

Home Credit Default Risk



Outline

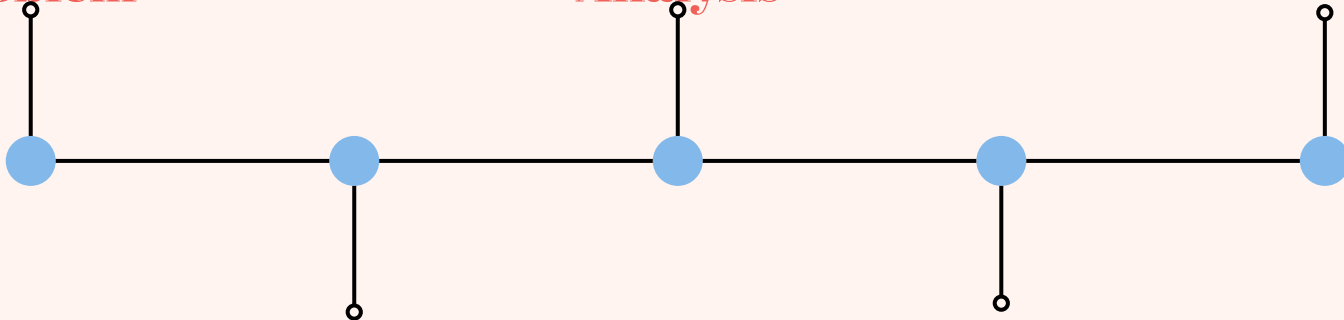
**Business
Problem**

**Feature
Analysis**

Recommendations

**Data &
Methods**

Model Results



Business Problem

Minimise the Default risk of Home Credit for Wells Fargo



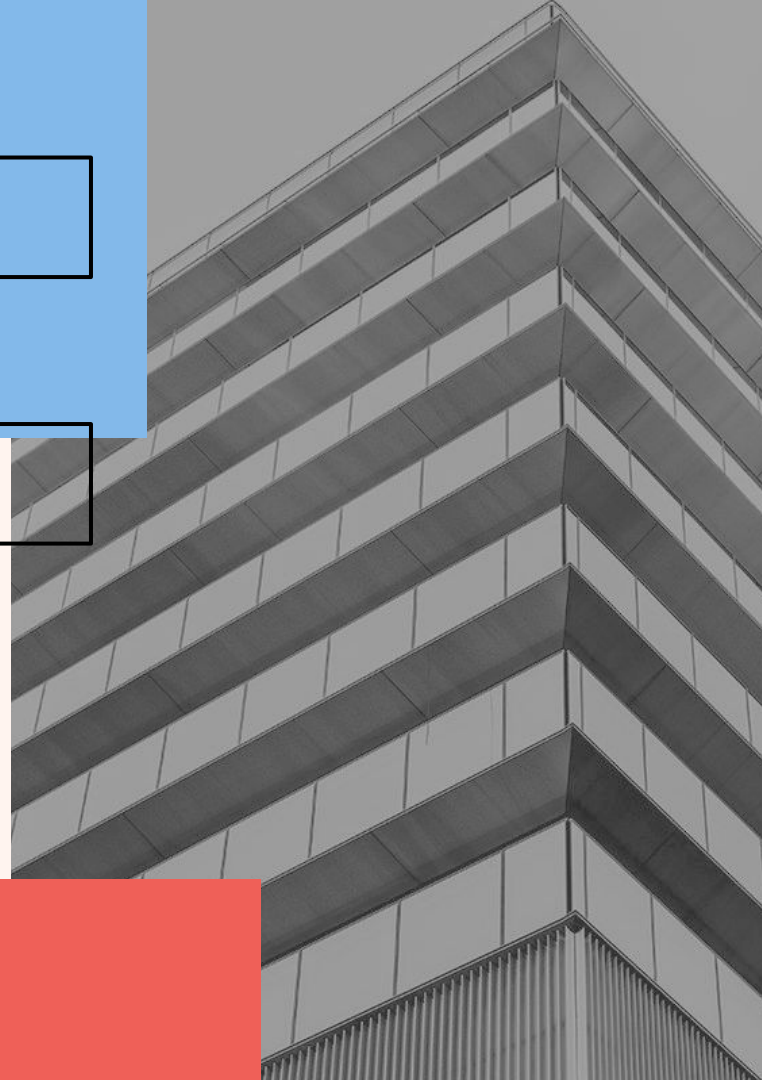
Data

- Kaggle

Methods

Four types of classifiers used-

- Logistic Regression
- Random Forest
- ADA Boosting
- Gradient Boosting



Feature Analysis

