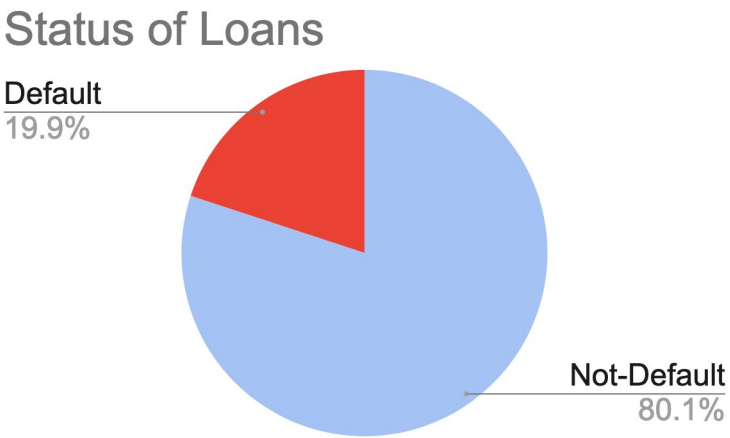
Proposal Toward More Profitable Business Practice

- customized solution for home equity credit lending business department

Ami Kang

Feb 2023

The challenge (problem)



- The 20% of accounts become default during the journey of loan terms.
- The business is running on the level of *negative ROI**.

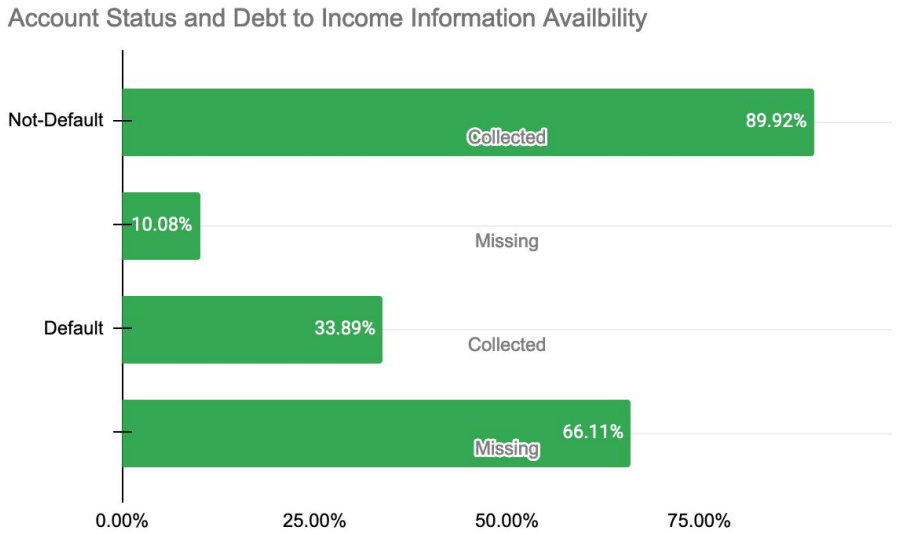
* **Assumption:** If it is assumed that 70% of the loan principal will not be recollected from the default accounts and the average interest income rate from the remaining loans are 10%

Solution Development

- Goal: Minimize the portion of default accounts in future portfolio
- Approach: Develop a methodology which helps the department identify credit applicants with high possibility of default using information only that is available at the time of lending decision making
- Validation: What-if analysis - if the new methodology had been applied to the current portfolio

