

Protect Your Investment and Maximize Your Return

DO LEE, Data Scientist

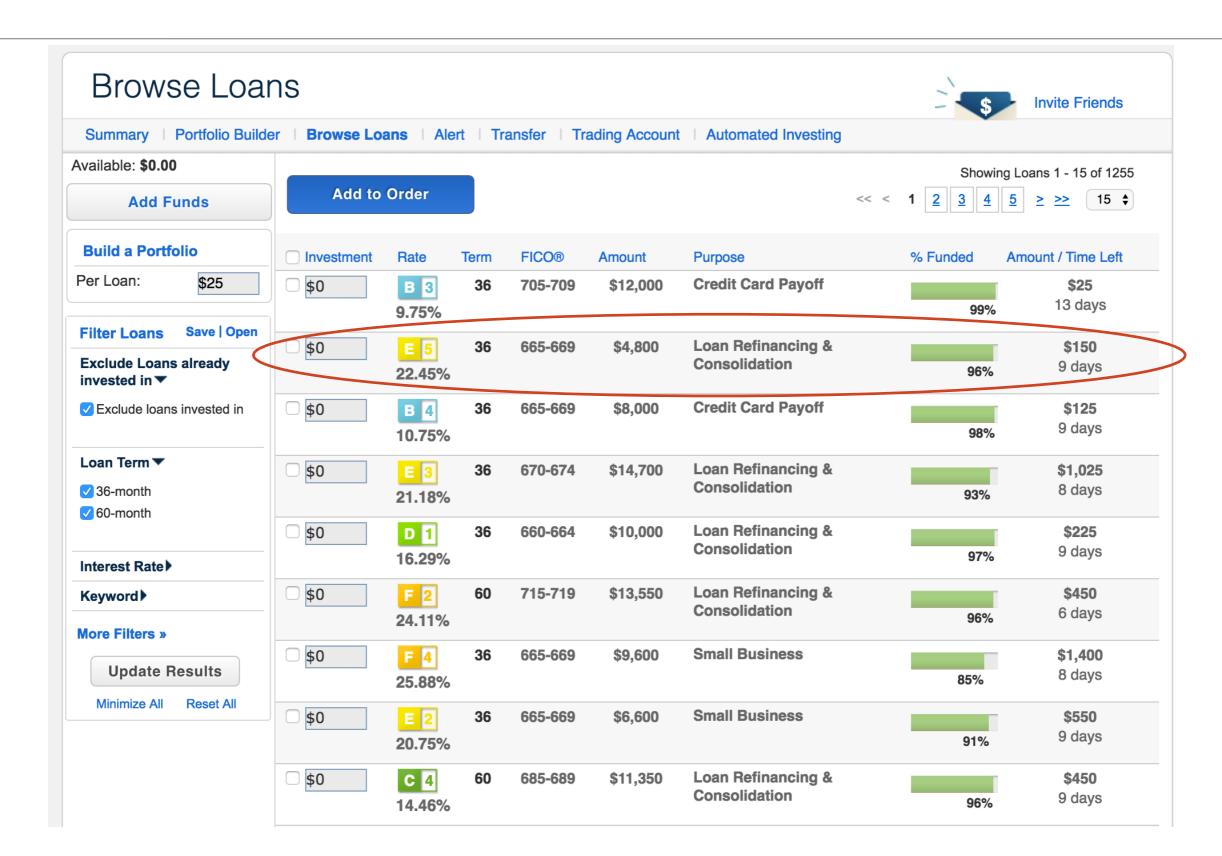
Lending Club: Peer-to-Peer Lending

Borrowers

Investors



Investor's View



Problem: Minimizing loan default risk

- Limited online features to gauge default risk
- Limited relevant metrics to maximize your ROI
- Example: 13% return 5% default 1% fees = 7%
- Goal: Minimize default risk and Maximize ROI