01

Gender

Male vs. Female

02

Loan Type

Cash vs Revolving loans

03

Clients' Age

04

Clients' Income

05

Type of Client

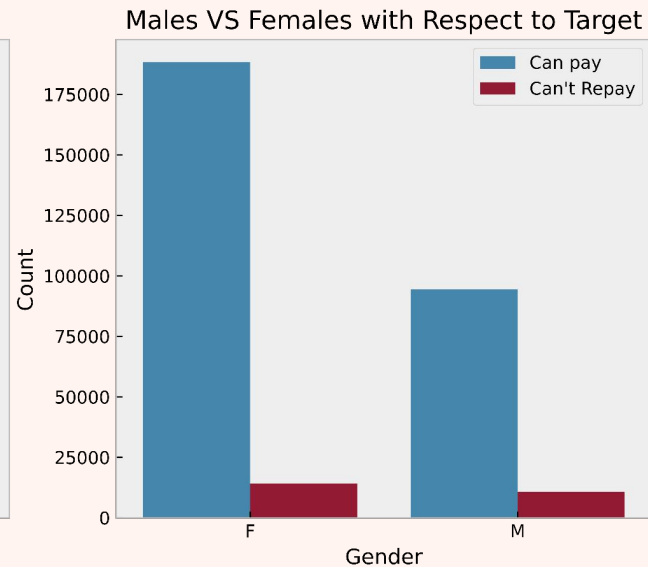
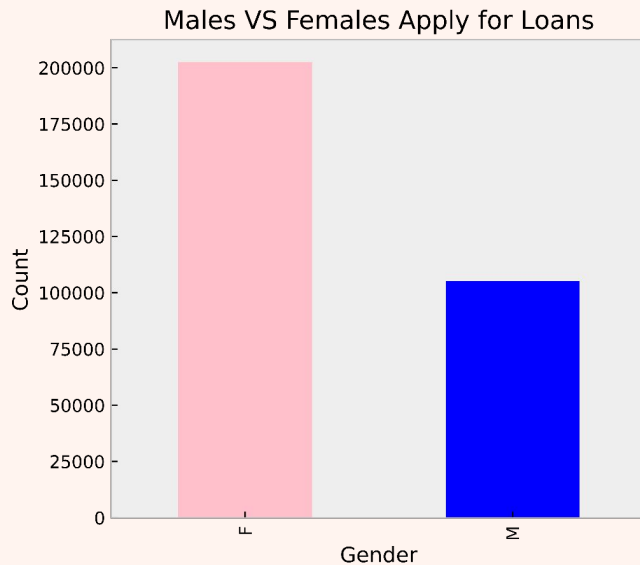
Working, State Servant,
Commercial associate,
Pensioner etc

06

Clients' employment

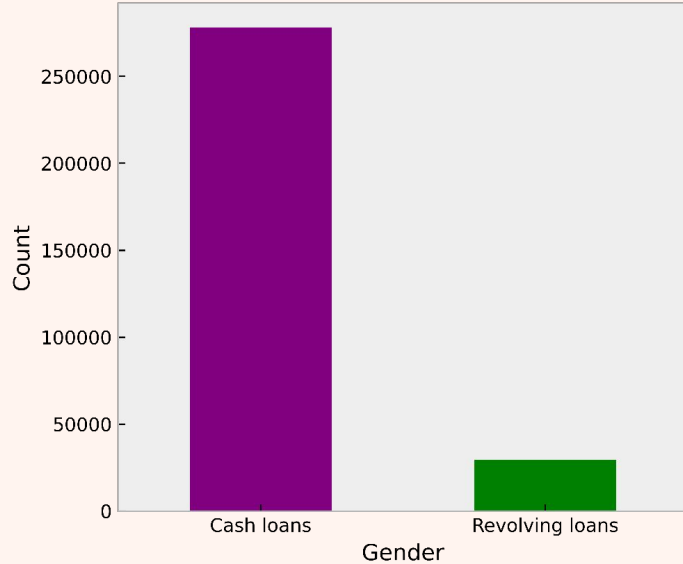
Gender

- ❑ Females apply for more loans than Males
- ❑ Default is not affected by the gender

