

## 1. Pendahuluan

### 1.1. Latar Belakang

Kardiovaskular atau disebut juga sakit jantung merupakan kondisi dimana terjadi gangguan pada jantung dan pembuluh darah. Kardiovaskular ditandai dengan adanya gangguan pada jantung, otot skelet, kelainan fungsi ginjal, stimulai sistem saraf simpatis dan perubahan neurohormonal yang kompleks[1]. Angka kematian yang disebabkan oleh kardiovaskular terus meningkat seiring dengan pertumbuhan usia dimana terdapat 100 kasus per 1000 orang pada usia di atas 60 tahun. Faktor penyebab kardiovaskular antara lain adalah diabetes, hipertensi, kebiasaan merokok, kelebihan berat badan, *stress* dan kolesterol. Selain itu, faktor yang melekat pada diri pasien atau tidak dapat dihindari sama sekali yakni faktor keturunan dan latar belakang keluarga, jenis kelamin serta usia yang banyak ditemui pada kasus kardiovaskular khususnya gagal jantung [1]. Kurangnya kesadaran terhadap pola hidup yang sehat dan kurangnya informasi terkait penyakit jantung menyebabkan kebanyakan pasien penderita penyakit jantung tidak mengetahui gejala-gejala awal yang dirasakannya. Hal ini mendorong banyaknya penelitian terhadap penyakit kardiovaskular, salah satunya menggunakan metode berbasis komputer.

Beberapa penelitian yang membahas tentang penyakit kardiovaskular diantaranya, penelitian yang dilakukan oleh Siti Holis Sumartini dkk pada tahun 2015, tentang klasifikasi rekurensi pasien kanker serviks menggunakan metode *Classification and Regression Trees* (CART). Pada penelitian ini didapatkan hasil bahwa klasifikasi CART mampu menghasilkan rata-rata nilai akurasi untuk data *learning* sebesar 77,75% dan 69,14% untuk data prediksi.

Selanjutnya pada tahun 2018 dilakukan penelitian oleh Abdul Rohman dkk yang memprediksi penyakit jantung menggunakan algoritma C4.5. Pada penelitian ini digunakan pengukuran dengan *confusion matrix* dan *curve ROC* dan diketahui bahwa algoritma C4.5 untuk klasifikasi menghasilkan nilai akurasi sebesar 86,59%.

Pada tahun 2020, Derisma juga melakukan penelitian tentang perbandingan kinerja algoritma untuk prediksi penyakit jantung menggunakan teknik data mining. Algoritma yang digunakan yaitu *Naive Bayes*, *Random Forest* dan *Neural Network*. Hasil dari penelitian ini menunjukkan bahwa algoritma *Naive Bayes* adalah algoritma yang tepat dan akurat digunakan untuk dapat melakukan prediksi penyakit jantung dengan persentase 83%.

Penelitian terkait penyakit kardiovaskular telah banyak dilakukan sebelumnya, tetapi penelitian terkait prediksi penyakit kardiovaskular dengan membandingkan kinerja antara algoritma *Classification and Regression Tree* (CART) dan *Random Forest* belum banyak dilakukan. Selain itu rata-rata akurasi yang dihasilkan pada penelitian terdahulu masih dapat ditingkatkan. Berdasarkan hal tersebut, penulis ingin melakukan penelitian terhadap kinerja algoritma CART dan *Random Forest* dalam memprediksi penyakit kardiovaskular, serta melihat apakah dapat dihasilkan kinerja yang lebih baik dari penelitian terdahulu.

### 1.2. Topik dan Batasannya

Penelitian ini dilakukan untuk mengetahui faktor utama penyakit kardiovaskular dan menghitung serta membandingkan akurasi algoritma *Classification and Regression Tree* (CART) dan *Random Forest* dalam memprediksi penyakit kardiovaskular. Dataset yang digunakan memiliki 13 parameter dan dua *class*. Melalui proses pemangkasan pohon (*pruning*) pada pohon maksimal, akan dihasilkan pohon optimal oleh algoritma *Classification and Regression Tree* (CART), yang merupakan hasil prediksi faktor penyebab penyakit kardiovaskular, kemudian kinerja dari *Classification and Regression Tree* (CART) dan *Random Forest* akan dibandingkan untuk mengetahui algoritma terbaik diantara keduanya untuk memprediksi penyakit kardiovaskular.

### 1.3. Tujuan

Tujuan yang ingin dicapai dalam tugas akhir ini adalah membangun sistem yang dapat menentukan faktor utama penyakit kardiovaskular dengan menggunakan algoritma *Classification and Regression Tree* (CART) dan *Random Forest*, kemudian menentukan dan membandingkan akurasi dari kedua algoritma tersebut.

### 1.4. Organisasi Tulisan

Tugas akhir ini terdiri dari beberapa bagian. Bagian pertama berisi latar belakang dan identifikasi masalah yang menjadi acuan dalam tugas akhir ini. Bagian selanjutnya berisi penjelasan tentang studi literatur dengan teori-teori yang digunakan. Dan bagian terakhir menjelaskan tentang sistem yang dibangun dan bagaimana sistem dapat mengimplementasikan hasil dan evaluasi yang dilakukan, kemudian penjelasan hasil dan evaluasi yang didapatkan dari penelitian yang dilakukan.