

BAB I PENDAHULUAN

1.1 Latar Belakang

Jantung merupakan salah satu organ vital dalam tubuh manusia dan merupakan pusat dari sistem peredaran darah, dengan melibatkan pembuluh darah sebagai salurannya. Oleh karena itu, jantung memegang peranan penting dalam menyediakan oksigen bagi seluruh tubuh dan memiliki tugas untuk membersihkan tubuh dari hasil metabolisme. Namun, sejalan dengan peran pentingnya, jantung juga merupakan salah satu organ tubuh yang paling rentan untuk terkena penyakit. Penyakit kardiovaskular atau yang biasa disebut penyakit jantung umumnya mengacu pada kondisi yang melibatkan penyempitan atau pemblokiran pembuluh darah yang bisa menyebabkan serangan jantung, nyeri dada (*angina*) atau stroke (WHO,2017).

Penyakit ini merupakan pembunuh paling berbahaya saat ini yang menjadikannya sebagai penyebab kematian nomor satu di dunia (WHO,2017). Berbagai macam penelitian *machine learning* telah dilakukan, salah satunya adalah penelitian pemanfaatan *machine learning* untuk memprediksi penyakit jantung. *Machine Learning* adalah bagian dari penelitian tentang kecerdasan buatan, mencari untuk memberikan pengetahuan kepada komputer melalui data, pengamatan dan berinteraksi dengan dunia. Pengetahuan yang diperoleh memungkinkan komputer untuk melakukan generalisasi dengan benar ke pengaturan baru (Yoshua Bengio, 2016).

Terdapat beberapa metode dalam pengimplementasian *machine learning*, antara lain Naïve Bayes dan Decision Tree. Naïve Bayes adalah salah satu algoritma klasifikasi berdasarkan teorema Bayesian pada statistika (Suntoro,Wahyu, & Indriyawati,2018). Algoritma Naïve Bayes dapat digunakan untuk memprediksi probabilitas keanggotaan suatu kelas (Han & Kamber,2012). Metode Decision Tree merupakan metode klasifikasi yang menyerupai pohon.

Algoritma Decision Tree mengonstruksi pohon keputusan dari sebuah data training yang berupa record-record dalam basis data (Larose,2005).

Kelebihan dari algoritma Naïve Bayes adalah algoritma ini mudah digunakan, lebih cepat dalam perhitungan, dan menghasilkan akurasi yang tinggi (Hamzah, 2012). Selain itu algoritma Naïve Bayes sangat cocok untuk digunakan pada tipe data nominal. Algoritma Decision Tree juga memiliki beberapa kelebihan, diantaranya adalah memiliki akurasi yang baik, dapat secara eksplisit menggambarkan suatu pola/pengetahuan/informasi dalam bentuk pohon keputusan (Suntoro & Indah,2017), dapat menghilangkan perhitungan yang tidak diperlukan. Karena dengan metode ini, sampel hanya diuji berdasarkan kriteria atau kelas tertentu.

Pada penelitian prediksi penyakit kegagalan jantung dengan menggunakan metode data mining Artificial Neural Network (ANN) dan K Mean Clustering yang dilakukan oleh Andre D'Souza (2015), memberikan hasil yaitu, kedua metode tersebut menghasilkan rata-rata akurasi masing-masing 79.38% untuk *Artificial Neural Network* (ANN) dan 63.299% untuk K Means Clustering dengan tiga belas atribut *predictor*, satu atribut target dengan jumlah data sebanyak 303 data.

Sedangkan menurut penelitian prediksi dan diagnosis penyakit jantung dengan teknik data mining menggunakan tools Weka yang dilakukan oleh Boshra Bahrami dan Mirsaed Hosseini Shirvani (2015), memberikan hasil yaitu, untuk metode Naïve Bayes memberikan akurasi rata-rata sebesar 81.8% dan Decision Tree dengan jenis j48 menghasilkan rata-rata akurasi sebesar 83.732%.

Menurut penelitian yang dilakukan oleh Fahd Saleh Altaibi (2019) tentang implementasi *machine learning* untuk prediksi penyakit kegagalan jantung, dimana dilakukannya pengujian diantaranya menggunakan metode Naïve Bayes dan Decision Tree, menghasilkan rata-rata akurasi yang tinggi untuk kedua metode tersebut, dimana Naïve Bayes menghasilkan rata-rata akurasi sebesar 87.27% dan Decision Tree menghasilkan rata-rata akurasi sebesar 93.19%.

Menurut Andre D'Souza (2015), bahwa yang menyebabkan pengujian nilai akurasi yang kurang maksimal pada proyeknya adalah bahwa kurangnya data set karena tidak tersedia. Dataset yang tersedia tidak diperbarui dan tidak mengandung parameter yang relevan melalui penelitian terbaru.

Oleh karena itu, berdasarkan latar belakang dan hasil penelitian sebelumnya, penelitian ini bertujuan untuk meningkatkan akurasi dalam studi kasus prediksi penyakit jantung dengan cara menambahkan jumlah data yang diuji dan membuktikan metode mana yang memberikan akurasi lebih baik antara Naïve Bayes atau Decision Tree. Pengambilan data pada penelitian ini berupa data sekunder penyakit jantung yang didapat dari website Kaggle. Hasil dari penelitian ini akan membandingkan akurasi mana yang lebih baik antara metode Naïve Bayes atau Decision Tree.

1.2 Rumusan Masalah

Berdasarkan latar belakang pada uraian di atas maka rumusan masalah yang akan dibahas adalah bagaimana menerapkan machine learning untuk memprediksi penyakit jantung dengan menggunakan metode Naïve Bayes dan Decision Tree ?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah diatas, tujuan diadakannya penelitian ini adalah untuk mengetahui bagaimana cara menarapkan *machine learning* untuk prediksi penyakit jantung menggunakan Naïve Bayes dan Decision Tree dan membuktikan antara metode Naïve Bayes atau Decision Tree yang memberikan hasil akurasi yang lebih.

1.4 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut

1. Mempermudah klasifikasi penyakit kegagalan jantung sehingga dapat membantu dokter dalam memberikan penanganan.

2. Dapat mengetahui algoritma mana yang memberikan akurasi prediksi lebih baik antara Naïve Bayes atau Decision Tree.

1.5 Batasan Masalah

Ditentukannya batasan masalah bertujuan agar ruang lingkup penelitian tidak terlalu luas, maka ditetapkan batasan masalah sebagai berikut:

1. Data yang digunakan adalah data sekunder yang didapat dari website Kaggle.
2. Data jantung yang digunakan adalah pada orang dewasa dengan rentang usia 29-77 tahun.
3. Metode yang digunakan adalah Naïve Bayes dan Decision Tree dengan jenis CART.