Predictive Analysis of Loan Repayment Capacity in the Unbanked Population Using the Home Credit Default Risk Dataset

ABSTRACT

The Home Credit Default Risk dataset provides an opportunity to explore financial inclusion by predicting loan repayment capabilities among individuals with minimal credit histories. This research focuses on applying classification techniques to assess the viability of loan applicants based on alternative data sources, such as telecommunications and transaction data. Through this analysis, we aim to understand the effects of machine learning parameter adjustments—including learning rate and optimizer changes—on model performance. We will also evaluate the impact of different classifiers, such as Stochastic Gradient Descent (SGD), on predicting outcomes. This study promotes fair financial access in addition to improving prediction accuracy.

-Mohammed