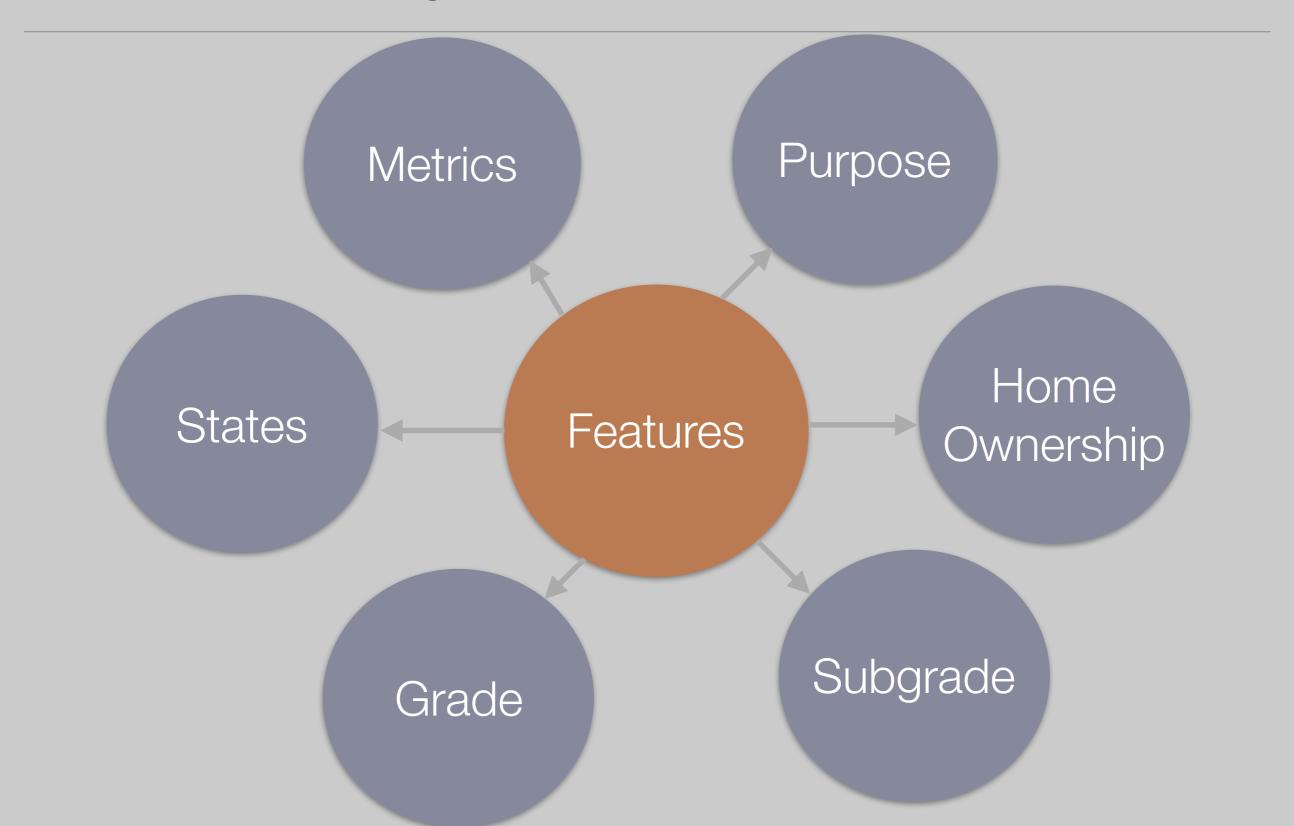
Solution: OptiReturn-Predictive Analytics Platform

- Classification Model: Logistic Regression
- Observations: Defaulted & Fully Paid Loans (~100k)
- Number of Features: 144
- Prediction Threshold: 0.50
- Year 2007 2014

