

ANALISIS PERBANDINGAN ALGORITMA *RANDOM FOREST* DAN *XGBOOST* UNTUK KLASIFIKASI PENYAKIT KARDIOVASKULAR

ALIF FAQIH

ABSTRAK

Jantung adalah salah satu organ berotot yang berfungsi sebagai alat untuk pemompa oksigen dan darah menuju ke seluruh organ tubuh. Berdasarkan laman resmi dari *World Health Organization* (WHO), terdapat 17,9 juta orang meninggal dunia setiap tahunnya disebabkan oleh penyakit *kardiovaskular* yang menyerang bagian jantung manusia. Maka dari itu dibutuhkan model *machine learning* yang memiliki performa yang baik dan dapat melakukan klasifikasi pada penyakit kardiovaskular dengan cepat agar apabila terdapat orang yang terdeteksi berisiko terkena penyakit kardiovaskular dapat langsung menemui dokter untuk dilakukan pemeriksaan lebih lanjut dan penyakit dapat lebih cepat untuk ditangani. Penelitian ini menggunakan *dataset* Cardiovascular Disease Risk Prediction dengan *Random Forest* dan *XGBoost* sebagai model untuk melakukan klasifikasi. Penelitian ini dilakukan dengan tujuan untuk menemukan model dengan performa terbaik antara *Random Forest* dan *XGBoost*. Hasil dari penelitian ini merupakan evaluasi dari *Random Forest* dan *XGBoost* dalam melakukan klasifikasi penyakit kardiovaskular. Didapatkan performa terbaik pada model *Random Forest* dengan nilai akurasi sebesar 0.95, *precision* sebesar 0.96, dan *recall* sebesar 0.93.

Kata Kunci: Penyakit Kardiovaskular, *Random Forest*, *XGBoost*

COMPARATIVE ANALYSIS OF RANDOM FOREST AND XGBOOST ALGORITHMS FOR CARDIOVASCULAR DISEASE CLASSIFICATION

ALIF FAQIH

ABSTRACT

Heart is a muscular organ that functions as a tool to pump oxygen and blood to all organs of the body. Based on the official website of the World Health Organization (WHO), 17.9 million people die every year due to cardiovascular disease which attacks the human heart. Therefore, we need a machine learning model that has good performance and can classify cardiovascular disease quickly so that if someone is detected to be at risk of developing cardiovascular disease, they can immediately see a doctor for further examination and the disease can be treated more quickly. This research uses the Cardiovascular Disease Risk Prediction. *dataset* with Random Forest and XGBoost as models for classification. This research was conducted with the aim of finding the model with the best performance between Random Forest and XGBoost. The results of this study are an evaluation of Random Forest and XGBoost in classifying cardiovascular disease. The best performance was obtained from the Random Forest model with an accuracy value of 0.95, precision of 0.96, and recall of 0.93.

Keyword : Cardiovascular Disease, Random Forest, XGBoost