

PERBANDINGAN ALGORITMA K-NEAREST NEIGHBOR DAN GAUSSIAN NAÏVE BAYES PADA KLASIFIKASI PENYAKIT DIABETES MELITUS

Puput Tri Rahayu , Daryanto, Qurrota A'yun
Prodi Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember
Jl. Karimata No. 49, Jember, 68121, Indonesia
Email :

Abstrak

Data mining merupakan pemrosesan suatu data menggunakan cara statistik, matematik, dll untuk mengidentifikasi suatu informasi pengetahuan potensial dan berguna yang tersimpan dalam basis data besar. Klasifikasi adalah salah satu tugas dari data mining yang bertujuan untuk memprediksi label kategori benda yang tidak diketahui sebelumnya, dalam membedakan antara objek yang satu dengan yang lainnya berdasarkan atribut atau fitur Terdapat beberapa algoritma yang dapat digunakan untuk mengklasifikasikan seperti algoritma K-Nearest Neighbor yang memiliki keunggulan yaitu lebih efektif didata training yang besar, dapat menghasilkan data yang lebih akurat dan Gaussian Naïve Bayes dimana metode ini hanya membutuhkan jumlah data trainng yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. Setelah dilakukan penelitian menggunakan kedua algoritma tersebut, penelitian menggunakan algoritma K-Nearest Neigboar menghasilkan *accuracy* sebesar 71.27%, *sensitivity* sebesar 76.86%, *specifity* sebesar 81.59%%, *precision* sebesar 76.15% dan *error rate* sebesar 35.05% sedangkan pada saat mengguakan algoritma Gaussian Naïve Bayes menghasilkan *accuracy* sebesar 73.50%, *sensitivity* sebesar 96.14%, *specifity* sebesar 86.45%, *precision* sebesar 78.98% dan *error rate* sebesar 29.47%. berdasarkan hasil penelitian tersebut menunjukkan untuk algoritma Gaussian Naïve Bayes mempunyai *accuracy* yang lebih akurat bila dibandingkan dengan algortima K-Nearest Neighbor dalam mengklasifikasi data Diabetes Melitus.

Kata kunci : Diabetes Mellitus, Gaussian Naïve Bayes, K-Nearest Neighbor

COMPARISON OF K-NEAREST NEIGHBOR AND GAUSSIAN NAÏVE BAYES ALGORITHM ON THE CLASSIFICATION OF DIABETES MELLITUS

Puput Tri Rahayu , Daryanto, Qurrota A'yun
Informatics Engineering Department, Faculty of Engineering, University of
Muhammadiyah Jember
Jl. Karimata No. 49, Jember, 68121, Indonesia
Email :

Abstract

Data mining is a statement of data using statistical, mathematical, etc. methods to identify potential and useful knowledge information stored in large databases. Classification is one of the tasks of data mining that aims to predict the labels of previously unknown object categories, in distinguishing between objects from one another based on attributes or features there are several algorithms that can be used to classify such as the K-Nearest Neighbor algorithm which has advantages. which is more effective in large training data, can produce more accurate data and Gaussian Naïve Bayes where this method only requires a small amount of training data to determine the parameters needed in the classification process. After doing research using these two algorithms, research using the K-Nearest Neighbor algorithm produces an accuracy of 71.27%, sensitivity of 76.86%, specificity of 81.59%, precision of 76.15% and error rate of 35.05 % while using the Gaussian Naïve Bayes algorithm, the accuracy is 73.50%, the sensitivity is 96.14%, the specificity is 86.45%, the precision is 78.98% and the error rate is 29.47%. Based on the results of this study, the Gaussian Naïve Bayes algorithm has more accuracy than the K-Nearest Neighbor algorithm in classifying Diabetes Mellitus data.

Keywords: *Diabetes Mellitus, Gaussian Naïve Bayes, K-nearest Neighbor*