

## ABSTRAK

Penyakit kardiovaskular (CVD) merupakan salah satu penyebab utama kematian di seluruh dunia, terutama di negara berpenghasilan rendah dan menengah. Untuk mengurangi angka kematian dini, prediksi risiko penyakit ini sangat penting. Penelitian ini berfokus pada algoritma *Random Forest*, sebuah metode *Machine Learning* yang telah terbukti efektif dalam memproses data medis kompleks, untuk meningkatkan akurasi prediksi kelangsungan hidup pasien dengan penyakit kardiovaskular (CVD). Penelitian ini akan membangun model untuk memprediksi waktu kelangsungan hidup pasien yang menderita penyakit kardiovaskular (CVD) selama 10 tahun kedepan. Hasil penelitian menunjukkan bahwa nilai MAE sebesar 5.37, yang menunjukkan bahwa terdapat perbedaan rata-rata sebesar 5.37 tahun antara prediksi model dengan nilai sebenarnya. Nilai C-index yang didapatkan adalah 0.5, menunjukkan bahwa model ini belum mampu secara efektif merangking prediksi kelangsungan hidup pasien. Maka diperlukan evaluasi dan pengembangan lebih lanjut untuk meningkatkan akurasi dan kehandalan prediksi, khususnya dalam merangking prediksi secara lebih tepat. Penelitian ini diharapkan dapat meningkatkan perencanaan perawatan yang lebih baik bagi pasien dengan penyakit jantung, serta memberikan prediksi yang akurat dan handal tentang komponen utama yang mempengaruhi kelangsungan hidup pasien.

Kata kunci : CVD, *Machine Learning*, *Random Forest*, *Mean Absolute Error (MAE)*, *C-Index*

## ABSTRACT

### ***IMPLEMENTATION OF THE RANDOM FOREST ALGORITHM TO PREDICT SURVIVAL OF PATIENTS WITH CARDIOVASCULAR DISEASE (CVD)***

**By:**

**MEILIN NEDA FARERA**

**2011050078**

**[meilinnedafarera@gmail.com](mailto:meilinnedafarera@gmail.com)**

Cardiovascular disease (CVD) is one of the leading causes of death worldwide, particularly in low- and middle-income countries. To reduce premature mortality, predicting the risk of this disease is crucial. This study focused on the Random Forest algorithm, a machine learning method that had proven effective in processing complex medical data, to improve the accuracy of predicting survival in patients with cardiovascular disease (CVD). The study developed a model to predict the survival time of patients with cardiovascular disease (CVD) over the next 10 years. The result showed that the model achieved a Mean Absolute Error (MAE) value of 5.37, indicating that there was an average difference of 5.37 years between the model's predictions and the actual values. The C-index value obtained was 0.5, indicating that this model had not yet effectively ranked the patients' survival predictions. Further evaluation and development are needed to improve the accuracy and reliability of the predictions, particularly in ranking them more precisely. This research was expected to contribute to better care planning for heart disease patients and provide accurate and reliable predictions regarding the key factors affecting patients' survival.

***Keywords:*** CVD, Machine Learning, Random Forest, Mean Absolute Error (MAE), C-Index