

**ANALISIS METODE *K NEAREST NEIGHBOR* TERHADAP
KLASIFIKASI DATA PASIEN PENDERITA GAGAL JANTUNG**

Jurusan Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember

Umi Nur Hidayah, Hardian Octavianto, Lutfi Ali Muharom

ABSTRAK

Penelitian ini menjelaskan penggunaan metode *K Nearest Neighbor* terhadap klasifikasi data pasien penderita gagal jantung. Dalam penelitian ini dilakukan beberapa tahapan dan metode. Pada tahapan *preprocessing* penelitian ini menggunakan metode *MinMax Scaler* untuk normalisasi serta metode *LogX* untuk mengurangi tingkat *skewness* dan *outlier*. Pada tahap skenario uji dalam penelitian ini menggunakan metode *Cross Fold Validation* dengan nilai $k = 3, 5, 7$ dan 9 dan sebagai pembeda dalam penelitian ini dilakukan dua proses yang berbeda yaitu proses implementasi K-NN menggunakan Teknik *SMOTE* sebelumnya dan implementasi K-NN tanpa Teknik *SMOTE*. Pada pengukuran hasil klasifikasi berfokus pada tingkat akurasi dan presisi. Dalam penelitian ini data yang digunakan berjumlah 299 pasien yang berasal dari *UCI Machine Learning*. Pengukuran jarak pada K-NN dalam penelitian ini menggunakan *vector Euclidean Distance* dengan nilai ketetanggan 3, 5, 7 dan 9. Dari implementasi yang dilakukan diperoleh hasil yaitu klasifikasi *K Nearest Neighbor* terhadap data pasien gagal jantung tanpa *SMOTE* diperoleh hasil validasi terbaik terdapat pada *Nearest Neighbor $k = 9$* dengan rata-rata akurasi sebesar 71.59% dan akurasi tertinggi yaitu sebesar 83.33% pada *fold* kedua. Sedangkan pada data uji akurasi tertinggi diperoleh pada *Nearest Neighbor $k = 3$* dengan nilai akurasi 71.66%. klasifikasi *K Nearest Neighbor* terhadap data pasien gagal jantung dengan menggunakan *SMOTE* diperoleh hasil validasi terbaik terdapat pada *Nearest Neighbor $k = 3$* dengan rata-rata akurasi sebesar 87.88% dengan akurasi tertinggi yaitu sebesar 80.14% pada *fold* keenam dan delapan. Sedangkan pada data uji akurasi tertinggi diperoleh pada *Nearest Neighbor $k = 3$* dengan nilai akurasi 63.33%.

Keyword : Klasifikasi, K-NN, Gagal Jantung, *SMOTE*, *Cross Fold Validation*.

ANALYSIS ON DATA CLASSIFICATION OF PATIENTS WITH HEART FAILURE USING K NEAREST NEIGHBOR METHOD

Department of Informatics, Faculty of Engineering, University of Muhammadiyah Jember

Umi Nur Hidayah , Hardian Octavianto , Lutfi Ali Muharom

ABSTRACT

This study describes the use of the K Nearest Neighbor method for data classification of patients with heart failure . In this research , several stages and methods were carried out . At the preprocessing stage of this study using the MinMax Scaler method for normalization and the LogX method to reduce the level of skewness and outliers . At stage s k enario test in research is using methods Cross Fold Validation with value $k = 3, 5, 7$ and 9 and as a differentiator in the research is carried out two processes are different , namely the implementation of the K-NN using techniques SMOTE earlier and implementation K- NN without SMOTE Technique . In measuring the classification results focus on the level of accuracy and precision . In a study of this data is used amounted to 299 patients who come from UCI Machine Learning . Measurement of the distance to the K-NN in the research is to use vector Euclidean Distance with a value of neighborhood 3, 5, 7 and 9. From the implementations are done obtained results that classification K Nearest Neighbor to the data of patients failing heart without SMOTE obtained results validate the best there is on the Nearest Neighbor $k = 9$ with an average accuracy of 71.59% and accuracy is the highest that is by 83.33% in the fold second . Whereas the highest accuracy test data was obtained at Nearest Neighbor $k = 3$ with an accuracy value of 71.66% . classification K Nearest Neighbor to the data of patients failing heart by using SMOTE obtained results validate the best there is on the Nearest Neighbor $k = 3$ with an average accuracy of 87.88% with accuracy the highest that is by 80.14% in the fold sixth and eighth . Whereas in the test data the highest accuracy was obtained at Nearest Neighbor $k = 3$ with an accuracy value of 63.33%.

Keyword: Classification , K-NN, Heart Failure , SMOT E , Cross Fold Validation.