

BAB 1 PENDAHULUAN

1.1 Latar Belakang

Jantung sebuah rongga organ berotot yang memompa darah melalui pembuluh darah dengan kontraksi berirama yang terus berulang merupakan salah satu organ manusia yang berperan dalam sistem peredaran darah. Darah kemudian menyuplai oksigen dan nutrisi pada tubuh, dimana darah juga membantu menghilangkan sisa-sisa dari metabolisme. Letak Jantung berada di rongga dada sekitar sebelah kiri. Jantung sebagai salah organ terpenting dalam tubuh memiliki resiko kematian jika ada kelainan yang terjadi pada jantung. Beberapa masalah pada jantung dibagi menjadi dua yaitu penyakit jantung dan serangan jantung. WHO berdasarkan data menyatakan bahwa ada sebanyak 7,3 juta penduduk di dunia yang meninggal dikarenakan penyakit jantung. (*World Health Organization*, (2020). Tipe penyakit jantung terjadi dikarenakan jantung tidak dapat melaksanakan tugasnya dengan baik, seperti Otot jantung yang lemah atau adanya celah antara serambi kanan dan serambi kiri. (Tegar, 2019)

Mengambil data dari tulisan *Center for Disease Control and Prevention*, penyakit jantung merupakan penyebab utama meninggalnya wanita, pria dan beberapa kelompok ras dan juga etnis di Amerika Serikat, paling tidak ada satu orang yang meninggal setiap 36 detik dari penyakit jantung. Indonesia berdasarkan data Riset Kesehatan Dasar (Riskesdas) pada tahun 2018 memiliki angka kejadian penyakit jantung dan juga pembuluh darah yang terus meningkat dari tahun ke tahun, 15 dari 1000 orang atau sekitar 4,2 juta individu di Indonesia menderita penyakit jantung. Penyakit jantung yang terjadi perlu didiagnosa oleh dokter dengan menjalankan serangkaian tes dan evaluasi, yaitu pemeriksaan fisik, tes darah, tes dalam kategori non invasif seperti *stress test* dan

elektrokardiogram dan tes invasif seperti katerisasi pada jantung. Semua tes ini akan menghasilkan data pasien seperti kadar kolesterol, tekanan darah dan beberapa data lainnya yang kemudian dapat membantu dalam proses mendiagnosis penyakit pasien. (Donovan, 2018)

Industri kesehatan telah berkembang sangat pesat salah satunya dikarenakan teknologi yang berkembang pesat, sehingga perkembangan ini membuka pintu untuk melakukan lebih banyak penelitian. Kesehatan merupakan hal yang penting untuk dijaga, dimana kesalahan pengobatan atau pencegahan dapat menyebabkan kehilangan nyawa, terutama pada penyakit jantung. Peningkatan teknologi pada industri kesehatan salah satunya adalah melakukan digitalisasi informasi medis terutama data pasien, dengan harapan informasi yang telah digitalisasi dapat digunakan untuk penelitian dan hasil penelitian tersebut dapat meningkatkan layanan industri kesehatan. (NEJM Catalyst, 2018)

Tentunya perkembangan teknologi yang pesat ini mempunyai beberapa masalah yang perlu diperhatikan, seperti bagaimana penanganan informasi medis yang didapatkan akan mempunyai ukuran data yang cukup besar dikarenakan perkembangan industri, hal yang perlu kita perhatikan adalah bagaimana bisa kita uraikan, kemudian analisa sehingga kita mendapatkan informasi yang penting dan berguna dimana kemudian dapat diaplikasikan dalam layanan pada industri kesehatan.

Machine Learning kemudian memegang peran penting pada proses penanganan dan analisa data medis, dengan *Machine Learning* kita dapat mencoba berbagai metode-metode yang tersedia untuk menemukan pola atau informasi yang penting dan juga berguna dimana diharapkan dapat meningkatkan kualitas layanan kesehatan saat ini.

Dalam beberapa penelitian yang memiliki kemiripan dalam melakukan klasifikasi penyakit jantung, metode-metode yang digunakan umumnya menghasilkan nilai akurasi yang berbeda.

Penelitian Mufti Ari Bianto, Kusri dan Sudarmawan yang diterbitkan pada tahun 2019. Menggunakan metode *Naive Bayes* menghasilkan nilai rata-rata akurasi sebesar 90,61%.

Penelitian oleh Jimmi Afriando Akbar, Zidane Ibrahim Fadela, Luthfi Fachruddin, Feby Ardiansyah dan Zuhdi Mukarram Bakhri yang diterbitkan pada tahun 2021 menjelaskan bahwa dengan menggunakan metode *CART Decision Tree*, mendapatkan akurasi sebesar 80%

Penelitian Madhumita Pal dan Smita Parija yang diterbitkan Maret 2021 menjelaskan bahwa penerapan klasifikasi pada 303 data penyakit jantung menggunakan algoritma *Random Forest*, mendapatkan hasil akurasi sebesar 86.69%.

