

ANALISIS DAMPAK PENERAPAN METODE DISKRETISASI DATA TERHADAP PERFORMA DIAGNOSIS PENYAKIT LIVER BERBASIS *K-NEAREST NEIGHBOR*

Candra Utomo Alatas
(08.04.1.1.1.00111)

Pembimbing 1 : Mula'ab, S.Si., M.Kom.
Pembimbing 2 : Iwan Santosa, S.T., M.T.

ABSTRAK

Akurasi selalu menjadi isu penting dalam klasifikasi data. Semakin tinggi akurasi suatu metode klasifikasi, semakin baik hasil klasifikasinya. Dalam kasus diagnosis penyakit liver, metode *K-Nearest Neighbor* (KNN) menunjukkan akurasi yang tinggi. Tentu akan sangat baik apabila performa KNN ini dapat ditingkatkan. Dengan begitu, diharapkan diagnosis yang dihasilkan oleh KNN akan semakin akurat.

Dalam penelitian Tugas Akhir ini, himpunan data pasien penyakit liver diklasifikasikan menggunakan KNN. Klasifikasi dilakukan terhadap himpunan data asli dan himpunan data yang telah terlebih dahulu melalui proses diskretisasi berbasis *entropy*. Kemudian, akurasi klasifikasi himpunan data asli dibandingkan dengan akurasi klasifikasi himpunan data terdiskretisasi. Hal ini dimaksudkan untuk melihat dampak penerapan diskretisasi berbasis *entropy* terhadap akurasi diagnosis penyakit liver berbasis KNN.

Dari uji coba yang dilakukan, terlihat bahwa akurasi klasifikasi himpunan data terdiskretisasi lebih tinggi daripada akurasi klasifikasi himpunan data asli pada seluruh nilai k yang diuji (1, 3, 5, 7, 9, 11, 13 dan 15). Oleh karena itu, disimpulkan bahwa penerapan diskretisasi berbasis *entropy* mampu meningkatkan akurasi diagnosis penyakit liver berbasis KNN. Peningkatan yang diperoleh mulai dari 1,03% hingga 4,8%. Adapun akurasi tertinggi diperoleh pada klasifikasi himpunan data terdiskretisasi dengan nilai $k = 13$, yakni sebesar 71,7%.

Kata kunci : Akurasi, diagnosis penyakit liver, KNN, diskretisasi berbasis *entropy*

ANALYSIS OF THE EFFECT OF APPLYING DATA DISCRETIZATION METHOD ON THE PERFORMANCE OF K-NEAREST NEIGHBOR-BASED DIAGNOSIS OF LIVER DISEASE

Candra Utomo Alatas
(08.04.1.1.1.00111)

Supervisor 1 : Mula'ab, S.Si., M.Kom.
Supervisor 2 : Iwan Santosa, S.T., M.T.

ABSTRACT

Accuracy has always been an important issue in data classification. The higher the accuracy of a classification method, the better the classification result. In the case of liver disease diagnosis, K-Nearest Neighbor (KNN) method showed high accuracy. It would be very nice if that performance could be increased. It was hoped that the diagnosis generated by KNN would be more accurate.

In this research, liver patient data set was classified with KNN. The classification performed on original and entropy-based-discretized data set. Then, accuracy obtained from the original data set classification was compared with accuracy obtained from the discretized data set classification. It was intended to look at the effect of applying entropy-based discretization on the accuracy of KNN-based diagnosis of liver disease.

From the experiments performed, it could be seen that the accuracy of the discretized data set classification was higher than the accuracy of the original data set classification on all of the parameter k values observed. Thus, it was concluded that applying entropy-based discretization could increase accuracy of KNN-based diagnosis of liver disease. The increase obtained varied from 1,03 % to 4,8%. The highest accuracy was obtained at k=13, which was 71,7%.

Keywords : Accuracy, diagnosis of liver disease, KNN, entropy-based discretization