BAB I
PENDAHULUAN

1.1. Latar Belakang

Penyakit kardiovaskular (PKV) menurut definisi dari *World Health Organization* (*WHO)* dalam pusdatin Kemenkes RI 2014 yaitu penyakit yang disebabkan oleh gangguan fungsi jantung dan pembuluh darah, seperti penyakit jantung koroner, hipertensi, stroke dan penyakit gagal jantung. Setiap tahunnya lebih dari 36 juta orang meninggal karena Penyakit Tidak Menular (PTM) 63% dari semua kematian. Secara global PTM menjadi pemicu kematian nomor satu setiap tahunnya. Kematian dini yang disebabkan oleh penyakit jantung terjadi sebesar 4% di negara berpendapatan tinggi dan 42% terjadi di negara berpendapatan rendah (Kemenkes, 2014). Data *WHO* mengatakan, lebih dari 17 juta manusia di dunia meninggal disebabkan penyakit jantung dan pembuluh darah (Kemenkes, 2019).

Penyakit gagal jantung masuk ke kategori penyakit kardiovaskular. Gagal jantung merupakan suatu keadaan ketika jantung tidak dapat memompa darah yang mencukupi untuk kebutuhan tubuh. Penderita gagal jantung merupakan pasien paling banyak di rumah sakit. Prevalensi gagal jantung di Amerika dan Eropa sekitar 1-2%. Di Indonesia belum ada data epidemiologi untuk gagal jantung, tetapi pada survei kesehatan nasional 2003 ditemukan bahwa penyakit sistem sirkulasi merupakan penyebab kematian utama di Indonesia (26,4%) dan pada profil Kesehatan Indonesia tahun 2003 disebutkan bahwa penyakit jantung berada pada urutan kedelapan (Sari, Yonata, Haryadi, & Swadharma , 2016)

Mengingat berharganya organ vital seperti jantung, memprediksi gagal jantung telah menjadi prioritas bagi dokter dan tenaga medis, namun sampai saat ini prediksi perihal gagal jantung dalam praktik klinis biasanya gagal mendapatkan akurasi yang tinggi (Chicco & Jurman, 2020).

Pendekatan *data mining* menjadi sangat krusial dalam aspek Kesehatan untuk mengambil keputusan berlandaskan data klinis yang besar (Fiandra, Defit, & Yuhandri, 2017). *Data mining* yaitu proses mencari pola atau informasi menarik dalam data terpilih dengan memakai teknik atau metode tertentu (Mardi, 2017).

Diterapkannya *data mining* diharapkan dapat menjadi suatu informasi untuk mencegah atau mengurangi penderita penyakit gagal jantung di Indonesia dan dunia, sehingga angka penderita penyakit gagal jantung dapat menurun.

Algoritma *K-Nearest Neighbor* (*KNN*) yakni algoritma klasifikasi berlandaskan kedekatan jarak suatu data dengan data yang lain. 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. Nilai K pada *K-Nearest Neighbor* (*KNN*) berarti K-data terdekat dari data uji (Lestari, 2014).

Telah banyak penelitian tentang prediksi dengan memakai algoritma *K-Nearest Neighbor* (*KNN*) salah satunya penelitian yang di lakukan oleh Gultom dan Simanjuntak (2020), yaitu prediksi tingkat kelancaran pembayaran kredit bank. Algoritma yang digunakan yaitu *naïve bayes dan K-Nearest Neighbor* (*KNN*), dataset yang digunakan mengandung 363 data dan 14 atribut. Hasil yang di peroleh menyatakan bahwa pengujian yang dilakukan dengan menggabungkan algoritma *Naïve Bayes* dan *K-Nearest Neighbor* (*KNN*) mampu melakukan prediksi dengan baik dengan akurasi 80%. Sehingga gabungan kedua algoritma tersebut cocok diginakan untuk prediksi kelancaran pembayaran kredit pada bank, dan dapat membantu petugas bank dalam menentukan nasabahnya.

Selain itu ada juga penelitian Lestari (2014), penelitian ini meneliti tentang menditeksi penyakit jantung. Algoritma yang digunakan dalam penelitian ini yaitu *K-Nearest Neighbor* (*KNN*), dataset yang digunakan mengandung 100 data pasien penyakit jantung dan 13 atribut. Hasil dari penelitian ini yaitu, algoritma *K-Nearest Neighbor* (*KNN*) dengan memakai *cofusion matrix* dan kurva ROC, diperoleh nilai akurasi 70% dan termasuk klasifikasi baik sebab memiliki nilai AUC 0.875.

Teknik-teknik penelitian yang telah dilakukan Gultom dengan Simanjuntak (2020) dan Lestari (2014) Beserta penjelasan yang sudah di paparkan, maka pada penelitian ini akan mengambil judul "Penerapan Algoritma *K-Nearest Neighbour* (*KNN*) Untuk Prediksi Kematian Akibat Penyakit Gagal Jantung".

1.2. Rumusan Masalah

Berdasarkan latar belakang di atas, maka dapat di rumuskan masalah pokok yaitu:

1. Bagaimana menerapkan algoritma *K-Nearest Neighbor* (*KNN*) pada prediksi kematian akibat penyakit gagal jantung?
2. Seberapa besar tingkat akurasi algoritma *K-Nearest Neighbor* (*KNN*) dalam memprediksi kematian akibat penyakit gagal jantung?

1.3. Tujuan Penelitian

Berdasarkan rumusan masalah di atas tujuan penelitian ini yaitu:

1. Menerapkan algoritma *K-Nearest Neighbor* (*KNN*) pada prediksi kematian akibat penyakit gagal jantung.
2. Mengetahui tingkat akurasi algoritma *K-Nearest Neighbor* (*KNN*) dalam memprediksi kematian akibat penyakit gagal jantung.

1.4. Manfaat

1. Manfaat Bagi Akademisi

   Hasil dari penelitian ini diharapkan dapat dijadikan referensi bagi penelitian selanjutnya. Serta mampu mendorong berkembangnya penelitian-penelitian selanjutnya pada bidang *data mining*.
2. Manfaat Bagi Praktisi

   Hasil dari penelitian ini, diharapkan dapat digunakan untuk membantu rumah sakit dalam menilai tingkat keparahan pasien dengan penyakit gagal jantung.