Predictive Analytics for Loan Repayment

Date: 2023

$ 11,017,023

DOLLARS

Rachel Roggenkemper: *rroggenk@calpoly.edu*

Matteo Shafer: *mshafe01@calpoly.edu*

Anagha Sikha: *arsikha@calpoly.edu*

Cameron Stivers: *ctstiver@calpoly.edu*

Sucheen Sundaram: *sssundar@calpoly.edu*

CalPoly

Authorized Signature

12345678 12345678 12345

Agenda

01 Data Overview

02 Summary of Models

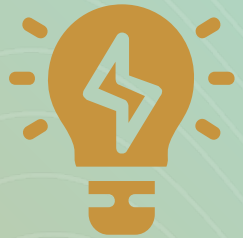
03 Final Model Implementation

04 Final Takeaways

Understanding the Task

Goal

Predict whether an applicant will be able to repay their loan



Objective

Analyzing patterns of previous applicants to understand what factors drive whether they will repay their loan

01

Data Overview

Data Overview

Data Source:

- Home Credit Group:
  - https://www.kaggle.com/competitions/home-credit-default-risk/data?select=application_train.csv

Data Preparation + Feature Engineering:

- Collapsed categories of occupation into broader categories
- Transformed variables in days into years

Variables:

Financial Profile		Asset Ownership	
Personal and Demographic Information		Loan Specifics	
Employment and Occupation Details		Housing Situation	

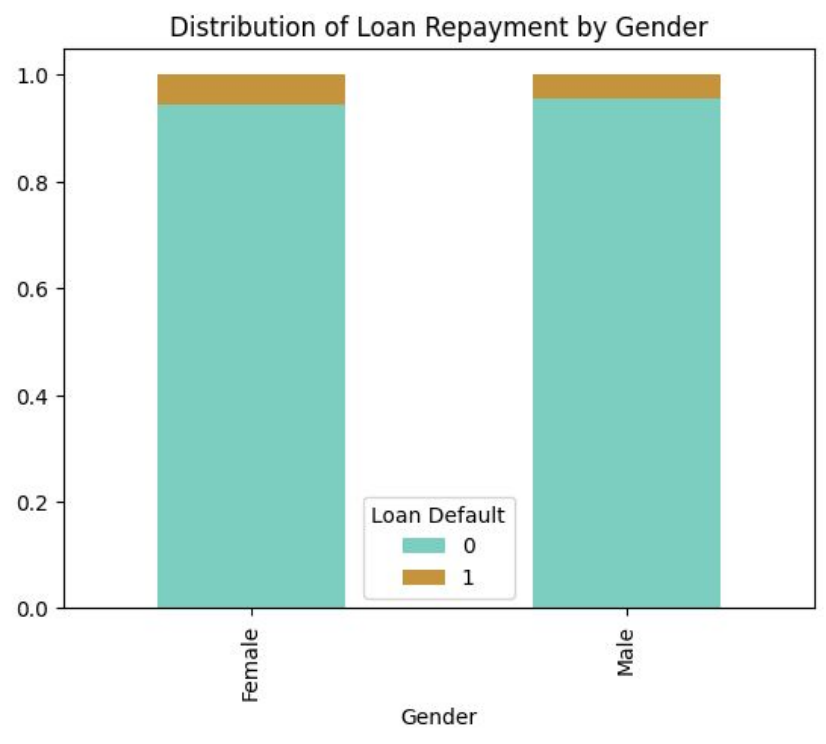
Glimpse of Dataset

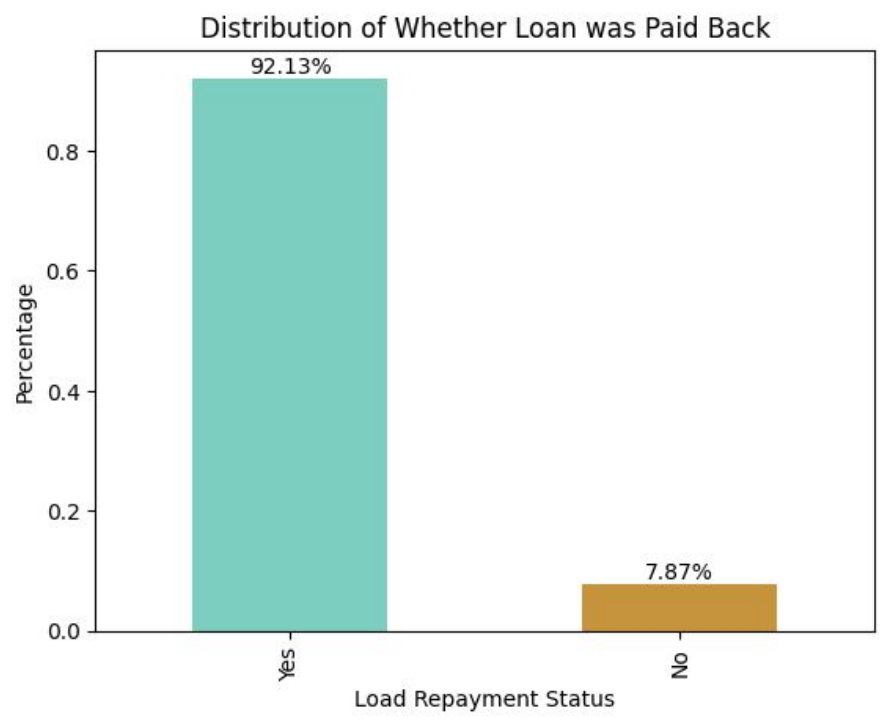
Target Variable:

