

INTISARI

MODEL PREDIKSI TINGKAT BERTAHAN HIDUP PASIEN GAGAL JANTUNG MENGGUNAKAN AUTOMATED MACHINE LEARNING

Oleh

Abiyyu Barraq

17/412629/PA/17948

Kelompok penyakit kardiovaskular menjadi penyebab nomor satu kematian di dunia, termasuk salah satu di dalamnya adalah gagal jantung. Dengan banyaknya praktis klinis dan observasi yang dilakukan terkait gagal jantung, membuat koleksi data terkait gagal jantung makin kaya, sehingga dengan kumpulan data tersebut para peneliti banyak melakukan penelitian termasuk penelitian tingkat bertahan hidup pasien gagal jantung dengan menggunakan metode *Machine Learning (ML)*. Sayangnya, penggunaan metode ML dalam dunia medis untuk pengolahan data masih sangat terbatas karena pada implementasinya tenaga medis memerlukan seseorang yang memiliki keahlian baik di bidang data maupun ML.

Dalam berbagai penelitian, metode ML terbukti memiliki potensi untuk digunakan pada bidang medis, khususnya pada kasus gagal jantung. Hal ini mengindikasikan bahwa dengan penggunaan ML terutama untuk kasus prediksi dapat membantu para tenaga medis untuk mengambil keputusan. Penelitian ini akan membuat model *Automated Machine Learning (AutoML)* dengan berbasis *Random Search* yang bertujuan untuk mengotomasi sistem kerja dari metode ML untuk memprediksi tingkat bertahan hidup pasien gagal jantung. Penelitian ini juga akan mengevaluasi model AutoML yang dibuat dengan algoritma ML yang biasa digunakan untuk melakukan klasifikasi.

Hasil penelitian menunjukkan bahwa model AutoML memiliki performa lebih baik dari beberapa model ML standar. Model AutoML yang dikembangkan mencapai nilai akurasi 0,8555 pada data uji, yang sebanding dengan performa terbaik dari model ML standar seperti XGBoost dan SVM. Model AutoML juga berhasil mengoptimalkan parameter secara otomatis dan lebih efektif, sehingga mampu menghasilkan model dengan performa yang lebih baik dibandingkan model ML standar yang menggunakan parameter default dari *Scikit-learn*. Hal ini terlihat dari hasil AutoML dengan hasil Random Forest sebagai estimator terbaik, memberikan performa yang lebih baik daripada Random Forest pada model ML standar.

Kata Kunci: Pembelajaran Mesin, Rekam Medis, *Cardiovascular heart diseases*, Heart failure, *Automated Machine Learning*, Klasifikasi, Optimasi *Hyperparameter*

ABSTRACT

SURVIVAL RATE PREDICTION MODEL FOR HEART FAILURE PATIENT USING AUTOMATED MACHINE LEARNING

By

Abiyyu Barraq

17/412629/PA/17948

Cardiovascular diseases are the leading cause of death worldwide, with heart failure being one of the major contributors. The abundance of clinical practices and observations related to heart failure has resulted in a rich collection of data, enabling researchers to conduct numerous studies, including those on the survival rates of heart failure patients using Machine Learning (ML) methods. Unfortunately, the application of ML methods in the medical field for data processing is still very limited, as it requires medical personnel to have expertise in both data science and ML.

Various studies have demonstrated that ML methods have the potential to be used in the medical field, particularly in cases of heart failure. This indicates that the use of ML, especially for predictive cases, can assist medical personnel in making decisions. This study aims to develop an Automated Machine Learning (AutoML) model based on Random Search, with the objective of automating the workflow of ML methods to predict the survival rates of heart failure patients. The study will also evaluate the AutoML model created with commonly used ML algorithms for classification.

The research findings show that the AutoML model outperforms several standard ML models. The developed AutoML model achieved an accuracy score of 0.8555 on the test data, comparable to the best performance of standard ML models such as XGBoost and SVM. The AutoML model also successfully optimized parameters automatically and more effectively, resulting in a model with better performance compared to standard ML models using default parameters from Scikit-learn. This is evident from the AutoML results, with Random Forest as the best estimator, providing better performance than Random Forest in standard ML models.

Keywords: Machine Learning, Medical Record, *Cardiovascular heart diseases*, Heart failure, *Automated Machine Learning*, Classification, *Hyperparameter Optimization*