

## **BAB 1**

### **PENDAHULUAN**

#### **1.1. LATAR BELAKANG**

Jantung merupakan salah satu organ yang memiliki peranan yang sangat penting bagi seorang manusia. Jantung bertanggung jawab untuk memompa dan mengedarkan darah ke seluruh tubuh. Jantung yang tidak sehat sangat mempengaruhi fungsi organ tubuh lainnya. Penyakit jantung koroner adalah penyempitan atau tersumbatnya pembuluh darah arteri jantung yang disebut pembuluh darah koroner. Sebagaimana halnya organ tubuh lain, jantung pun memerlukan zat makanan dan oksigen agar dapat memompa darah ke seluruh tubuh, jantung akan bekerja baik jika terdapat keseimbangan antara pasokan dan pengeluaran. Jika pembuluh darah koroner tersumbat atau menyempit, maka pasokan darah ke jantung akan berkurang, sehingga terjadi ketidakseimbangan antarakebutuhan dan pasokan zat makanan dan oksigen, makin besar persentase penyempitan pembuluh koroner makin berkurang aliran darah ke jantung, akibatnya timbulah nyeri dada(Saragih, n.d.). Penyumbatan pada pembuluh darah jantung adalah penyebab utama penyakit ini(Sitanggang et al. 2022). Penyumbatan ini terjadi karena penumpukan lemak, kolesterol, dan pembentukan plak di dinding arteri. Plak ini dapat pecah dan menyebabkan gumpalan darah, yang menghambat aliran darah ke jantung, menyebabkan kerusakan otot jantung atau bahkan kematian karena kekurangan oksigen dan nutrisi yang dibawa oleh darah.

Masalah yang dihadapi saat ini adalah kelemahan dalam memprediksi penyakit jantung dan mengetahui apakah seseorang menderita penyakit tersebut. Dibutuhkan teknik yang lebih efisien dan sesuai untuk membuat prediksi yang akurat tentang kemungkinan penyakit jantung. Penyakit jantung sering dianggap sebagai penyakit yang muncul tiba-tiba tanpa data kesehatan sebelumnya, menurut *Harvard Health Publishing*. Penelitian sebelumnya yang dilakukan oleh Sepharni dengan judul klasifikasi penyakit jantung dengan menggunakan algoritma C4.5, hasil penelitian yang telah dilaksanakan menggunakan algoritma C4.5 diterapkan menggunakan *k-fold cross validation* dengan jumlah  $k = 10$ , makan dari itu

pengujian ini memiliki 10 skenario pengujian dan 10 skenario pembagian dataset (Sepharni, Hendrawan, and Rozikin 2022). Pengujian 10 skenario tersebut mendapatkan Accuracy yang baik.

Metode data mining yang digunakan dalam penelitian ini adalah *metode K-Nearest Neighbor* (KNN). Algoritma KNN merupakan algoritma yang umum digunakan untuk klasifikasi, namun dapat juga digunakan untuk estimasi dan prediksi. Metode *K-Nearest Neighbors* merupakan model pembelajaran sederhana yang mempertahankan data pelatihan, memungkinkan klasifikasi data tidak terarah dibandingkan dengan informasi pelatihan yang disimpan (Azis et al. 2020). Pada algoritma *K-Nearest Neighbor* (KNN), data berdimensi q, jarak dari data tersebut ke data yang lain dapat dihitung. Nilai jarak inilah yang ditetapkan sebagai nilai kedekatan atau kemiripan antara data uji dengan data latih (Herman, Widiyanto, and Ernawati 2020). Nilai K pada *K-Nearest Neighbor* (KNN) berarti K-data terdekat dari data uji (Argina 2020). Telah banyak penelitian tentang prediksi dengan memakai algoritma *K-Nearest Neighbor* (KNN) salah satunya penelitian yang di lakukan oleh Esty Purwaningsih dan Ela Nurelasari pada tahun 2021, yaitu Penerapan *K-Nearest Neighbor* Untuk Klasifikasi Tingkat Kelulusan Pada Siswa. *K-Nearest Neighbor* (KNN), dataset yang digunakan mengandung 104 data dan 5 atribut (Purwaningsih and Nurelasari 2021). Hasil yang di peroleh menyatakan bahwa pengujian yang dilakukan dengan menggunakan *K-Nearest Neighbor* (KNN) mampu melakukan prediksi dengan baik. Sehingga algoritma tersebut cocok digunakan untuk prediksi tingkat kelulusan pada siswa.