0	Client with no payment difficulties
1	Client with payment difficulties

Target	AMT_INCOME_TOTAL	AMT_CREDIT	...	AMT_ANNUITY	CODE_GENDER
1	$202,500	$406,597.50	...	$24,700.50	M
0	$270,000	$1,293,502.50	...	$35,698.50	F
0	$67,500	$135,000	...	$6750	M

Data Exploration





02

Summary of Models

1) Logistic Regression
2) Support Vector Machines
3) Linear Discriminant Analysis

03

Final Model Implementation

Final Model: Logistic Regression

ROC-AUC	Accuracy	Recall	F1 Score
Percentage model correctly differentiates between the two classes	Percentage correctly classified	Percentage of loan defaults model correctly classified	Balance of precision and recall
59.7%	75.6%	27.6%	17.1%

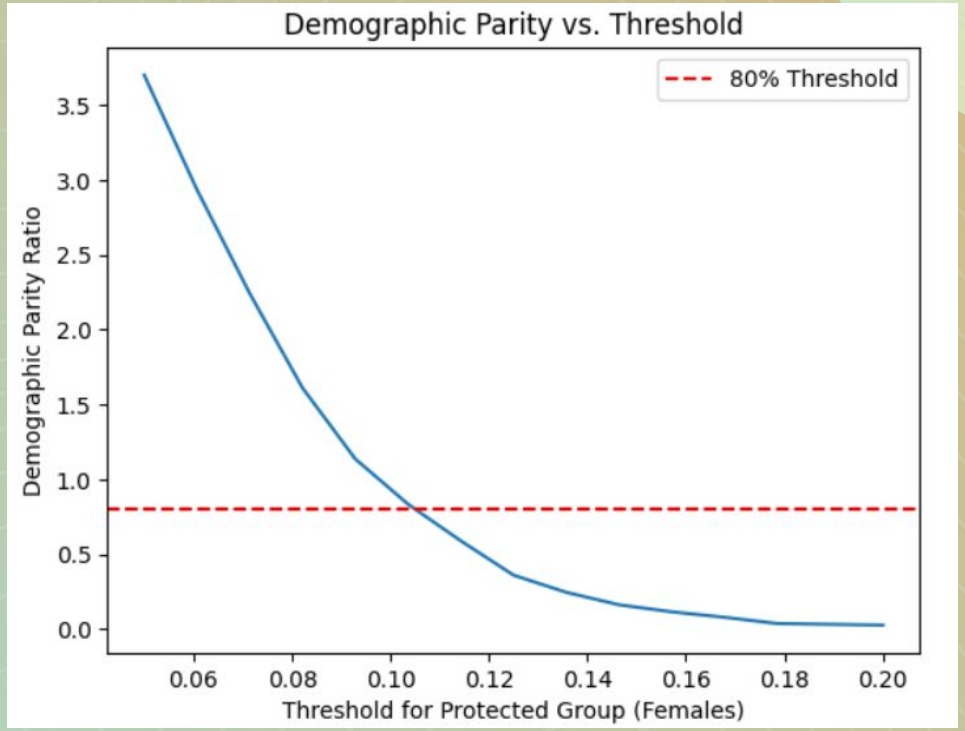
Fairness Metric: Demographic Parity

Gender:

Demographic Parity

The ratio of predicted loan defaults between males and females

1.04



Your Example Customer #1

Gender

Woman

Age

38

Occupation

CEO of Pritchett's Closets & Blinds

More Information

Education Type	Higher education
Income	$176,000
Amount of Credit	$100,000
Number of Children	3

Prediction:

Will repay her loan with no payment difficulties

Your Example Customer #2

Gender

Man

Age

26

Occupation

Temp at Dunder Mifflin

More Information

Education Type	Higher education
Income	$21,000
Amount of Credit	$200,000
Number of Children	0

Prediction:

Will not repay his loan and will have payment difficulties

04

Final Takeaways

Conclusions + Ethical Considerations

Our model demonstrates a reasonable capability in classifying loan defaults

Our model ensures that the loan approval process is equitable across different genders

Our model should be treated as an aid

Thanks!

Date: 2023

$ 11,017,023

DOLLARS

Do you have any questions?
Rachel Roggenkemper:*rroggenk@calpoly.edu*
Matteo Shafer: *mshafe01@calpoly.edu*
Anagha Sikha: *arsikha@calpoly.edu*
Cameron Stivers: *ctstiver@calpoly.edu*
Sucheen Sundaram: *ssundar@calpoly.edu*

CalPoly

12345678 12345678 12345

Authorized Signature