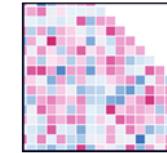


Optimizing Profit By Predicting Loan Default

Jay Cheong

Cost-Benefit Analysis on Classification is Essential

-  Predict the probability of loan default
-  Combine with Cost-Benefit Analysis (CBA) for profit optimization



kaggle

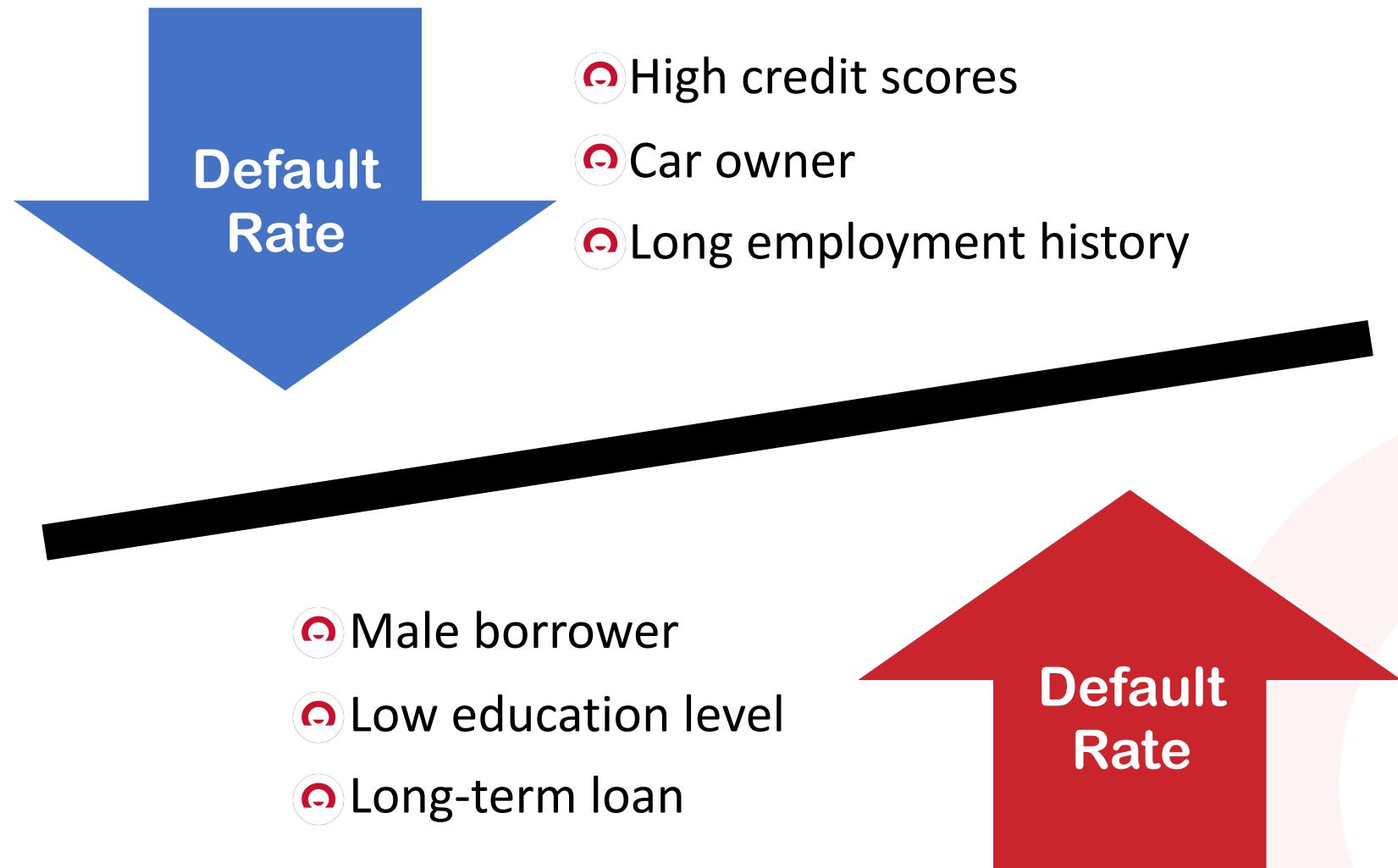
58 million rows of
imbalanced data



45 Engineered Features

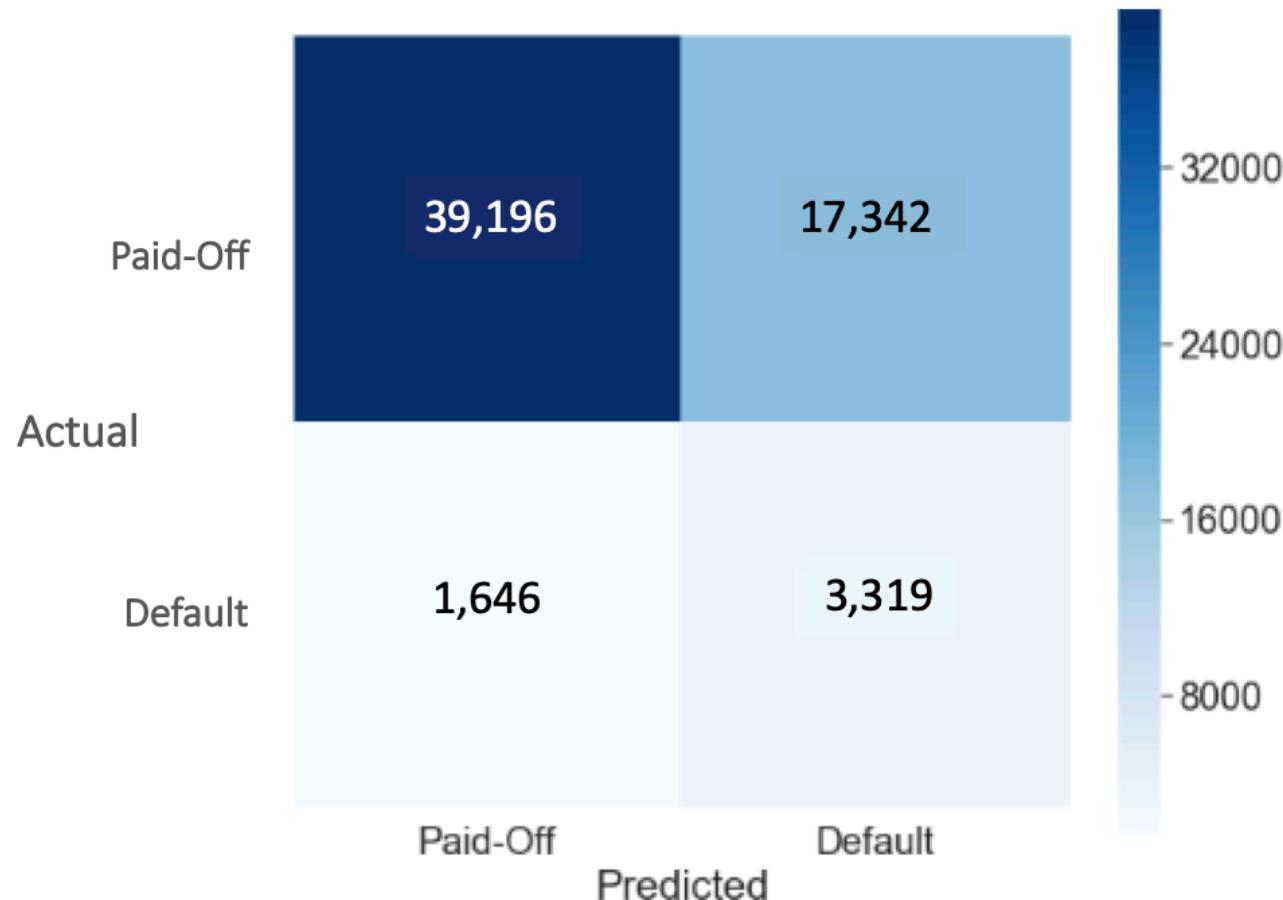
Logistic Regression
Bernoulli Naïve Bayes
Random Forest
XGBoost

