Home Credit Default Risk

Business Problem Statement

Home Credit aims to expand financial inclusion by providing loans to individuals with limited or no credit history. The current loan approval system is unable to accurately differentiate between applicants who can repay loans and those who cannot. This results in qualified applicants being denied and unqualified ones receiving loans, increasing default rates and missed opportunities for underserved populations.

The purpose of this project is to develop a predictive model that improves the accuracy of loan approval decisions by utilizing alternative data sources such as transactional information.

Benefits of improving Loan Approval Decisions

The benefits of a more accurate loan approval system include:

- **Increasing Approval Rates:** Ensuring that creditworthy and or cash flow applicants are granted loans, expanding the customer base.
- **Reducing Defaults:** Lowering the financial risks associated with loan repayments.
- **Enhancing Financial Inclusion:** Supporting underserved populations with a fair and safe borrowing experience.
- **Operational Efficiency:** Streamlining the approval process to save time and resources.

Analytics Approach

The approach will involve developing a supervised classification model to predict the probability of loan repayment.

- **Data Sources:** Utilize alternative data such as current or future assets, credit history, and transactional behavior.
- **Target Variable:** Loan repayment status (TARGET: 1 for default, 0 for non-default).
- **Methods:** Explore and compare machine learning models, such as Logistic Regression, Random Forests, and Gradient Boosting.
- **Evaluation Metric:** Models will be assessed using the Area Under the receiver operating characteristic Curve to ensure accuracy and reliability.

Success metrics

The success of this project will be determined by the model's ability to outperform the current loan approval process in several key areas. The model should demonstrate a high AUC-ROC score, indicating strong predictive accuracy and reliability. It should also increase the approval rates for qualified applicants while effectively identifying and reducing defaults among high-risk borrowers. The solution must also process loan applications

quickly and efficiently, ensuring seamless integration into Home Credit's existing operational workflow. The goal is to deliver a model that is better across all these metrics compared to the current system.

Scope/Deliverables

Primary deliverables for this project will be the following.

- Pervading a predictive classification model that improves the accuracy of loan approval decisions.
- Give a comprehensive report detailing findings, model performance, and actionable recommendations.
- Have a GitHub repository of all code & history of the project.
- Python or R code for inference on new stores.

Project Details

This project will be executed by Adam Edwards on or before April 22, 2025. Following is a list of project milestones.

- Business problem statement delivery
- Exploratory analysis
- Presentation draft
- Final presentation