

PERBANDINGAN ALGORITMA NAIVE BAYES DENGAN ALGORITMA K-NEAREST NEIGHBOR UNTUK KLASIFIKASI PENYAKIT JANTUNG

¹Agil Langga Dikayanto (1410651223), ²Agung Nilogiri, S.T., M.Kom.

Jurusan Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember

agillangga95@gmail.com

agungnilogiri@unmuhjember.ac.id

ABSTRAK

Saat ini kebutuhan terhadap analisis suatu data sangat dibutuhkan. Salah satu pendekatan yang dapat dilakukan untuk menganalisis sekumpulan data yaitu dengan mengklasifikasi data. Beberapa metode klasifikasi yang umum digunakan seperti Artificial Neural Network (ANN), Support Vector Machines (SVM), Decision Tree, Bayesian, dan sebagainya. Klasifikasi Naive Bayes terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam basis data dengan jumlah yang besar. Selain Naive Bayes, algoritma K-Nearest Neighbor juga memiliki tingkat akurasi yang tinggi. Oleh karena itu peneliti ingin melakukan sebuah penelitian dengan membandingkan algoritma Naive Bayes dan algoritma K-Nearest Neighbor menggunakan data diagnosa penyakit jantung yang diambil dari penyedia layanan dataset publik UCI Machine Learning dengan menggunakan teknik pengujian Cross Validation. Dari keseluruhan pengujian, nilai K optimal untuk algoritma K-Nearest Neighbor yaitu dengan jarak ketetanggaan sebesar 7. Algoritma Naive Bayes menghasilkan performa tertinggi pada 10-fold Cross Validation pada folder data pengujian ke-4 dengan akurasi yaitu 90.00%, dan nilai presisi sebesar 86,67%. Algoritma K-Nearest Neighbor juga menghasilkan akurasi tertinggi dengan jumlah ketetanggaan sebesar 7 pada pengujian 10-fold Cross Validation dengan akurasi tertinggi pada folder data pengujian ke-4 yaitu sebesar 80.00%, dan presisi sebesar 90.00%. Berdasarkan pengujian yang telah dilakukan, algoritma Naive Bayes lebih akurat dan lebih baik dalam klasifikasi penyakit jantung dibandingkan algoritma K-Nearest Neighbor.

Kata Kunci : Algoritma Naive Bayes, Algoritma K-Nearest Neighbor, Analisis, Klasifikasi, Perbandingan, Cross Validation.

COMPARISON OF NAIVE BAYES ALGORITHM WITH K-NEAREST NEIGHBOR ALGORITHM FOR CLASSIFICATION OF HEART DISEASE

¹Agil Langga Dikayanto (1410651223), ²Agung Nilogiri, S.T., M.Kom.
Jurusan Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember
agillangga95@gmail.com
agungnilogiri@unmuhjember.ac.id

ABSTRACT

Currently analysis of a data is very needed. One approach that can be done to analyze a data set is to classify data. Several classification methods are commonly used such as Artificial Neural Network (ANN), Support Vector Machines (SVM), Decision Tree, Bayesian, and so on. The Naive Bayes classification proved to have high accuracy and speed when applied to a large number of databases. Besides Naive Bayes, the K-Nearest Neighbor algorithm also has a high level of accuracy. Therefore the researcher wants to conduct a study by comparing the Naive Bayes algorithm and the K-Nearest Neighbor algorithm using heart disease diagnostic data taken from the public dataset provider UCI Machine Learning using Cross Validation testing techniques. From the whole test, the optimal K for the K-Nearest Neighbor algorithm that is neighboring distance is 7. The Naive Bayes algorithm produces the highest performance at 10-fold Cross Validation in the 4th test data folder with an accuracy of 90.00%, and a precision value of 86, 67%. The K-Nearest Neighbor Algorithm also produces the highest accuracy with a neighbor value of 7 in the 10-fold Cross Validation test with the highest accuracy in the 4th test data folder which is equal to 80.00%, and a precision of 90.00%. Based on the tests that have been done, the Naive Bayes algorithm is more accurate and better in the classification of heart disease than the K-Nearest Neighbor algorithm.

Keywords: Naive Bayes Algorithm, K-Nearest Neighbor Algorithm, Analysis, Classification, Comparison, Cross Validation.