

Analisis Kinerja Model XGBoost dalam Mendeteksi Congestive Heart Failure pada Short-term Heart Rate Variability = Analyzing the Performance of XGBoost Model to Detect Congestive Heart Failure on Short-term Heart Rate Variability

Gregorino Al Josan, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920541684&lokasi=lokal>

Abstrak

Cardiovascular diseases (CVD) merupakan salah satu penyebab utama kematian di dunia. WHO memperkirakan angka 17,9 juta kematian pada tahun 2021 disebabkan oleh CVD. Di Indonesia sendiri, prevalensi penyakit jantung mencapai angka 1,5% atau sekitar 2,7 juta orang pada tahun 2018. CVD mencakup berbagai macam jenis penyakit jantung. Salah satu tipe penyakit jantung tersebut adalah congestive heart failure. Congestive heart failure (CHF) adalah kondisi dimana jantung tidak dapat memompa darah yang cukup ke seluruh bagian tubuh. CHF dapat terjadi dikarenakan melemahnya kemampuan otot jantung untuk memompa darah sehingga mempengaruhi heart rate atau detak jantung manusia. Heart rate dapat direpresentasikan menggunakan sinyal yang dapat diukur menggunakan alat rekaman electrocardiogram (ECG/EKG). EKG adalah rekaman aktivitas elektrik jantung yang ditangkap melalui bagian permukaan tubuh. Heart rate variability (HRV) diketahui berkorelasi dengan berbagai penyakit jantung dan salah satunya adalah CHF. Dengan berkembangnya teknologi, terdapat beberapa penelitian mengenai implementasi artificial intelligence (AI) untuk mendeteksi keberadaan CHF menggunakan model machine learning dan HRV sebagai fitur bagi model. Pada penelitian ini, akan dibangun dan dievaluasi kinerja model XGBoost untuk mendeteksi eksistensi penyakit CHF pada short-term HRV dari rekaman EKG 5 menit. Dataset yang digunakan berasal dari empat database yang berbeda yang diambil dari situs PhysioNet, yaitu NSRDB dan NSR2DB sebagai kelas sehat dan CHFDB dan CHF2DB sebagai kelas CHF. Masing-masing database memiliki rekaman long-term EKG. Seluruh rekaman tersebut dilakukan segmentasi selama 5 menit pada 2 jam pertama rekaman. Dari hasil segmentasi rekaman 5 menit tersebut akan dihitung nilai HRV yang akan menjadi fitur bagi model XGBoost. XGBoost dilatih menggunakan kombinasi teknik Grid Search dan K-Fold Cross Validation dengan nilai $\delta_{3/4} = 10$. Terdapat 4 metrik yang dijadikan objektif optimisasi Grid Search, yaitu akurasi, sensitivitas, spesifisitas, dan skor AUC. XGBoost yang dilatih dengan mengoptimasi akurasi berhasil mencapai nilai akurasi sebesar 0,954, sensitivitas sebesar 0,935, spesifisitas sebesar 0,96, dan skor AUC sebesar 0,947. XGBoost yang dilatih dengan mengoptimasi sensitivitas berhasil mencapai nilai akurasi sebesar 0,966, sensitivitas sebesar 0,977, spesifisitas sebesar 0,963, dan skor AUC sebesar 0,97. XGBoost yang dilatih dengan mengoptimasi spesifisitas berhasil mencapai nilai akurasi sebesar 0,962, sensitivitas sebesar 0,931, spesifisitas sebesar 0,971, dan skor AUC sebesar 0,951. Kemudian XGBoost yang dilatih dengan mengoptimasi skor AUC berhasil mencapai nilai akurasi sebesar 0,955, sensitivitas sebesar 0,935, spesifisitas sebesar 0,962, dan skor AUC sebesar 0,948.

.....Cardiovascular diseases (CVD) is one of the major causes of death in the world. WHO estimated that 17.9 million of deaths during 2021 are caused by CVD. In Indonesia alone, the prevalence of heart diseases reached 1.5% or around 2,7 million people in 2018. CVD consists of various types of heart disease. Congestive heart failure is one of them. Congestive heart failure (CHF) is a condition where the heart cannot

pump enough blood for the entire body. CHF can occur due to a weakening of the heart muscle's ability to pump blood, thereby affecting the human heart rate. Heart rate can be represented using signal that can be measured using electrocardiogram (ECG/EKG) recording. EKG is a recording of the heart's electrical activity captured through the surface of the body. Heart rate variability (HRV) have been known to be correlated with various heart diseases with CHF is one of it. With the advance of technology, there have been various research regarding the implementation of artificial intelligence (AI) to detect the presence of CHF using machine learning model and HRV as features for the model. In this research, we built and evaluated the performance of XGBoost model to detect the existence of CHF on short-term HRV from 5 minutes EKG recording. The dataset came from four different databases that can be accessed from PhysioNet website. Those are NSRDB and NSR2DB datasets to represent healthy class and CHFDB and CHF2DB to represent CHF class. Each database contains long-term EKG. All records are segmented by 5 minutes on the first 2 hours of the recording. HRV metrics are calculated from those 5 minutes segments to become features for the XGBoost model. XGBoost was trained using a combination of Grid Search and K-Fold Cross Validation techniques with $k = 10$. There are 4 metrics that become the objective scoring function for the Grid Search. Those are accuracy, sensitivity, specificity, and AUC score. XGBoost trained to optimize accuracy managed to achieve 0.954 accuracy, 0.935 sensitivity, 0.96 specificity, and 0.947 AUC score. XGBoost trained to optimize sensitivity managed to achieve 0.966 accuracy, 0.977 sensitivity, 0.963 specificity, and 0.97 AUC score. XGBoost trained to optimize specificity managed to achieve 0.962 accuracy, 0.931 sensitivity, 0.971 specificity, and 0.951 AUC score. Lastly, XGBoost trained to optimize AUC score managed to achieve 0.955 accuracy, 0.935 sensitivity, 0.962 specificity, and 0.948 AUC score.