PLAGIAT MERUPAKAN TINDAKAN TIDAK TERPUJI

ABSTRAK

Churn pelanggan merupakan salah satu permasalahan yang dapat memengaruhi keberlangsungan bisnis layanan *subscription*, termasuk pada platform streaming seperti Netflix. Kemampuan untuk mengidentifikasi pelanggan yang berpotensi berhenti berlangganan sejak dini dapat membantu perusahaan dalam menyusun strategi retensi pelanggan secara efektif. Penelitian ini bertujuan untuk menerapkan algoritma Light Gradient Boosting Machine (LightGBM) dalam memprediksi *churn* pelanggan Netflix serta menginterpretasikan hasil prediksi menggunakan metode Shapley Additive Explanations (SHAP) sehingga faktor-faktor yang memengaruhi keputusan model dapat dijelaskan. Penelitian menggunakan *dataset* Netflix *Customer Churn* yang diperoleh dari Kaggle dengan jumlah 5.000 data pelanggan. Tahapan penelitian meliputi *Exploratory Data Analysis* (EDA), *pre-processing*, pembagian data *training* dan *testing* dengan menggunakan *K-Fold Cross-Validation* pada nilai *fold* 3, 5, dan 10. Optimasi hyperparameter menggunakan *GridSearchCV*, evaluasi model menggunakan *confusion matrix*, serta pengembangan sistem berbasis web menggunakan Streamlit. Hasil penelitian menunjukkan bahwa model terbaik diperoleh pada skenario *K-Fold* = 10 dengan nilai *accuracy* sebesar 99,8%, *precision* sebesar 100%, *recall* sebesar 99,6%, dan *F1-score* sebesar 99,8%. Interpretasi menggunakan SHAP menunjukkan bahwa fitur avg_watch_time_per_day, last_login_days, dan watch_hours merupakan faktor yang paling berpengaruh terhadap prediksi *churn*, sehingga rendahnya aktivitas penggunaan layanan menjadi indikator utama pelanggan berpotensi berhenti berlangganan. Hasil pengujian *System Usability Scale* (SUS) memperoleh skor rata-rata 83,57 yang termasuk kategori Excellent, Grade A, dan Acceptable. Berdasarkan hasil tersebut, dapat disimpulkan bahwa algoritma LightGBM mampu menghasilkan model prediksi churn dengan performa yang baik, sedangkan SHAP mampu memberi interpretasi yang membantu dalam memahami faktor-faktor yang memengaruhi hasil prediksi.

Kata Kunci : *customer churn*, Light Gradient Boosting Machine (LightGBM), Shapley Additive Explanations (SHAP), Netflix.

PLAGIAT MERUPAKAN TINDAKAN TIDAK TERPUJI

ABSTRACT

Customer churn is one of the issues that can affect the sustainability of subscription-based businesses, including streaming platforms such as Netflix. The ability to identify customers who are likely to cancel their subscriptions early on can help companies develop effective customer retention strategies. This study aims to apply the Light Gradient Boosting Machine (LightGBM) algorithm to predict Netflix customer churn and to interpret the prediction results using Shapley Additive exPlanations (SHAP), thereby explaining the factors that influence the model's decisions. The study uses the Netflix Customer Churn dataset obtained from Kaggle, consisting of 5,000 customer data. The research stages include Exploratory Data Analysis (EDA), preprocessing, and data splitting using K-Fold Cross-Validation with fold values of 3, 5, and 10. Hyperparameter optimization is performed using GridSearchCV, model evaluation using a confusion matrix, and the development of a web-based system using Streamlit. The results show that the best model was obtained in the K-Fold = 10 scenario, with an accuracy of 99.8%, precision of 100%, recall of 99.6%, and an F1-score of 99.8%. Interpretation using SHAP revealed that the features avg_watch_time_per_day, last_login_days, and watch_hours were the most influential factors in churn prediction, indicating that low user activity is the primary indicator of potential customer churn. The System Usability Scale (SUS) evaluation yielded an average score of 83.57, which falls into the Excellent, Grade A, and Acceptable categories. Based on these results, it can be concluded that the LightGBM algorithm is capable of producing a high-performance churn prediction model, while SHAP provides interpretations that help in understanding the factors influencing the prediction results.

Keywords: customer churn, Light Gradient Boosting Machine (LightGBM), Shapley Additive Explanations (SHAP), Netflix.