percobaan kedua memberikan tingkat akurasi sebanyak 80%.

Dalam pengelolaan klasifikasi data penelitian ini menggunakan algoritma seperti *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN). Menurut (Maisat et al., 2023) algoritma *Support Vector Machine* (SVM) merupakan suatu metode klasifikasi pada *data linear* dan juga *non-linear* yang memanfaatkan fungsi dari suatu kernel secara sistematis yang membuat *support vector classifier* ke dalam bentuk suatu dimensi yang lebih tinggi. Cara kerja dari metode tersebut yaitu dengan cara menentukan jarak terjauh dari suatu *hyperplane* dengan cara berulang kali hingga mendapatkan suatu jarak *hyperplane* yang terbaik pada kedua *class* tersebut. Menurut (Taufiq et al., 2022) algoritma *K-Nearest Neighbor* (KNN) yaitu suatu metode klasifikasi yang digunakan dengan cara mengumpulkan beberapa objek hasil dari data yang telah dilakukan proses *data training* yang mempunyai jarak terdekat dengan kemiripan tertinggi pada data pengujian. Dengan digunakan kedua algoritma tersebut, menurut (Aulia et al., 2015) perbandingan tersebut memiliki kelebihan yang sama yaitu algoritma *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN) yang proses pengelompokkan data secara sederhana dan umum digunakan, memiliki kemudahan dalam pengimplementasian, serta memiliki sebuah performansi yang baik pada metode tersebut.

Oleh karena itu, penelitian ini akan mengumpulkan data serangan jantung terlebih dahulu menggunakan *datasets* yang diperoleh dari situs web www.Kaggle.com, dan akan diimplementasikan lewat klasifikasi pada algoritma *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN), hasil dari klasifikasi tersebut akan dilakukan perbandingan untuk mendapatkan nilai dengan

tingkat akurasi pada kedua algoritma yang lebih baik dalam mendeteksi penyakit serangan jantung.

1.2 RUMUSAN MASALAH

Berdasarkan latar belakang yang telah dijelaskan diatas, dapat ditarik rumusan dari permasalahan yang harus di selesaikan pada penelitian ini yaitu :

- 1) Bagaimana menganalisis algoritma *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN) pada pengklasifikasian penyakit serangan jantung?
- 2) Bagaimana cara mengimplementasikan algoritma *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN) pada klasifikasi penyakit serangan jantung?
- 3) Bagaimana hasil dari tingkat akurasi algoritma *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN) pada klasifikasi penyakit serangan jantung?

1.3 TUJUAN PENELITIAN

Adapun tujuan dari dilakukannya penelitian ini adalah sebagai berikut :

- 1) Menganalisis dan memprediksi menggunakan algoritma *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN) pada klasifikasi penyakit serangan jantung.
- 2) Membandingkan tingkat akurasi algoritma *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN) pada klasifikasi penyakit serangan jantung.

1.4 BATASAN MASALAH

Adapun permasalahan yang harus dibatasi guna dalam menghindari permasalahan yang akan meluas, maka batasan masalah dalam penelitian tersebut adalah sebagai berikut :

- 1) Dataset yang akan digunakan merupakan data gejala penyakit serangan jantung yang diperoleh dari *website* Kaggle.com.
- 2) File yang digunakan dengan format .csv.
- 3) Penelitian ini dilakukan hanya menggunakan algoritma *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN).

1.5 MANFAAT PENELITIAN

Manfaat yang dapat diberikan pada penelitian ini adalah sebagai berikut :

1. Memberikan pengetahuan tentang cara pengklasifikasian penyakit serangan jantung dengan algoritma *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN).
2. Mengetahui hasil tingkat akurasi yang lebih akurat dalam algoritma klasifikasi *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN) pada penyakit serangan jantung.

1.6 SISTEMATIKA PENULISAN

Adapun sistematika dalam penulisan yang terdiri dari 5 (lima) bab yang akan diberikan penjelasan sebagai berikut :

BAB I : PENDAHULUAN, Pada bab ini berisikan tentang penjelasan dari latar belakang, rumusan masalah, tujuan penelitian, batasan masalah, manfaat penelitian, dan sistematika penulisan.

BAB II : LANDASAN TEORI, Pada bab ini berisikan tentang penjelasan penyakit serangan jantung, *dataset*, *machine learning*, *data mining*, metode klasifikasi, *Support Vector Machine* (SVM), *K-Nearest Neighbor* (KNN), dan penelitian terdahulu sebagai acuan di dalam penelitian agar lebih terstruktur dan mudah dipahami.

BAB III : METODOLOGI PENELITIAN, Pada bab ini berisikan mengenai kerangka pikir, metode pengumpulan data, hipotesis, tahapan penelitian, tahapan penyelesaian, prediksi algoritma *Support Vector Machine* (SVM), prediksi algoritma *K-Nearest Neighbor* (KNN), dan *confusion matrix*.

BAB IV : HASIL PEMBAHASAN, Pada bab ini menjelaskan mengenai hasil pengumpulan data, hasil penelitian, dan pembahasan sebuah klasifikasi dari prediksi penyakit serangan jantung dan melakukan perbandingan tingkat akurasi dari algoritma *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN).

BAB V : PENUTUP, Pada bab ini menjelaskan kesimpulan dari hasil penelitian yang didapatkan diatas, keterbatasan dari penelitian, dan saran dari penulis.