

BAB I

PENDAHULUAN

1.1 Latar Belakang

Jantung merupakan organ vital manusia yang merupakan perlindungan terakhir manusia untuk bisa hidup selain otak dan memiliki fungsi dalam tubuh manusia sebagai pusat beredarnya darah [1]. Penyakit jantung adalah kondisi dimana jantung manusia tidak bisa bekerja secara normal dan tidak dapat menjalankan tugasnya dengan baik, meskipun penyakit jantung tidak dapat menular namun penyakit jantung adalah penyakit yang serius dan mematikan pertama di dunia. Data WHO mengatakan bahwa sejumlah 7,3 juta penduduk meninggal dunia disebabkan oleh penyakit jantung [2].

Tingkat kematian yang disebabkan oleh penyakit ini semakin bertambah seiring berkembangnya pola hidup yang kurang sehat di dalam kehidupan masyarakat. Berbagai hal yang menjadi penyebab munculnya penyakit jantung [3], faktor utama penyebab penyakit jantung antara lain adalah merokok, mengkonsumsi minuman beralkohol, mengkonsumsi makanan berkolesterol tinggi, kelebihan berat badan, diet yang tidak sehat, umur, dan mempunyai tekanan darah tinggi [2].

Prosedur untuk mengetahui penyakit jantung dizaman sekarang umumnya masih memakai cara manual yakni dengan cara tatap muka langsung ke dokter spesialis jantung dan juga perlu menjalani pemeriksaan lab yang selanjutnya perlu dikonsultasikan kembali ke dokter spesialis penyakit jantung. Metode tersebut masih mengeluarkan banyak biaya seperti untuk konsultasi ke dokter spesialis dan pemeriksaan laboratorium belum juga menghabiskan banyak waktu untuk menyesuaikan dengan jadwal praktek dokter spesialis jantung yang akan ditemui, oleh karenanya dibutuhkannya sebuah sistem yang bisa memprediksi untuk membantu dalam mendeteksi dini penyakit jantung sehingga tidak perlu melakukan proses diatas [3].

Untuk mempermudah mendeteksi penyakit jantung bisa dilakukan dengan berbagai metode klasifikasi, pada penelitian ini menggunakan metode *K-Nearest Neighbor* (KNN). Metode KNN adalah salah satu dari algoritma klasifikasi dengan

menghitung nilai kedekatan jarak suatu data terhadap data yang ada sesuai dengan nilai parameter K yang ditentukan, nilai kedekatan jarak tersebut menjadi patokan kemiripan suatu data terhadap data yang lain [2].

Metode K -NN adalah metode yang diunggulkan untuk klasifikasi dengan *training* data yang mempunyai banyak *noise* dan juga baik apabila digunakan untuk *training* data yang luas, namun metode K -NN memiliki kelemahan yaitu harus menentukan nilai K yang terlalu subjektif selain itu waktu komputasi cukup lama karena menghitung jarak setiap data terhadap data training [2].

Berdasarkan penjelasan tersebut maka penulis akan melaksanakan penelitian yang berjudul “Identifikasi Penyakit Jantung Menggunakan Metode *K-Nearest Neighbor*”. Pada penelitian ini peneliti akan melakukan identifikasi penyakit jantung dengan menggunakan metode KNN. Diharapkan dengan penelitian ini dapat mendapatkan nilai akurasi yang tinggi.

1.2 Perumusan Masalah

1.2.1 Permasalahan

Banyak orang meninggal dunia disebabkan oleh penyakit jantung karena keterlambatan dalam menanganinya, oleh karena itu dibutuhkan tindakan awal sebagai solusi dari permasalahan penyakit jantung, yaitu dengan melakukan diagnosa guna mendeteksi dini penyakit jantung. Maka dibutuhkan sistem identifikasi penyakit jantung.

1.2.2 Metode Usulan

Implementasi metode *K-Nearest Neighbor* sebagai metode klasifikasi untuk mendeteksi dini penyakit jantung.

1.2.3 Pertanyaan Penelitian

Berapa akurasi metode KNN dalam identifikasi penyakit jantung?

1.3 Tujuan dan Manfaat Penelitian

1.3.1 Tujuan Penelitian

1. Membuat dan merancang sistem identifikasi klasifikasi diagnosis penyakit jantung dengan menggunakan metode KNN.

2. Mengetahui akurasi yang dihasilkan oleh metode KNN dalam identifikasi penyakit jantung.

1.3.2 Manfaat Penelitian

Manfaat dari penelitian ini diharapkan dapat mengetahui nilai akurasi dari metode *K-Nearest Neighbor* (KNN). Dan untuk membantu para ahli mendiagnosa penyakit jantung.

1.4 Batasan Masalah

Mengenai batasan masalah dari penelitian ini yakni:

1. Sistem yang dibuat berbasis Web dengan menggunakan Bahasa pemrograman *Python*.
2. Data yang digunakan sebanyak 4240 data dengan jumlah fitur 15 yaitu, *Gender, Age, Education, currentSmoker, cigsPerDay, BPMeds, PrevalentStroke, PrevalentHyp, Diabetes, TotChol, SysBP, DiaBP, BMI, HeartRate, Glucose*, dan 2 label yaitu 1 dan 0, 1 berarti positif penyakit jantung, sedangkan 0 berarti negatif penyakit jantung. Data ini diperoleh dari *Framingham Heart Study* Dataset dari situs [Kaggle.com](https://www.kaggle.com/amanajmera1/framingham-heart-study-dataset/data) (<https://www.kaggle.com/amanajmera1/framingham-heart-study-dataset/data>).

1.5 Sistematika Laporan

Mengenai sistematika yang digunakan..dalam..penulisan skripsi terdiri..dari beberapa..bab..yaitu:

BAB I PENDAHULUAN

Terkait dalam bab..ini akan menjelaskan latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, serta sistematika penulisan laporan.

BAB II KAJIAN PUSTAKA

Pada bab akan menjelaskan tentang konsep dan teori yang mendukung dalam penyusunan penelitian ini. Serta landasan teori dan penelitian sebelumnya yang akan digunakan dalam mengatasi permasalahan.

BAB III METODE USULAN

Pada bab akan menjelaskan terkait dengan rancangan sistem yang akan dibangun, tertera adanya arsitektur sistem, data penelitian, flowchart metode yang akan digunakan, usecase sistem, serta rencana pengujian.

BAB IV HASIL PENELITIAN

Pada bagian ini menjelaskan tentang implementasi sistem yang dibuat berdasarkan hasil rancangan sebelumnya beserta hasil pengujian.

BAB V PENUTUP

Pada bab ini berisi kesimpulan dari implementasi sistem yang telah dibuat serta saran untuk pengembangan sistem.