PROSPER MARKETPLACE INVESTMENTS

D. Chris Young

PROBLEM IDENTIFICATION

6.5%

ACHIEVE TOTAL PORTFOLIO RETURN OF AT LEAST 6.5% WITHIN THE NEXT 3 YEARS

ABOUT PROSPER MARKETPLACE

COMPANY BACKGROUND

Prosper is a peer-to-peer lending marketplace founded in 2005 that allows individuals to invest in loans.

HIGH RISK

Personal loans are unsecured and typically have higher interest rates to account for the higher risk of default.

HIGH RETURN

Historically have yielded return of 5.3% and offer diversification beyond traditional investments.

WHAT IS THE RISK?

A charge off occurs when the borrower fails to make their required loan payments.

The loan is closed and investors do not receive unpaid interest payments and principal amount.

18.5%

~1:5 PROSPER LOANS ARE NOT PAID BACK

WHAT FACTORS HAVE THE HIGHEST RISK?

27.8%

INTEREST RATE

New listings with an interest rate >= 15% have the highest default risk of 27.8%.

24.8%

LISTING TERM

Loans with a 60 month term have the highest default risk of 24.8%.

26%

DEBT TO INCOME

Borrowers with debit obligations of at least 55% of their income have a default risk of 26%.

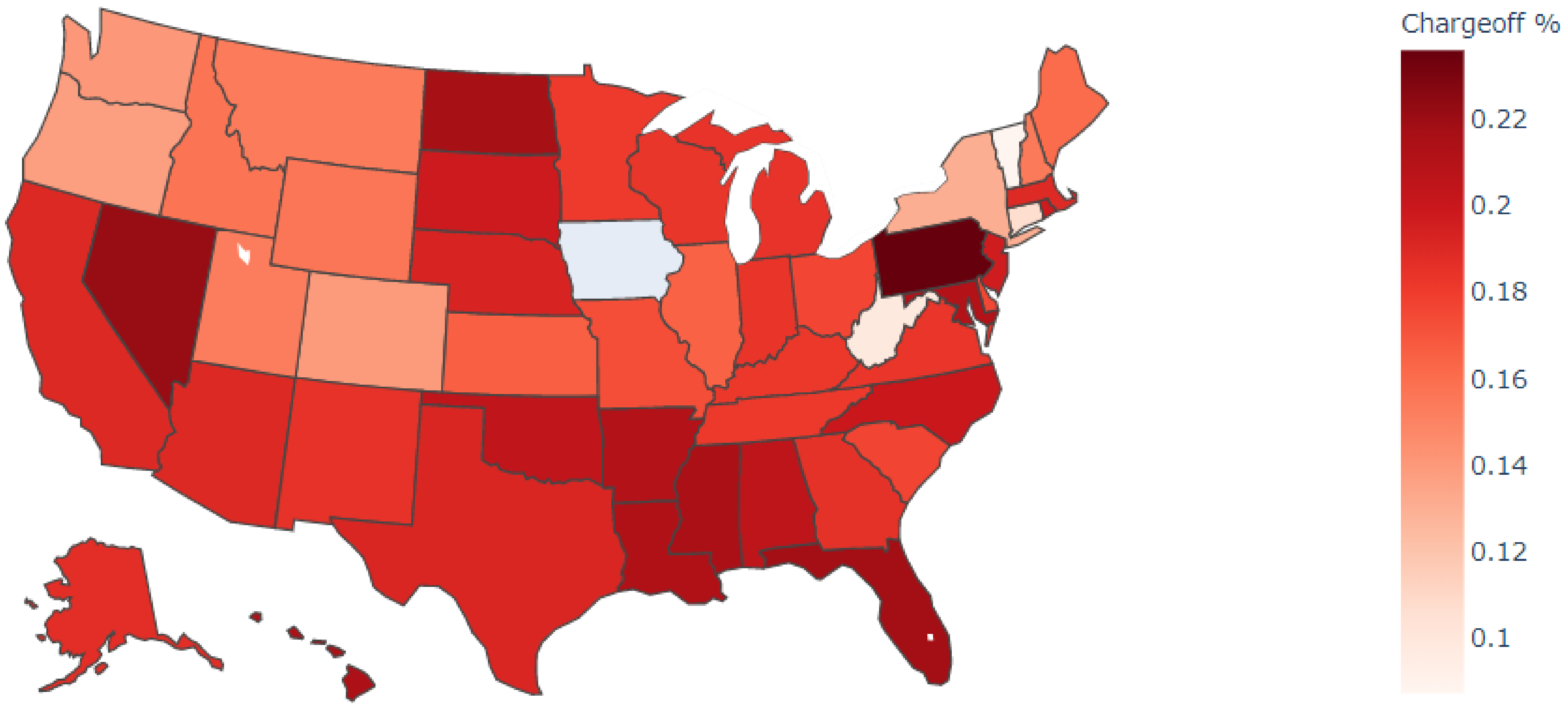
31.5%

PROSPER RATING

Prosper rates all new listings ranging from AA (highest) to HR (lowest). Ratings D-HR have a default risk of 31.5%.

DOES GEOGRAPHIC LOCATION DECREASE RISK?

Mountain region states, Pacific Northwest and West Virginia/New York have the lowest default risk of all geographic regions.



WHAT ARE THE CHARACTERISTICS OF A STRONG CREDIT PROFILE?

