

ABSTRACT

FEATURE IMPORTANCE AND MODEL PERFORMANCE ANALYSIS OF LOGISTIC REGRESSION AND MULTILAYER PERCEPTRON IN ALTERNATIVE CREDIT CLASSIFICATION WITH TELECOMMUNICATIONS DATA

by

Almas Mirzandi Ramadhan

21/472900/PA/20357

Dalam lanskap layanan keuangan yang terus berkembang, kebutuhan akan model penilaian kredit yang adil dan transparan semakin meningkat, terutama di wilayah yang sedang berjuang untuk inklusi keuangan yang lebih besar. Dalam dunia penilaian kredit, model klasifikasi kredit digunakan sebagai salah satu aspek penilaian. Model ini akan menilai apakah penerima pinjaman akan *default* atau tidak. *Default*, dalam hal ini adalah ketika penerima pinjaman gagal membayar kembali pinjaman yang telah diberikan oleh lembaga keuangan terkait.

Penelitian ini meneliti bagaimana data telekomunikasi dapat meningkatkan klasifikasi kredit dengan membandingkan kinerja dan kemampuan interpretasi dari model Regresi Logistik (LR) dan Multilayer Perceptron (MLP).

Berfokus pada Indonesia, penelitian ini membahas tantangan dalam menggunakan data alternatif. Dengan mengeksplorasi perbedaan fitur yang penting di antara model-model ini, penelitian ini menjelaskan kemampuan mereka untuk memberikan wawasan yang dapat ditindaklanjuti untuk penilaian kelayakan kredit. Kinerja dievaluasi melalui metrik-metrik utama, untuk memastikan pandangan yang seimbang terhadap kekuatan dan keterbatasan masing-masing model. Hasilnya menyoroti potensi data alternatif untuk mendukung upaya inklusi keuangan, menawarkan jalan menuju solusi penilaian kredit yang lebih transparan dan efektif.

Secara ringkas, penelitian ini menunjukkan bahwa model MLP secara konsisten mengungguli LR. Tanpa SMOTE, MLP mencapai AUC 0,6618 dan skor K-S 0,2688, dibandingkan dengan AUC LR 0,628 dan K-S 0,2081; dengan SMOTE, angka-angka tersebut meningkat menjadi 0,7443 dan 0,3750 untuk MLP, serta 0,7293 dan 0,3649 untuk LR. Kedua model ini sangat bergantung pada jumlah rata-rata transaksi per hari sebagai prediktor utama mereka - SHAP menunjukkan kisaran sekitar -0.06 hingga +0.07 untuk MLP dan -0.04 hingga +0.05 untuk LR - sementara variabel terkait produk yang dibeli orang (terdiri dari lebih dari 25% fitur) dan jendela perilaku 360 hari juga secara signifikan berkontribusi dalam memprediksi risiko gagal bayar.

Keywords: Logistic Regression, Multilayer Perceptron, SHAP values, Credit Scoring, Alternative Credit Scoring, Feature Importance, Model Explainability, Financial Inclusion