

ABSTRAK

KLASIFIKASI PENYAKIT *DIABETES MELLITUS* TIPE II BERBASIS *MACHINE LEARNING* MENGGUNAKAN LIGHTGBM

oleh

SALMA IRENA FEBRIASTIA

Diabetes mellitus merupakan sebuah kondisi metabolis serius dan kronis yang terjadi karena kenaikan kadar glukosa darah akibat tubuh tidak dapat memproduksi hormon insulin atau tidak dapat menggunakan insulin yang dihasilkan dengan efektif. Sekitar 537 juta orang di berbagai bagian dunia mengalami diabetes. Pada tahun 2019, diabetes menjadi salah satu penyebab kematian utama, mengalami peningkatan sebesar 70% sejak tahun 2000. Di Indonesia, diabetes menempati peringkat tiga sebagai penyebab kematian pada tahun 2019, mencapai persentase 6.23%. Penelitian ini bertujuan untuk melakukan klasifikasi penyakit *diabetes mellitus* tipe II menggunakan *Light Gradient-Boosting Machine* (LightGBM) dengan dua skenario, yaitu menggunakan fitur yang mengandung hasil tes laboratorium berupa FPG (skenario 1) dan tanpa fitur hasil tes laboratorium (skenario 2). Pada kedua skenario, dilakukan evaluasi menggunakan konfigurasi seluruh fitur, fitur terseleksi berdasarkan mutual *information*, dan fitur rekomendasi dari *expert*. Lalu membuat sebuah antarmuka *website* untuk memprediksi diagnosis penyakit *diabetes mellitus* tipe II. Metode yang digunakan adalah *Cross-Industry Standard Process for Data Mining* (CRISP-DM). Hasil evaluasi menunjukkan bahwa model memiliki kinerja yang baik dalam semua konfigurasi fitur pada kedua skenario. Tujuh fitur rekomendasi berdasarkan *expert* pada skenario 2 (AGE, Nocturia, Polyuria, Weight_loss, Polydipsia, Polyphagia, BMI) dengan akurasi 94,87% digunakan sebagai bahan untuk memprediksi diagnosis pada antarmuka *website* karena relevan dengan ilmu kedokteran dan pengguna dapat melakukan *prescreening* awal *diabetes mellitus* tipe II sebelum melakukan pemeriksaan lebih lanjut ke dokter, sehingga meningkatkan efisiensi waktu dan biaya.

Kata kunci: CRISP-DM, *diabetes mellitus*, klasifikasi, LightGBM, klasifikasi, *mutual information*

ABSTRACT

MACHINE LEARNING-BASED TYPE II DIABETES MELLITUS DISEASE CLASSIFICATION USING LIGHTGBM

By

SALMA IRENA FEBRIASTIA

Diabetes mellitus is a serious and chronic metabolic condition that occurs due to elevated blood glucose levels because the body cannot produce the hormone insulin or cannot use the insulin produced effectively. About 537 million people in various parts of the world have diabetes. In 2019, diabetes became one of the leading causes of death, having increased by 70% since 2000. In Indonesia, diabetes ranked third as a cause of death in 2019, reaching a percentage of 6.23%. This study aims to classify type II diabetes mellitus using Light Gradient-Boosting Machine (LightGBM) with two scenarios, namely using features containing laboratory test results in the form of FPG (scenario 1) and without laboratory test results features (scenario 2). In both scenarios, evaluation was carried out using the configuration of all features, selected features based on mutual information, and expert recommendation features. Then create a website interface to predict the diagnosis of type II diabetes mellitus. The method used is Cross-Industry Standard Process for Data Mining (CRISP-DM). The evaluation results showed that the model performed well in all feature configurations in both scenarios. Seven expert-based recommendation features in scenario 2 (AGE, Nocturia, Polyuria, Weight_loss, Polydipsia, Polyphagia, BMI) with an accuracy of 94.87% are used as configuration for predicting diagnosis on the website interface because they are relevant to medical science and users can perform initial prescreening of type II diabetes mellitus before conducting further examinations to doctors, thereby increasing time and cost efficiency.

Keywords: *Classification, CRISP-DM, diabetes mellitus, LightGBM, mutual information*