

ABSTRAK

Penyakit jantung disebabkan karena adanya plak di dinding pembuluh darah arteri yang mengakibatkan gangguan aliran darah. Angka kematian akibat penyakit jantung semakin tinggi dikarenakan penderita terlambat melakukan pengobatan. Berdasarkan kondisi tersebut, maka diperlukan diagnosis awal penyakit jantung untuk mengetahui kemungkinan bahwa seseorang menderita penyakit jantung atau tidak. Dengan melakukan diagnosis awal, para penderita penyakit jantung dapat segera mengambil tindakan pengobatan dan meminimalisir risiko kematian akibat penyakit ini. Pada penelitian ini menerapkan algoritma *K-Nearest Neighbor* dengan perhitungan jarak berbasis *Overlap Metric* untuk mendiagnosis penyakit jantung. *Dataset* yang digunakan pada penelitian ini bernama *Heart Failure Prediction* diperoleh dari website <https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction>. *Dataset* tersebut terdiri dari 918 data, 11 atribut, dan 2 kelas sebagai hasil klasifikasi yaitu nilai 1 berarti positif penyakit jantung dan 0 berarti negatif. Penelitian ini melakukan uji coba terhadap beberapa proporsi pembagian data latih dan data uji yaitu 70%:30%, 80%:20%, dan 90%:10%. Berdasarkan analisis terhadap hasil uji coba diperoleh proporsi yang tepat untuk diterapkan pada algoritma *K-Nearest Neighbor* berbasis *Overlap Metric* dalam melakukan diagnosis penyakit jantung adalah 90% data latih dan 10% data uji, dengan rincian nilai performa yaitu *accuracy* sebesar 89%, *precision* senilai 88%, *recall* sebesar 94%, dan *F-score* senilai 91%. Hasil pengujian parameter *k* yaitu angka ganjil antara 1 sampai 49 pada proporsi data 90%:10% didapatkan nilai *k* terbaik untuk diterapkan pada algoritma KNN adalah *k* = 7.

Kata kunci: Penyakit Jantung, Diagnosis, *K-Nearest Neighbor*, *Overlap Metric*

ABSTRACT

Heart disease is caused by the presence of plaque on the walls of the arteries which results in impaired blood flow. The death rate from heart disease is higher because sufferers are late in taking treatment. Based on these conditions, an early diagnosis of heart disease is needed to find out the possibility that a person has heart disease or not. By making an early diagnosis, people with heart disease can immediately take action on treatment and minimize the risk of death from this disease. In this study, we apply the *K-Nearest Neighbor* algorithm with distance calculations based on *Overlap Metric* to diagnose heart disease. The dataset used in this study is called *Heart Failure Prediction* obtained from the website <https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction>. The dataset consists of 918 data, 11 attributes, and 2 classes as a result of classification, namely a value of 1 means positive for heart disease and 0 means negative. This study tested several proportions of the distribution of training data and test data, namely 70%:30%, 80%:20%, and 90%:10%. Based on the analysis of the trial results, it was found that the right proportion to be applied to the *K-Nearest Neighbor* algorithm based on *Overlap Metric* in diagnosing heart disease is 90% training data and 10% test data, with details of performance values, namely *accuracy* of 89%, *precision* of 88 %, *recall* of 94%, and *F-score* of 91%. The results of testing the *k*

parameter, namely an odd number between 1 and 49 in the proportion of data 90%: 10%, the best k value to be applied to the KNN algorithm is $k = 7$.

Keywords: Heart Disease, Diagnosis, K-Nearest Neighbor, Overlap Metric