Sistem ini menggunakan algoritma *K-Nearest Neighbor* pada *hearth disease* dataset yang diperoleh dari situs UCI *Machine Learning* untuk mengetahui risiko penyakit jantung berdasarkan atribut yang ada. Hasil algoritma tersebut akan diterapkan kedalam aplikasi sehingga pengguna dapat mengetahui risiko penyakit jantung. Tujuan penelitian ini adalah untuk melihat efektifitas algoritma *K-Nearest Neighbor* dalam melakukan prediksi penyakit jantung. Dalam hal ini Accuracy yang akan dibuat menjadi acuan dalam penilaian performa algoritma dalam prediksi penyakit jantung.

## 1.2. RUMUSAN MASALAH

Adapun rumusan permasalahan yang dapat disusun dalam penelitian ini adalah bagaimana cara memprediksi risiko penyakit jantung dengan menggunakan Algoritme *K-Nearest Neighbors* ?

## 1.3. BATASAN MASALAH

Pada kasus penelitian ini, dibutuhkan batasan masalah agar tetap fokus, terstruktur dan terarah. Berikut batasan-batasan masalah pada penelitian ini:

1. Dataset dalam penelitian ini adalah data *heart disease dataset* yang diperoleh dari situs repositori *UCI Machine Learning* dengan 100 *record* dan 10 atribut dengan 1 kelas.
2. Atribut yang digunakan mencakup *Age, Sex, CPT, Restbps, Chol, Fbs, Restecg, Maxhr, Exang, Oldpeak*.
3. Kriteria input yang dimasukan sesuai dengan atribut yang diambil dari data *UCI Machine Learning dataset*.

## 1.4. TUJUAN PENELITIAN

Tujuan yang ingin dicapai dalam penelitian ini adalah untuk mengetahui risiko pada penyakit jantung menggunakan Algoritma *K-Nearest Neighbor*.

## 1.5. MANFAAT PENELITIAN

Manfaat Dengan adanya Penelitian ini nantinya akan mempermudah atau membantu pengguna dalam melakukan prediksi pada risiko penyakit jantung.

## 1.6. METODOLOGI PENELITIAN

Beberapa tahapan untuk mendapatkan data dan informasi terkait penelitian :

1. Studi Literatur  
Melakukan pendalaman materi yang berkaitan dengan pengelompokan menggunakan Algoritme *K-Nearest Neighbors*.
2. Pengumpulan Data

Pada tahap ini dilakukan pengumpulan data. Data yang digunakan dalam penelitian adalah data *heart disease dataset* yang diperoleh dari situs repositori *UCI Machine Learning*.

### 3. Pengolahan Data

Mengolah data yang didapat dari *UCI Machine Learning*, lalu data tersebut diolah dengan Algoritme *K-Nearest Neighbors* yang dihitung secara manual untuk mengelompokkan *data heart disease* berdasarkan tingkat risiko.

### 4. Perancangan Sistem

Perancangan meliputi analisis kebutuhan sistem, bahasa pemrograman dan arsitektur yang akan digunakan.

### 5. Implementasi dan Pengujian

Menerapkan desain yang telah dirancang ke dalam perangkat sistem.

### 6. Penulisan Laporan

Proses ini dilakukan dengan mendokumentasikan seluruh perubahan serta hasil yang diteliti.

## 1.7. SISTEMATIKA PENULISAN

Adapun sistematika penulisan skripsi ini adalah sebagai berikut:

### BAB I : PENDAHULUAN

Pada bab ini menjelaskan tentang latar belakang, rumusan masalah, batasan masalah, tujuan dan manfaat, metodologi penelitian, serta sistematika penulisan laporan.

### BAB II : TINJAUAN PUSTAKA

Pada bab ini menjelaskan tentang dasar teoritis yang menjadi landasan dalam pelaksanaan sebuah skripsi.

### BAB III : ANALISIS DAN PERANCANGAN SISTEM

Pada bab ini menjelaskan tentang perancangan sistem prediksi penyakit jantung menggunakan metode *K-Nearest Neighbors*.

### BAB IV : IMPLEMENTASI DAN HASIL PENGUJIAN

Pada bab ini menjelaskan tentang hasil dari implementasi analisa cara bekerja program yang dibangun.

#### BAB V : KESIMPULAN DAN SARAN

Pada bab ini menjelaskan tentang kesimpulan dan saran dari keseluruhan bahasan.

#### DAFTAR PUSTAKA

Pada bab ini terdapat semua referensi-referensi yang telah didapat dari buku, jurnal, internet, dan lain sebagainya.