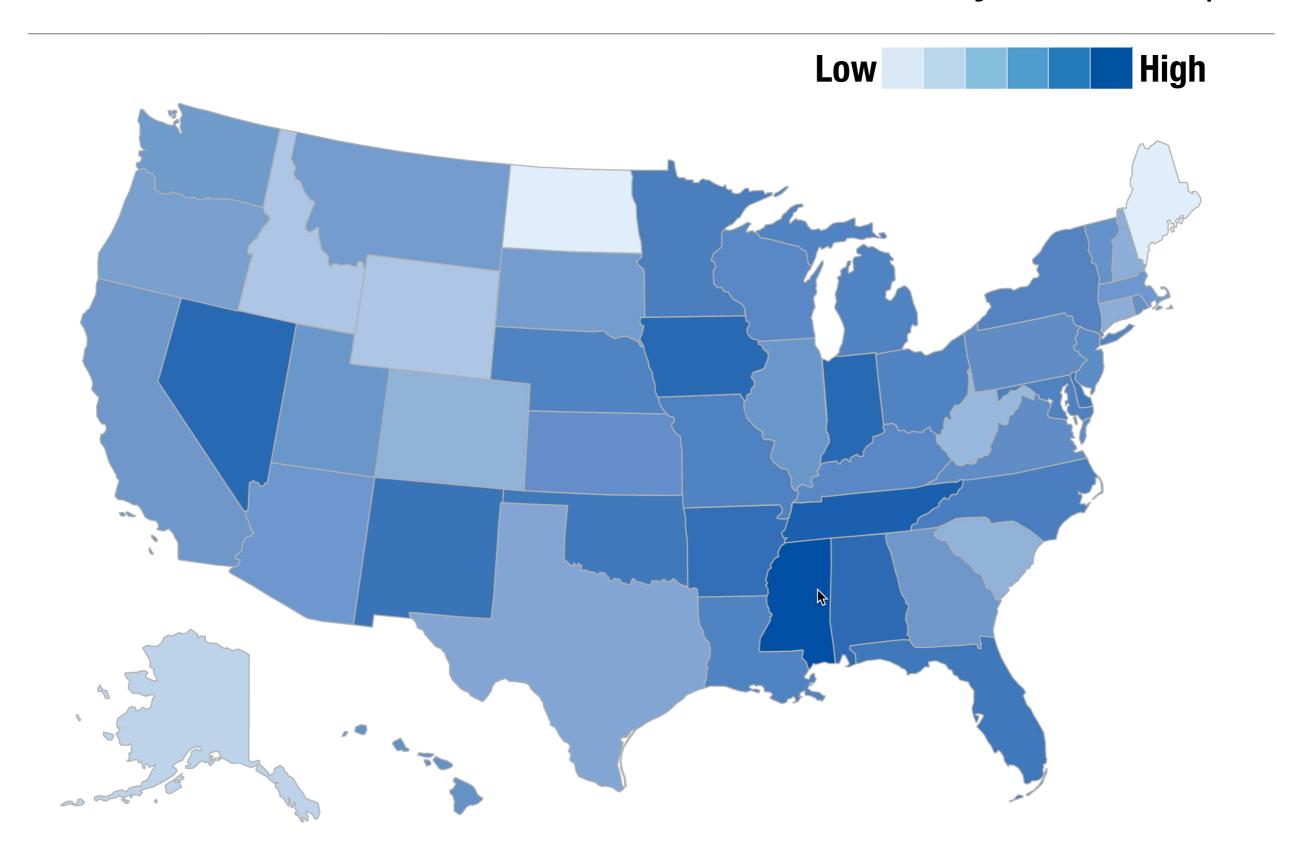
Predict Probability of Default

| | Default Probability | Paid Probability | Actual Status |
|------------|------------------------|---------------------|---------------|
| Borrower 1 | 0.73 | 0.27 | Defaulted |
| Borrower 2 | 0.35 | 0.65 | Fully Paid |
| Borrower 3 | 0.22 | 0.78 | Fully Paid |
| Borrower 4 | 0.81 | 0.19 | Defaulted |

Feature Groupings



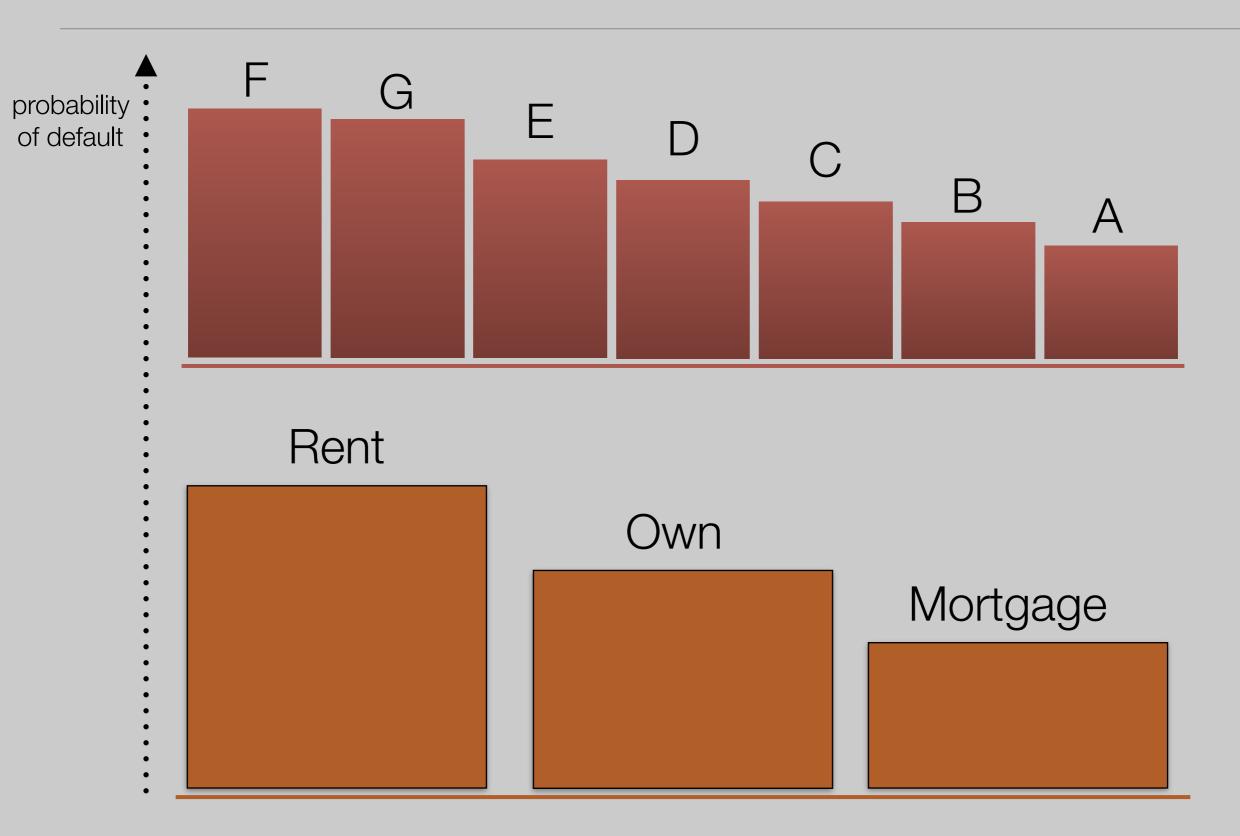
States: D3 Interactive Default Probability Heatmap