- No recent credit report inquires
- Low number of open trades on credit report
- No recent revolving/installment trades opened
- No 30 day past due accounts in the past 24 months

CAN WE MODEL CHARGE OFFS TO REDUCE DEFAULT RISK?

7 Classification Models were considered

Binary problem: 1 for charge off and 0 for paid off

Moderate imbalanced problem with 18% charge offs

Model performance was evaluated with an emphasis on the probability of charge off detection

DEFAULT MODEL PERFORMANCE

LGBM Classifier with balanced dataset performs best with fast training time

HIGHER F2 SCORE AND **LOWER** FALSE OMISSION RATE IS BETTER

FULL DATASET

Model	Training Time (Seconds)	F2 Score	False Omission Rate
Gaussian Naïve Bayes	0.41	0.469	0.141
XGB Classifier	15.59	0.165	0.183
LGBM Classifier	2.07	0.136	0.186
Gradient Boosting Classifier	92.14	0.101	0.19
AdaBoost Classifier	21.57	0.114	0.19
Logistic Regression	10.29	0.088	0.192
Random Forest Classifier	46.29	0.011	0.2

BALANCED DATASET

Model	Training Time (Seconds)	F2 Score	False Omission Rate
LGBM Classifier	1.01	0.58	0.103
Gradient Boosting Classifier	32.01	0.576	0.104
XGB Classifier	6.75	0.573	0.105
Random Forest Classifier	13.53	0.572	0.106
Logistic Regression	2.79	0.572	0.106
AdaBoost Classifier	8.83	0.569	0.107
Gaussian Naïve Bayes	0.13	0.538	0.121

FINAL MODEL SELECTION

LGBM Classifier with positive class weight tuning

.607

F2 SCORE

.055

FALSE OMISSION RATE

Parameter	Value
scale_pos_weight	10
colsample_bytree	0.964
learning_rate	0.0975
max_bin	262
max_depth	544
min_child_samples	25
num_leaves	46
subsample	0.746

SCENARIO ANALYSIS

-1.69%

RATE OF RETURN

RANDOM SELECTION

Randomly choosing investments is not a great strategy with 20% charge offs.

1.71%

RATE OF RETURN

RULE BASED SELECTION

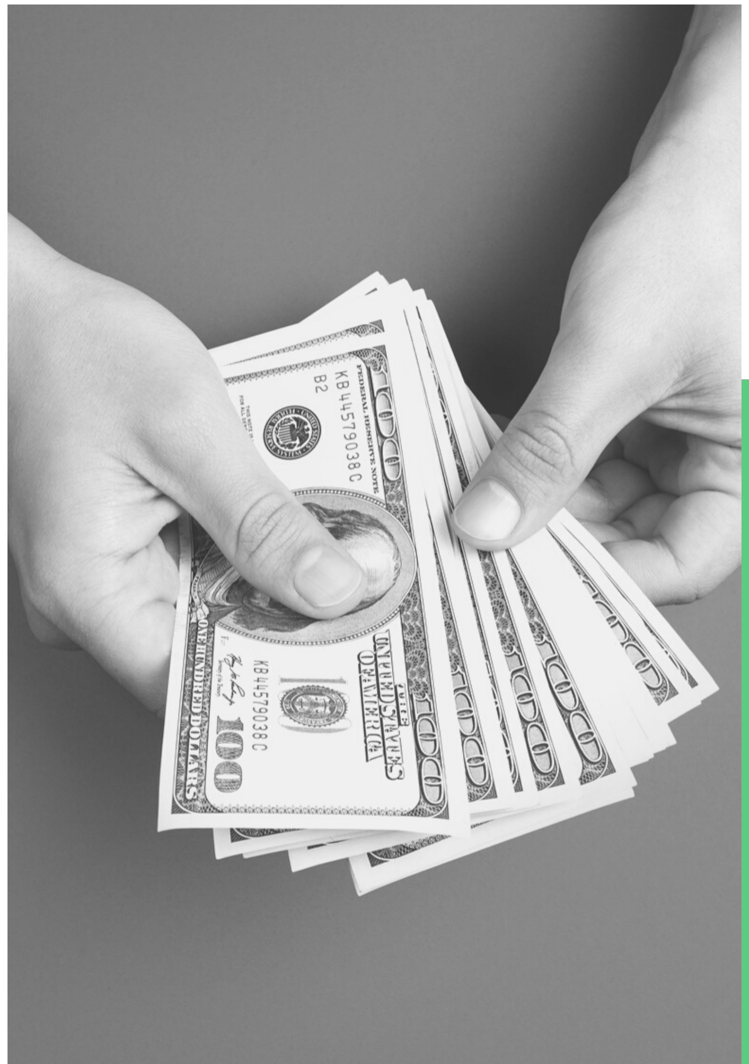
Rule based improves charge offs but selects only low yield investments.

7.25%

RATE OF RETURN

LOAN SCORECARD

Loan scorecard minimizes default risk exposure and maximizes rate of return.



FINAL SOLUTION

MODEL PREDICTIONS AND LOAN SCORECARD

7.25%

AVERAGE RATE OF RETURN

Model predictions reduces default risk exposure to approximately 5%

This allows investments to be selected from a pool of listings with a higher yield

An investment
in knowledge pays
the best interest.

BENJAMIN FRANKLIN