

ABSTRAK

Nugroho, Laurensius Ferdinan P. 2025. ***Perbandingan Algoritma KNN dan Naïve Bayes untuk Klasifikasi Gejala Penyakit Diabetes***. Skripsi. Yogyakarta: Informatika, FST, USD.

Penelitian ini membahas perbandingan algoritma KNN dan *Naïve Bayes* dalam proses klasifikasi gejala penyakit diabetes. Tujuan dari penelitian ini adalah (1) mengetahui bagaimana menerapkan algoritma *Naïve Bayes* dan KNN untuk klasifikasi gejala penyakit diabetes, (2) mengetahui hasil klasifikasi dari suatu gejala diabetes dengan data tertentu, dan (3) membandingkan hasil klasifikasi dari metode *Naïve Bayes* dan KNN.

Jenis penelitian ini adalah penelitian kuantitatif komparatif. Sumber data penelitian ini adalah dataset *Pima Indians Diabetes Database* dari situs *Kaggle* dengan data berupa kelas-kelas gejala penyakit diabetes. Metode pengumpulan data yang digunakan adalah metode studi dokumentasi. Sementara untuk analisa data, peneliti membandingkan hasil dari kedua proses klasifikasi.

Kesimpulan dari penelitian ini adalah (1) algoritma *Naïve Bayes* dan KNN dapat diterapkan untuk klasifikasi gejala penyakit diabetes dengan membuat program dengan menggunakan bahasa pemrograman *Python*, (2) hasil klasifikasi dari proses klasifikasi gejala penyakit diabetes bervariasi, tergantung dari kelas-kelas apa saja yang digunakan dalam proses dengan nilai akurasi KNN tertinggi adalah 72.7% dan nilai akurasi tertinggi *Naïve Bayes* adalah 74.7%, dan (3) algoritma *Naïve Bayes* memiliki hasil akurasi yang lebih akurat dibandingkan algoritma KNN.

Kata kunci: gejala penyakit diabetes, KNN, *naïve bayes*, klasifikasi, akurasi.

ABSTRACT

Nugroho, Laurensius Ferdinan P. 2025. *Comparison Algorithm of KNN and Naïve Bayes for Classifying Diabetes Symptoms*. Thesis. Yogyakarta: Informatics, FST, USD.

This study discusses the comparison of KNN and Naïve Bayes algorithms in the process of classifying diabetes symptoms. The objectives of this study are (1) to find out how to apply the Naïve Bayes and KNN algorithms for classifying diabetes symptoms, (2) to find out the classification results of a diabetes symptom with certain data, and (3) to compare the classification results of the Naïve Bayes and KNN methods.

This type of research is comparative quantitative research. The data source for this study is the Pima Indians Diabetes Database dataset from the Kaggle website with data in the form of classes of diabetes symptoms. The data collection method used is the documentation study method. Meanwhile, for data analysis, the researcher compared the results of the two classification processes.

The conclusion of this study were (1) the Naïve Bayes and KNN algorithms can be applied to classify diabetes symptoms by creating a program using the Python programming language, (2) the classification results from the diabetes symptom classification process vary, depending on the classes used in the process with the highest KNN accuracy value being 72.7% and the highest Naïve Bayes accuracy value being 74.7%, and (3) the Naïve Bayes algorithm has more accurate results than the KNN algorithm.

Keywords: diabetes symptoms, KNN, naïve bayes, classification, accuracy