

Pendekatan Diagnostik untuk Deteksi Penyakit Jantung Menggunakan Random Forest = Diagnostic Approach for Heart Disease Detection Using Random Forest

Tiara Anggraini Gaib, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920543563&lokasi=lokal>

Abstrak

Penyakit jantung, seperti yang didefinisikan oleh World Health Organization (WHO) sebagai kumpulan berbagai gangguan yang memengaruhi kesehatan jantung, merupakan salah satu masalah kesehatan global yang memerlukan deteksi dini dan penanganan yang efektif. Faktor risiko yang menyebabkan penyakit jantung seperti usia, jenis kelamin, nyeri dada (chest pain), tekanan darah (resting blood pressure), kolesterol (cholesterol), kadar gula darah (fasting blood sugar), hasil elektrokardiogram (resting electrocardiogram), detak jantung maksimum yang dicapai (maximum heart rate achieved), keberadaan angina yang diinduksi (exercise-induced angina), depresi segmen ST (oldpeak), bentuk kelengkungan pada kurva (slope of the peak exercise st segment), jumlah pembuluh darah utama yang diwarnai oleh flourosopy (Number of Major Vessels Colored by Fluoroscopy/CA), dan jenis thalassemia (thalassemia), memiliki peran signifikan dalam meningkatkan risiko terjadinya penyakit jantung. Penelitian ini dilakukan menggunakan dataset yang berasal dari Klinik Cleveland, yang terdiri dari 303 entri data. Dataset ini digunakan untuk melakukan deteksi terhadap keberadaan atau ketidakhadiran penyakit jantung berdasarkan sejumlah atribut klinis yang diukur. Atribut-atribut ini, atau fitur-fitur, mencakup berbagai informasi seperti tekanan darah, kadar kolesterol, usia, jenis kelamin, dan lainnya. Tujuan dari penelitian ini adalah untuk melakukan deteksi yang dapat memprediksi penyakit jantung berdasarkan informasi klinis pasien dengan akurasi terbaik. Untuk mencapai tujuan ini, model Random Forest dilatih dan dibandingkan dengan model lain meliputi Naive Bayes, Decision Tree, Logistic Regression, Support Vector Machine (SVM), Neural Network dan XGBoost. Hasil evaluasi menunjukkan bahwa Random Forest Algorithm memiliki akurasi yang paling tinggi, mencapai 96,77%. Ini berarti bahwa model Random Forest mampu memprediksi keberadaan atau ketidakhadiran penyakit jantung dengan tingkat keberhasilan yang sangat tinggi. Sebagai hasilnya, Random Forest dipilih sebagai model yang paling sesuai untuk melakukan deteksi penyakit jantung dalam dataset ini. Model ini diharapkan dapat memberikan kontribusi yang signifikan dalam deteksi dini, pencegahan, dan pengelolaan penyakit jantung secara global.

.....Heart disease, as defined by the World Health Organization (WHO) as a collection of various disorders affecting heart health, is one of the global health issues requiring early detection and effective management. Risk factors contributing to heart disease such as age, gender, chest pain (angina), blood pressure (trestbps), cholesterol (cholesterol), blood sugar levels (fbs), electrocardiogram results (restecg), maximum heart rate achieved (thalach), presence of induced angina (exang), ST segment depression (oldpeak), slope of the ST segment (slope), number of major vessels colored by fluoroscopy (ca), and type of thalassemia (thal), play a significant role in increasing the risk of heart disease. This research was conducted using a dataset obtained from the Cleveland Clinic, consisting of 303 data entries. This dataset was utilized to classify the presence or absence of heart disease based on various measured clinical attributes, including blood pressure, cholesterol levels, age, gender, among others. The aim of this study is to perform detection that can predict heart disease based on patient clinical information with the highest accuracy. To achieve this objective, the

Random Forest model was trained and compared with other models, including Naive Bayes, Decision Tree, Logistic Regression, Support Vector Machine (SVM), Neural Network, and XGBoost. Evaluation results demonstrate that the Random Forest Algorithm achieved the highest accuracy, reaching 96,77%. This implies that the Random Forest model can predict the presence or absence of heart disease with a very high success rate. Consequently, Random Forest was chosen as the most suitable model for classifying heart disease in this dataset. This model is anticipated to significantly contribute to the early detection, prevention, and management of heart disease globally.