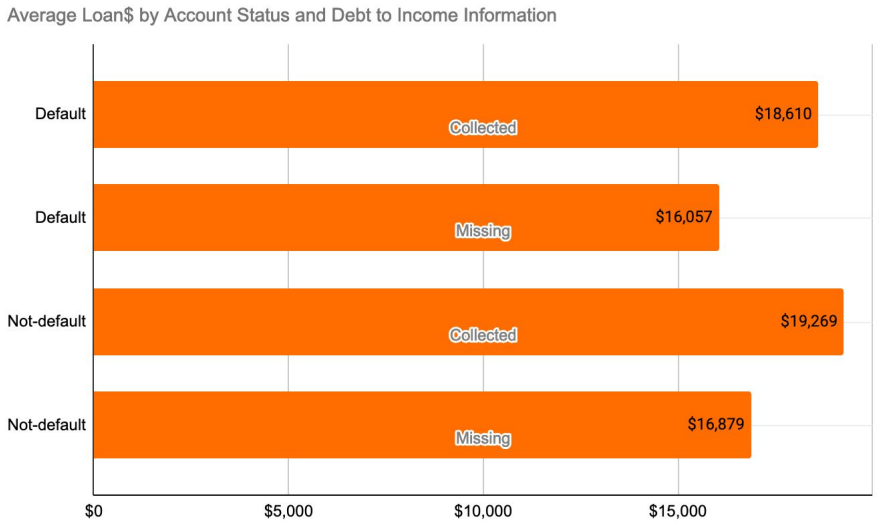
Current Business Practice Highlights 1

- 66% of the defaulted accounts did not provide debt-to-income information.



Current Business Practice Highlights 2

- Defaulted accounts received slightly lower loan than not-defaulted accounts.
- When debt-to-income information is not available, lower loan is granted.
- Debt-to-income missing indicator seems not used in loan amount allocation decision.



Model Development Methodology

Finalize the Customized Model

1

More feature-engineering and hyper-parameter tuning along with various decision threshold values to improve the performance of the final modeling framework

Compare the Performance of Candidate Models

2

Using the RECALL rate of the test data along with the precision, F1-rate and accuracy rates

Explore Binary-Classification Modeling Framework

3

Logistic regression
Decision Tree
Random Forest
Boosting Classification models
KNN
Linear Discriminant Analysis
Quadratic Discriminant Analysis

Final Model* Performance

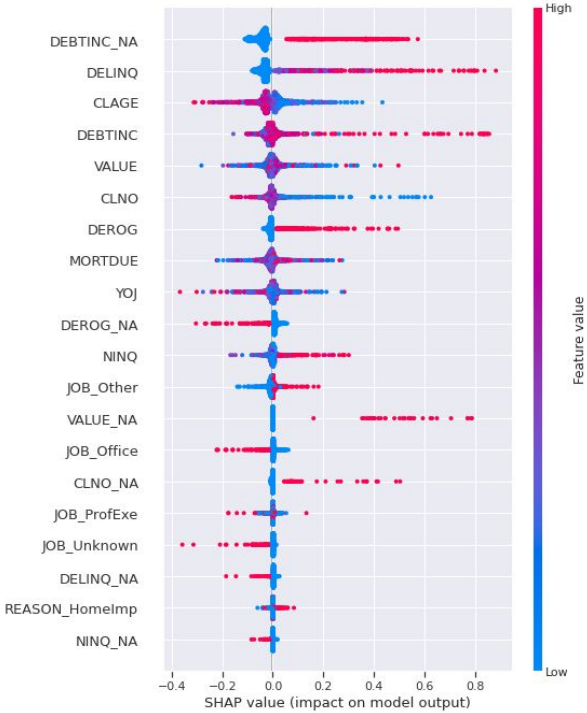
- The model correctly identifies 80% of defaulted accounts as future defaulters. (Recall rate=80%)

- 84% of the accounts identified as a future defaulter by the model turn out to be true default accounts. (Precision rate = 84%)

