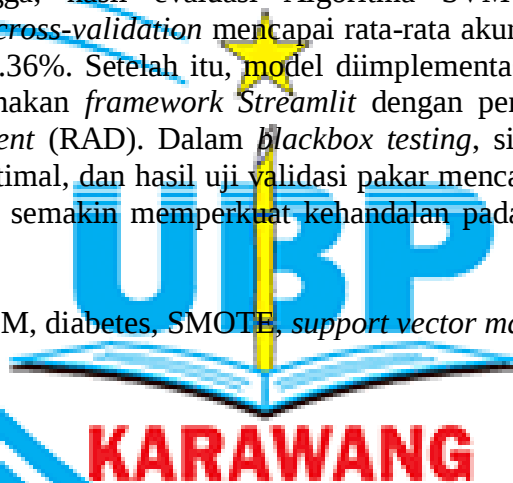


ABSTRAK

Diabetes menjadi salah satu penyakit yang dapat mengancam nyawa apabila tidak ditangani dengan cepat, karena dapat menyebabkan komplikasi serius pada jantung, ginjal, saraf, dan mata. Diabetes terjadi ketika kadar gula darah dalam tubuh tidak terkendali, dan pankreas tidak dapat memproduksi insulin dengan cukup. Selain itu, jumlah orang yang menderita diabetes terus meningkat secara global, dengan biaya perawatan di Indonesia mencapai 323,8 USD per-tahun. Berdasarkan jumlah kematian pasien yang mengidap diabetes, penelitian ini bertujuan mengembangkan sistem berbasis web untuk mendiagnosis seseorang terkena diabetes atau tidak. Pada penelitian ini, proses data mining dilakukan dengan enam tahapan dalam *Cross-Industry Standard Process for Data Mining* (CRISP-DM). Adapun pada dataset diabetes terdapat ketidakseimbangan yang diatasi dengan *Synthetic Minority Over-sampling Technique* (SMOTE). Selanjutnya, dilakukan proses pemodelan menggunakan Algoritma *Support Vector Machine* (SVM) dengan kernel linear. Sehingga, hasil evaluasi Algoritma SVM dengan SMOTE menggunakan *10-fold cross-validation* mencapai rata-rata akurasi 88.00%, presisi 88.67%, dan *recall* 87.36%. Setelah itu, model diimplementasikan dalam sistem berbasis web menggunakan *framework Streamlit* dengan pengembangan *Rapid Application Development* (RAD). Dalam *blackbox testing*, sistem menunjukkan fungsionalitas yang optimal, dan hasil uji validasi pakar mencapai tingkat akurasi sebesar 100%. Hal ini semakin memperkuat kehandalan pada sistem klasifikasi penyakit diabetes.

Kata Kunci: CRISP-DM, diabetes, SMOTE, *support vector machine*



ABSTRACT

Diabetes is a life-threatening disease if not treated quickly, as it can cause serious complications to the heart, kidneys, nerves and eyes. Diabetes occurs when blood sugar levels in the body go out of control, and the pancreas cannot produce enough insulin. In addition, the number of people suffering from diabetes continues to increase globally, with the cost of treatment in Indonesia reaching 323.8 USD per-year. Based on the number of deaths of patients with diabetes, this research aims to develop a web-based system to diagnose whether a person has diabetes or not. In this research, the data mining process is carried out with six stages in the Cross-Industry Standard Process for Data Mining (CRISP-DM). As for the diabetes dataset, there is an imbalance that is resolved with the Synthetic Minority Over-sampling Technique (SMOTE). Furthermore, the modeling process is carried out using the Support Vector Machine (SVM) algorithm with a linear kernel. Thus, the evaluation results of the SVM algorithm with SMOTE using 10-fold cross-validation achieved an average accuracy of 88.00%, precision of 88.67%, and recall of 87.36%. After that, the model was implemented in a web-based system using the Streamlit framework with Rapid Application Development (RAD). In blackbox testing, the system showed optimal functionality, and the expert validation test results reached an accuracy rate of 100%. This further strengthens the reliability of the diabetes classification system.

Keyword: CRISP-DM, diabetes, SMOTE, support vector machine