Mengacu pada Teorema *No Free Lunch* dimana dijelaskan bahwa semua algoritma umumnya akan bekerja secara optimal ketika kinerjanya rata-rata untuk setiap fungsi dan tujuan yang mungkin (Brownlee, 2021). Hal ini kemudian memberikan pertanyaan bagaimana jika ada tiga metode klasifikasi yang ingin digunakan, bagaimana caranya membandingkan kedua algoritma tersebut dan memilih algoritma yang terbaik.

Berdasarkan penelitian terdahulu di atas, maka pada penelitian akan dilaksanakan dengan mendapatkan data, kemudian mengolahnya berdasarkan atribut yang relevan, lalu menggunakan metode *Decision Tree*, *Naive Bayes* dan *Random Forest* untuk mendapatkan hasil akurasi pada setiap metode model prediktif yang optimal.

Penelitian ini diharapkan dapat membantu peneliti di bidang penyakit jantung dalam memilih model-model atau metode yang tepat di dalam penelitiannya dan dapat membantu peningkatan layanan ahli medis.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan diatas, maka, bisa dirumuskan permasalahan yang dapat diambil sebagai berikut:

1. Bagaimana penerapan metode *Decision Tree*, *Naive Bayes* dan *Random Forest* pada klasifikasi penyakit jantung?
2. Bagaimana hasil akurasi metode yang digunakan, dan bagaimana hasil perbandingan antara *Decision Tree*, *Naive Bayes* dan *Random Forest* yang digunakan?

1.3 Tujuan Penelitian

Penelitian ini mempunyai tujuan yaitu untuk mendapatkan bagaimana mengolah dan melakukan analisa data, bagaimana penerapan metode *Decision Tree*, *Naive Bayes* dan *Random Forest* pada klasifikasi penyakit jantung, kemudian bagaimana hasil akurasi metode-metode yang digunakan tersebut, bagaimana hasil perbandingan antara *Decision Tree*, *Naive Bayes* dan *Random Forest* yang digunakan dan metode apa yang merupakan terbaik dari klasifikasi penyakit jantung

1.4 Batasan Masalah

Adapun Batasan masalah di dalam penelitian ini dapat disimpulkan sebagai berikut:

1. Menggunakan kumpulan data penyakit jantung berjudul “*Personal Key Indicators of Heart Disease*”. Diajukan oleh Kamil Pytlak. format kumpulan data adalah *csv*.
<https://www.kaggle.com/datasets/kamilpytlak/personal-key-indicators-of-heart-disease>
2. Menggunakan algoritma *Decision Tree*, *Naive Bayes* dan *Random Forest* pada pembuatan model *machine learning* yang kemudian akan diperbandingkan dengan nilai evaluasi setiap model

1.5 Luaran Penelitian

Adapun luaran dari penelitian berupa informasi algoritma yang terbaik dalam melakukan klasifikasi dataset penyakit jantung berjudul “*Personal Key Indicators of Heart Disease*”

1.6 Sistematika Pembahasan

Skripsi ini terdiri dari 5 (Tiga) bab dengan beberapa sub bab, berikut arah dan gambaran dari setiap bab yang tertulis dengan sistematika penulisannya:

BAB I PENDAHULUAN membahas latar belakang, perumusan masalah, tujuan dari penelitian, batasan masalah dan juga sistematika penulisan pada setiap bab.

BAB II TINJAUAN PUSTAKA membahas tinjauan pustaka yang penulis gunakan berupa teori, pengertian, algoritma, metode, teknik yang membantu dan juga berkaitan dalam proses penelitian.

BAB III METODOLOGI PENELITIAN membahas setiap tahapan yang dijalankan pada proses penelitian, mulai dari menggunakan dan menggali informasi dari kumpulan data yang terkait, melakukan pra proses pada kumpulan data, hingga evaluasi hasil model.

BAB IV HASIL DAN PEMBAHASAN berisi mengenai pembahasan dari hasil setiap tahapan penelitian yang telah dilakukan, dimana hasilnya akan ditampilkan dalam bentuk visual grafik, tabel dan analisa data.

BAB V PENUTUP berisi mengenai kesimpulan dan saran dari penelitian ini. Isi kesimpulan akan didasarkan dari hasil penelitian yang telah dilakukan pada Bab IV, kemudian untuk saran akan berisikan apa saja saran-saran perbaikan atau hal yang dapat ditingkatkan terhadap penelitian yang telah dijalankan oleh penulis, untuk membantu pembaca sebagai panduan pada penelitian selanjutnya.