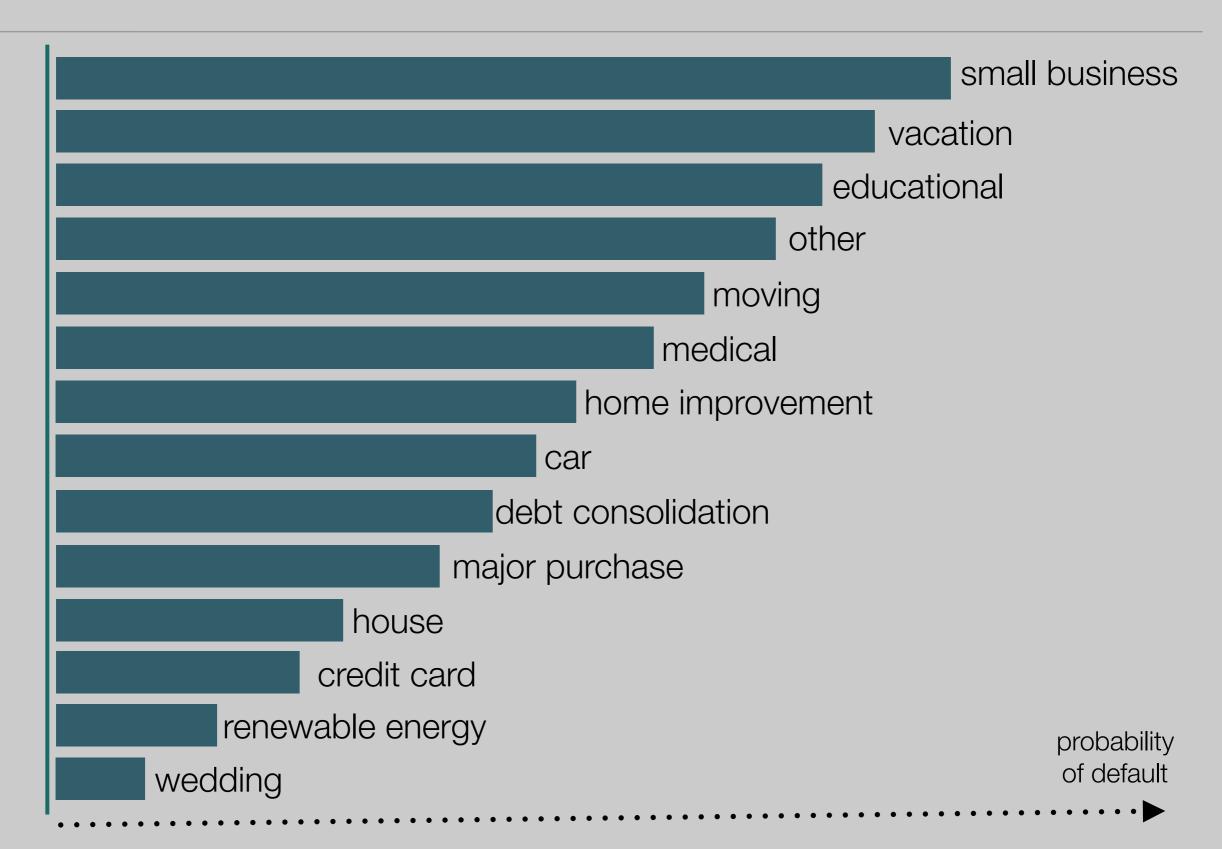
Metrics: Predictive Relevance to Default Risk