There are Six Strong Predictors



Unlocking Full Benefit of Prediction with Cost-Benefit Analysis

AUC-ROC = 0.74



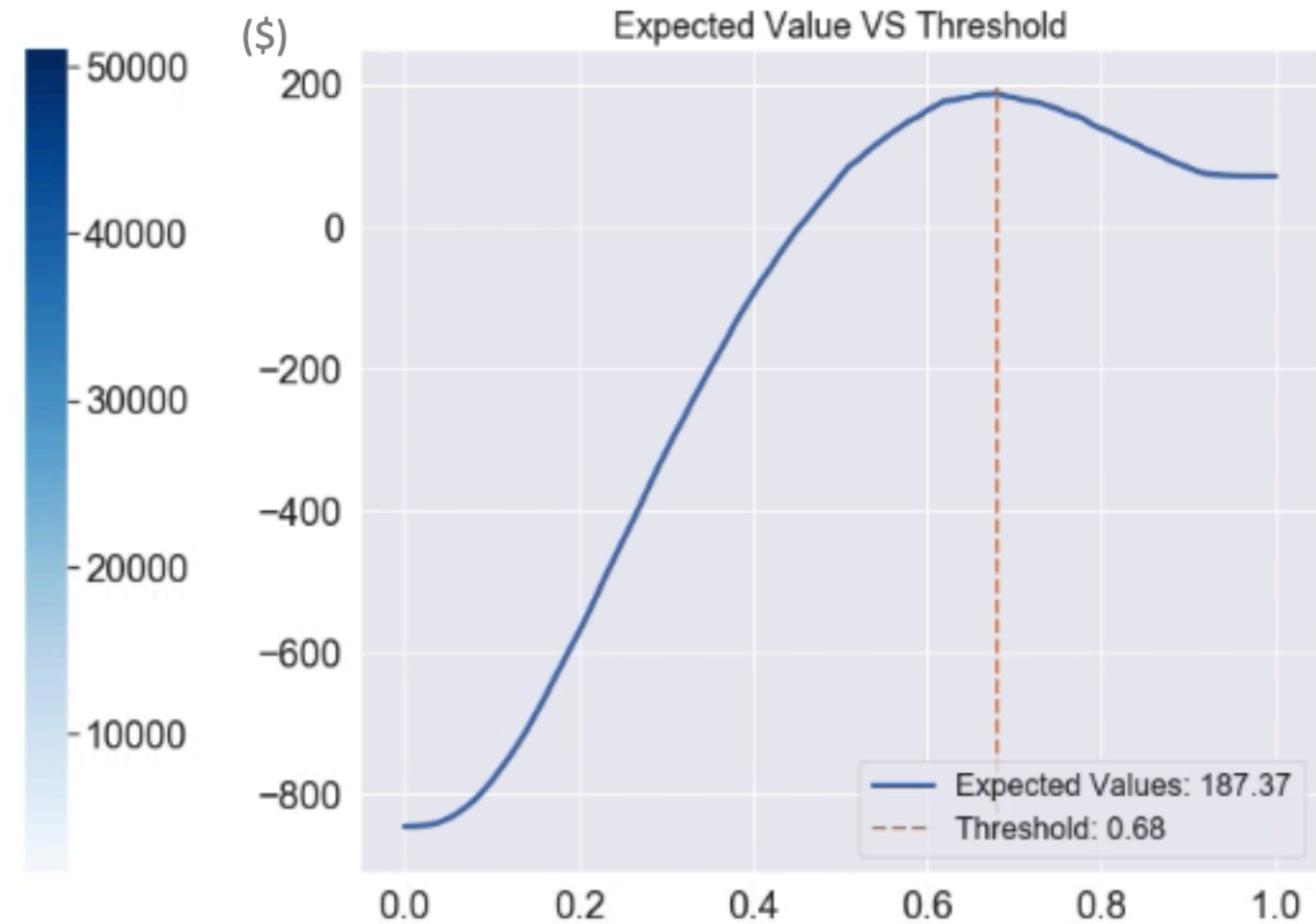
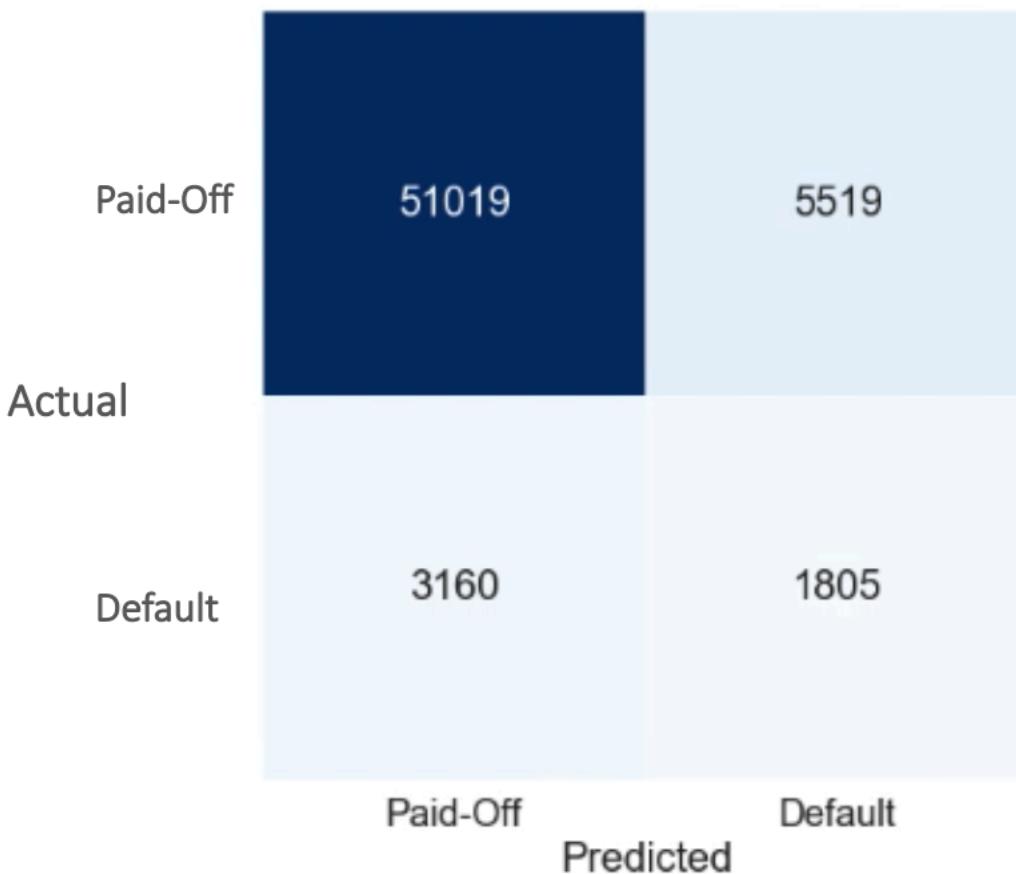
Cost-Benefit Analysis (CBA)

+ \$920	- \$920
x 64%	x 28%
- \$9,600	0
x 3%	x 5%

Expected Value = \$70

Optimum Expected Value is \$187 with threshold of 68%

threshold 0.68



Classification with CBA Beyond Lending Business

-  Medical device industry
 - Predict probability of device failure
-  Streaming service industry
 - Predict probability of churning
-  All industries
 - Predict probability of new hires quitting within the first year



Jay Cheong



yawjyec@yahoo.com



jay-cheong



yawjyec