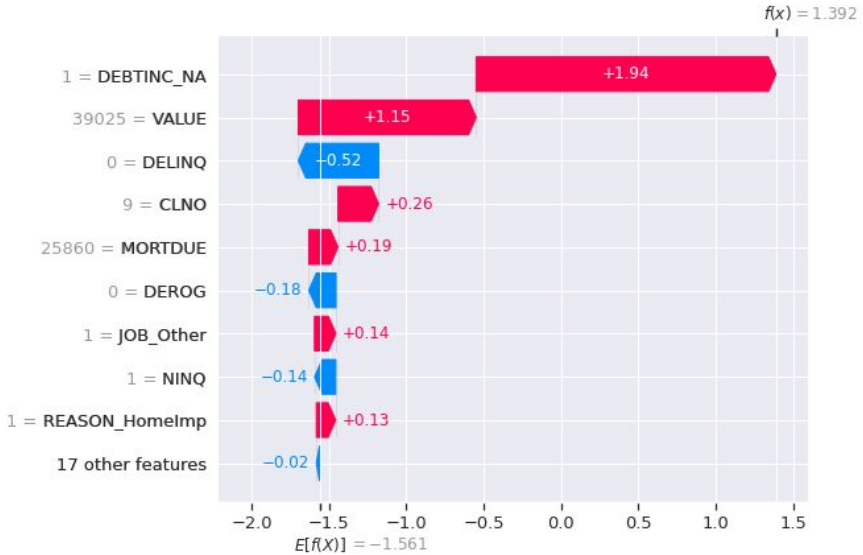
*Final model framework: Extreme Gradient Boosting Classifier

Learning and Usage of the Final Model

< Global Interpretability >



< Local Interpretability >



Status= Default
Default Probability = 80%
Decision Threshold = 40%

Business Solution Proposal

- Implement the (XGB) model developed in this study and calculate the default probability of an applicant prior to lending decisions.

- Decline the loan applications if a calculated default probability exceeds 40%.

- Utilize account-level analysis from the model to draft a comprehensive and reasonable explanation letter for adverse lending decisions.

What if the Proposal Had Been Applied

Expected Benefits

- **The expected default rates would have dropped from 20% to 3%.**
- The bank could have prevented 89% of the future default loans from the current portfolio while missing only 2% of booking non-default loans.
- **ROI would have been positive instead of negative (from -5% to 8%).**
- The bank would have lent $91M instead of $110M to clients while making the similar level of profits.
- The bank could have saved borrowing cost of extra $20M from the central bank or could have used that extra fund to more profitable business.

<Cost Benefit Analysis with the Adoption of the Proposed Solution>

	Investment	Safe-loans	Lost-loans	Expected Profit	Expected Loss	Net Profit	ROI
Current BAU	$110,903,500	$90,783,100	$20,120,400	$9,078,310	$14,084,280	-$5,005,970	***-5%***
Proposed Scenario	$91,049,900	$88,913,600	$2,136,300	$8,891,360	$1,495,410	$7,395,950	***8%***

Conservative Assumptions Applied to Cost Benefit Analysis

- Borrowers take out the full credit line available to them.
- An interest rate of 10% is applied (even though current home line equity loan interest rates are lower than 8%).
- It is assumed that the bank will incur a 70% loss on loan principal from defaulted accounts, although in reality the bank may incur an even greater loss.

Limitations of the Study

- Possibility of already-biased input data based on the current business practice
- Lack of data time stamp - assumed that all information was collected at the time of application except the loan amount
- Necessity of different observation period data for more scrutinized model validation
- Necessity of more accurate cost-benefit analysis incorporating the actual loss from the default loans and actual loans borrowed by the customers along with the implementation and operation cost of final (XGB) model
- Not taking the full advantage of XGB's capability to handle missing values
- Room to improve performance of the model and control possible overfitting (appropriate level of pruning and hyper parameter tuning)

Business Future Steps

- ❏ Carefully explore a possibility of unintended unfair treatment to a certain group of applicants by implementing the new model-based approach
- ❏ Develop an optimization model to decide the loan amount for each approved account
- ❏ Monitor the stability of the model and recalibrate the model on a regular basis
- ❏ Investigate the high proportion of current loans missing the critical income-to-debt information and reasoning of the current practice of granting loans to them
- ❏ Conduct more precise cost-benefit analysis before implementing the new method
- ❏ Explore opportunities to collect different informative data and apply them to the model

Thank you.