

ANALISIS ALGORITMA GAUSSIAN NAIVE BAYES TERHADAP KLASIFIKASI DATA PASIEN PENDERITA GAGAL JANTUNG

Prodi Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember

Quswatun Hasanah, Hardian Oktavianto, Yeni Dwi Rahayu

Abstrak

Penelitian ini menjelaskan penggunaan algoritma Gaussian Naive Bayes terhadap klasifikasi data pasien penderita gagal jantung. Dataset yang digunakan dalam penelitian ini berasal dari *Kaggle* dengan jumlah variabel sebanyak 30 yang diperoleh dari data pasien penderita gagal jantung sebanyak 1000 orang. Pada tahap skenario uji coba dalam penelitian ini menggunakan *Cross Fold Validation* dengan nilai $k=2, 4, 5$, dan 10 . Penentuan kriteria hasil klasifikasi terhadap output data asli (nilai aktual) menggunakan *confusion matrix*. Dari implementasi yang dilakukan menggunakan algoritma Gaussian Naive Bayes terhadap data pasien penderita gagal jantung diperoleh hasil validasi terbaik dengan nilai akurasi tertinggi didapatkan pada *Kfold 4* yaitu $59,4\%$, presisi tertinggi didapatkan pada *Kfold 4* yaitu $55,89\%$, sedangkan recall tertinggi didapatkan pada *Kfold 2* yaitu $86,52\%$.

Kata Kunci : Klasifikasi, Gaussian Naive Bayes, Gagal Jantung, *Cross Validation*.

GAUSSIAN NAIVE BAYES ALGORITHM ANALYSIS OF DATA CLASSIFICATION OF HEART FAILURE PATIENTS

*Informatics Engineering Study Program, Faculty of Engineering, University of
Muhammadiyah Jember*

Quswatun Hasanah, Hardian Oktavianto, Yeni Dwi Rahayu

Abstract

This study describes the use of the Gaussian Naive Bayes algorithm for data classification of heart failure patients. The dataset used in this study came from Kaggle with a number of variables as many as 30 data obtained from 1000 patients with heart failure. At the stage of the trial scenario in this study using Cross Fold Validation with values of $k = 2, 4, 5$, and 10. Determining the classification results criteria for the original output data (actual value) using a confusion matrix. From the implementation carried out using the Gaussian Naive Bayes algorithm on patient data failed to obtain the best validation results with the highest accuracy value obtained at Kfold 4, that is 59.4%, the highest precision was obtained at Kfold 4, that is 55.89%, while the highest recall was obtained at Kfold 2 that is 86.52% .

Keyword : *Classification, Gaussian Naive Bayes, Heart Failure, Cross Validation.*