Loan Default Prediction Project | Final Insights for the Stakeholders (Milestone 6)

Executive Summary Report

Prepared by: PhD Aleksandar Osmanli

Project Overview

I'm currently developing a data analytics project aimed at decreasing overall loan default among borrowers by predicting which client attributes most contribute their loan default. For the purposes of this project, loan default means borrowers failed to pay their loans to the financial institution. The ultimate goal for this project is to develop a machine learning (ML) model that predicts loan default. **This report offers details and key insights from all the project milestones and the future actions.**

Final Insights

- Older married clients with higher incomes, more months full-time employed and higher credit score should be segmented as **low-risk loan clients** and offer them best rates.
- On contrary, unemployed clients with loans with higher interest rates and more credit lines should be segmented as **high-risk loan clients**.
- Additionaly, in order to better predict the loan default, the following top 5 interactive features were identified with highest importance:

1. LoanAmount / Income,
2. Age / InterestRate,
3. InterestRate / Age,
4. Age*MonthsEmployed
5. Income*MonthsEmployed

Results

➔ **The best developed ML model showed excellent result in the loan default prediction.**
➔ **The model was able to correctly differentiate between loan defaulted and not in 75,77% cases.**
➔ **It fails to predict only 4% of all clients with loan default.**
➔ **It showed a well-calibrated likelihood score that can be used to target interventions and support most accurately.**