

ABSTRAK

Nama : Arikza Syahrin
NIM : 41519010070
Pembimbing TA : Wawan Gunawan, S. Kom, MT
Judul Laporan Skripsi : PERBANDINGAN ALGORITMA C5.0 DAN NAÏVE
BAYES TERHADAP KLASIFIKASI GANGGUAN IRAMA
JANTUNG

Atrial fibrilasi (AF) merupakan salah satu kondisi dimana serambi (atrium) jantung berdenyut dengan cepat dan tidak beraturan. Beberapa kondisi medis yang dapat menyebabkan terjadinya AF antara lain tekanan darah tinggi, penyakit paru-paru kelainan jantung bawaan, serta infeksi virus. AF merupakan salah satu gangguan irama jantung yang dapat meningkatkan 4 sampai 5 kali terjadinya stroke serta komplikasi lainnya. Jumlah kejadian yang terkait dengan AF dapat mencapai 5 sampai 10 kejadian tiap 1.000 orang dengan umur di atas 65 tahun. Penelitian ini berfokus pada klasifikasi data penyakit AF dengan menggunakan dua algoritma, yaitu C5.0 dan Naïve Bayes dan hasil klasifikasi dari kedua algoritma tersebut akan dibandingkan. Sebelum dilakukan proses klasifikasi, dilakukan tahap *pre-processing* terlebih dahulu terhadap data yang tersedia. Pada tahap *pre-processing* ini akan dilakukan *replace missing value* atau mengisi nilai yang kosong dengan nilai rata-rata (*mean*) dan normalisasi data. Untuk metode split data, hasil akurasi tertinggi yang diperoleh algoritma Naïve Bayes adalah pada percobaan split data 90:10 dengan nilai akurasi sebesar 81,25%. Hasil akurasi tertinggi yang diperoleh algoritma C5.0 adalah pada percobaan split data 70:30 dengan nilai akurasi sebesar 79,94%. Untuk metode *K-Cross Fold Validation*, algoritma C5.0 memperoleh hasil terbaik untuk nilai $K = 7$ dengan nilai akurasi sebesar 82,39%. Sedangkan untuk algoritma Naïve Bayes memperoleh hasil terbaik untuk nilai $K = 3$ dengan nilai akurasi sebesar 79,78%.

Kata kunci: Naïve Bayes, C5.0, klasifikasi, Atrial Fibrilasi, *K-Fold Cross Validation*.

ABSTRACT

Name : Arikza Syahrin
Student Number : 41519010070
Supervisor : Wawan Gunawan, S. Kom, MT
Thesis title : PERBANDINGAN ALGORITMA C5.0 DAN NAÏVE
BAYES TERHADAP KLASIFIKASI GANGGUAN IRAMA
JANTUNG

Atrial fibrillation (AF) is a condition where the chambers (atria) of the heart beat rapidly and irregularly. Some medical conditions that can cause AF include high blood pressure, lung disease congenital heart defects, and viral infections. AF is one of the heart rhythm disorders that can increase the occurrence of stroke and other complications by 4 to 5 times. The number of events associated with AF can reach 5 to 10 events per 1,000 people over the age of 65. This research focuses on the classification of AF disease data using two algorithms, namely C5.0 and Naïve Bayes and the classification results of the two algorithms will be compared. Before the classification process is carried out, a pre-processing stage is carried out first on the available data. In this preprocessing stage, we will replace missing values or fill in empty values with the average value (mean) and normalize the data. For the split data method, the highest accuracy result obtained by the Naïve Bayes algorithm is in the 90:10 split data experiment with an accuracy value of 81.25%. The highest accuracy result obtained by the C5.0 algorithm is in the 70:30 split data experiment with an accuracy value of 79.94%. For the K-Cross Fold Validation method, the C5.0 algorithm obtained the best results for the $K = 7$ value with an accuracy value of 82.39%. Meanwhile, the Naïve Bayes algorithm obtained the best results for the $K = 3$ value with an accuracy value of 79.78%.

Keywords: Naïve Bayes, C5.0, classification, Atrial Fibrillation, K-Fold Cross Validation.