

Credit Risk Prediction



The Business Challenge

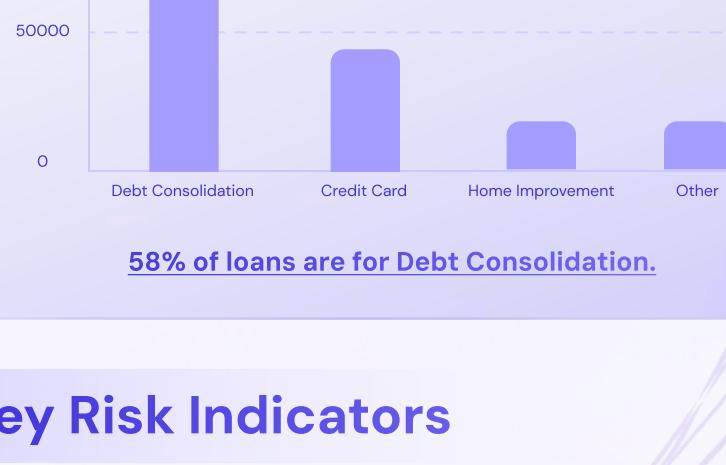
Reduce losses from loan defaults by building a predictive model to identify high-risk applicants before a loan is approved.

Our Data

Analyzed **466,285** loan records from **2007–2014** to find patterns distinguishing "Good" (Paid) and "Bad" (Defaulted) loans.

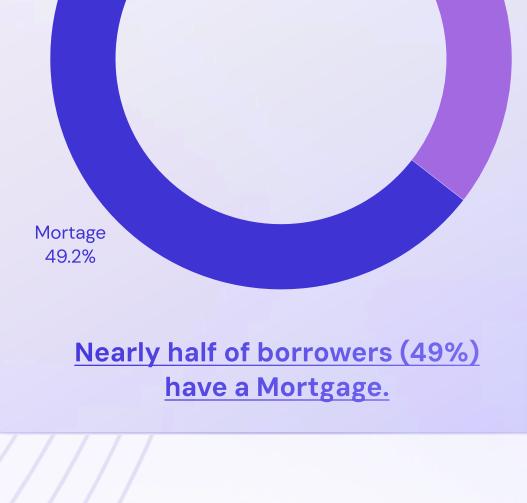
Who Are Our Borrowers?

Purpose



58% of loans are for Debt Consolidation.

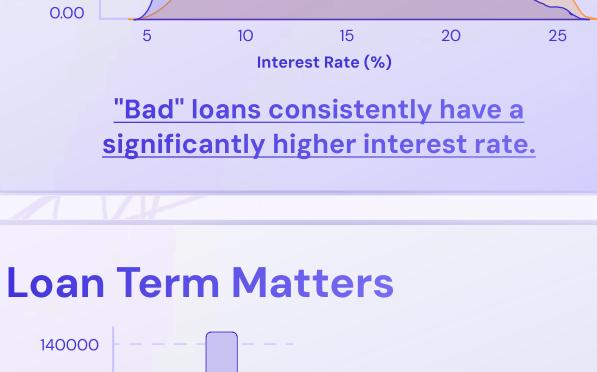
Home Ownership Status



Nearly half of borrowers (49%) have a Mortgage.

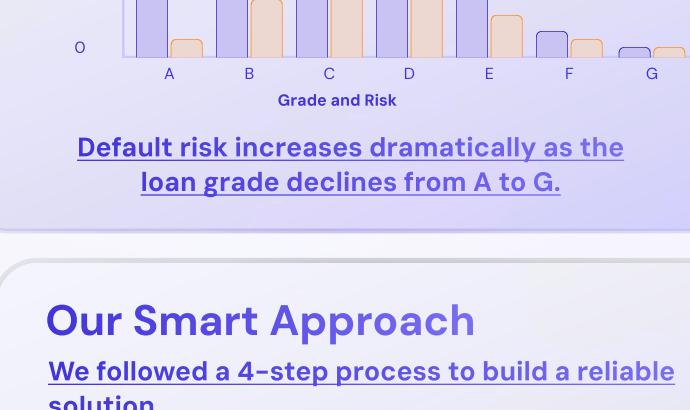
Key Risk Indicators

Interest Rate is a Key Differentiator



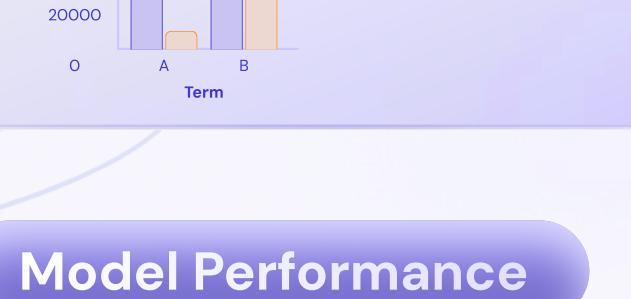
"Bad" loans consistently have a significantly higher interest rate.

Loan Grade is Highly Influential



Default risk increases dramatically as the loan grade declines from A to G.

Loan Term Matters



Loans with a 60-month term have a higher proportion of defaults.

Our Smart Approach

We followed a 4-step process to build a reliable solution.

- **Data Cleaning**
Selecting relevant features & removing unnecessary data.
- **Feature Engineering**
Creating new features like loan-to-income ratio for stronger predictions.
- **Modeling**
Training 3 different models: Logistic Regression, Random Forest, & LightGBM.
- **Optimization**
Performing hyperparameter tuning on the best model for maximum performance.

Model Performance



Contact

in Farrel Athalla Putra
o farrelathalla
✉ farrelxag@gmail.com