WEB-BASED DIABETES MELLITUS PREDICTION USING THE K-NEAREST NEIGHBOR ALGORITHM BASED ON CLINICAL DATA (CASE STUDY: PUSKESMAS MLATI II)

Neysa Devina

Program Studi Informatika Medis, Fakultas Sains dan Teknologi
Universitas Teknologi Yogykarta
Jl. Ringroad Utara Jombor Sleman Yogyakarta
E-mail : nysdvn@gmail.com

ABSTRACT

The lack of public awareness of the early symptoms of diabetes mellitus has led to many new cases being detected only after complications have developed. At Puskesmas Mlati II, data shows that only a small percentage of diabetic patients regularly undergo check-ups. To facilitate early detection, this study develops a web-based diabetes prediction system utilizing the K-NN algorithm. The dataset consists of 305 patient records that have undergone preprocessing and are divided into 183 training data points and 122 testing data points. This system utilizes clinical variables, including fasting blood sugar, HbA1c, urinary albumin, and others. In the modeling process, two encoding scenarios were compared: manual encoding and label encoding. The model using label encoding with a K value of 9 achieved a training accuracy of 81% and a testing accuracy of 80%. In contrast, the model with manual encoding at K = 7 resulted in a training accuracy of 91% and a testing accuracy of 97%. The system was developed using the CodeIgniter framework for the frontend and Python Flask for the backend. Black box testing showed that all features functioned as expected without errors. The results of the study indicate that the preprocessing method has a significant impact on model performance, making the selection of the correct technique crucial for developing an accurate prediction system. This system is expected to become a practical tool for early diabetes detection, supporting the improvement of healthcare services at Puskesmas Mlati II.

Keywords: Diabetes Mellitus, Prediction, K-Nearest Neighbor, Website, Clinical Data.