- FICO Score
- · Ratio: Monthly debt payment / Monthly income
- Total fee received to date
- · Ratio: Loan / Income
- Total number of credit lines
- % used in revolving credit line

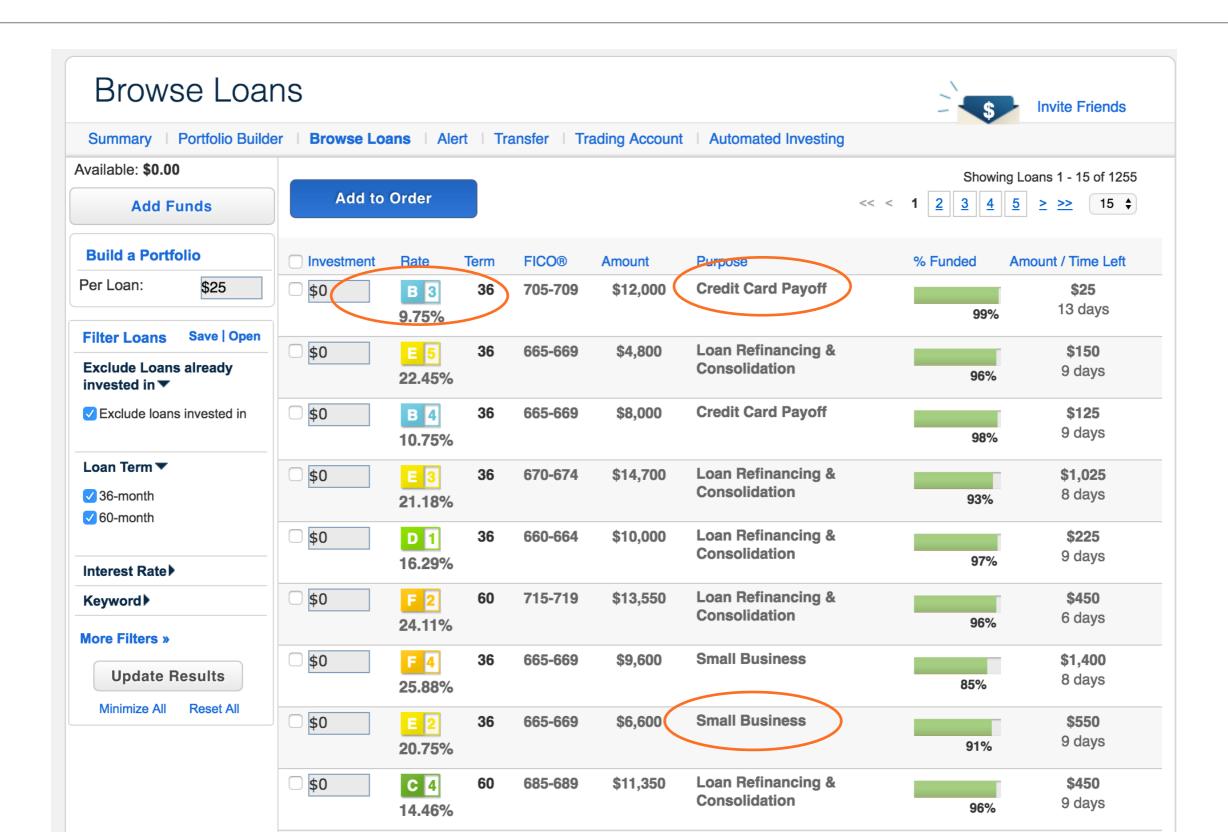
Grade & Home Ownership Probability of Default



Purpose Probability of Default



Enhance Your Understanding





Next Steps

- Custom feature engineering
- TF-IDF using borrowers' description as well as employment title
- Develop a web app to predict live using trained logistic regression algorithm
- Other applications:
 - Credit cards
 - Corporate bonds
 - Defaults on other financial instruments



www.github.com/doman1248/optireturn

APPENDIX

ROC Curve