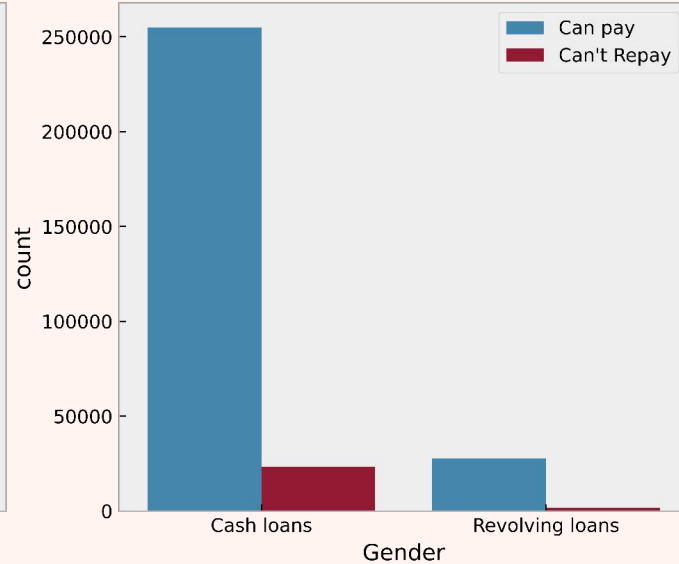


Loan Type

Cash VS Revolving Loans

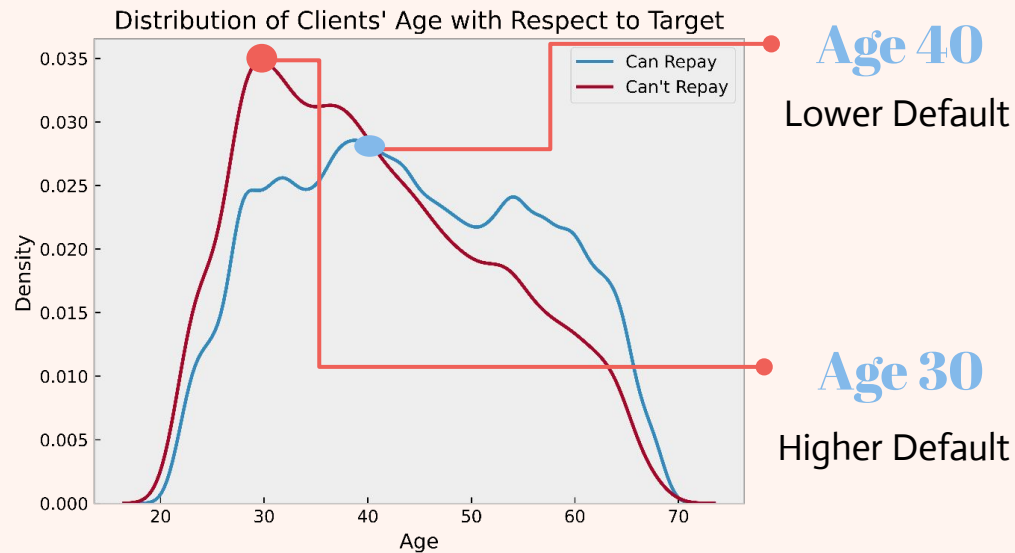
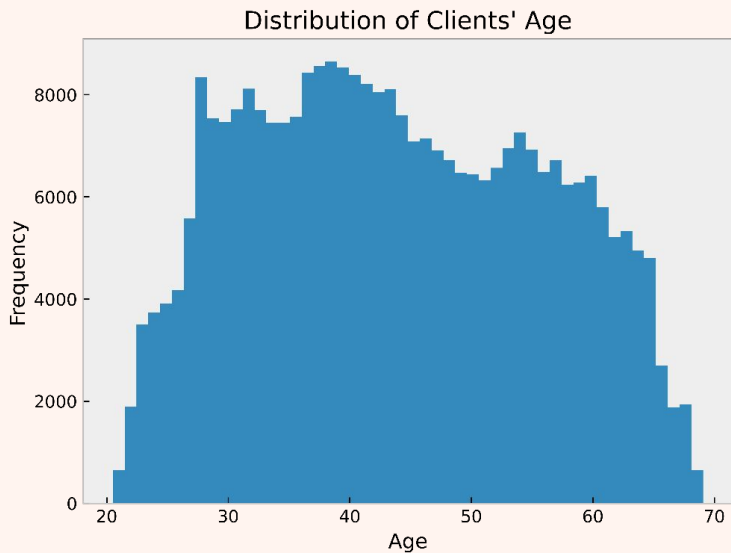


Cash VS Revolving Loans with Respect to Target



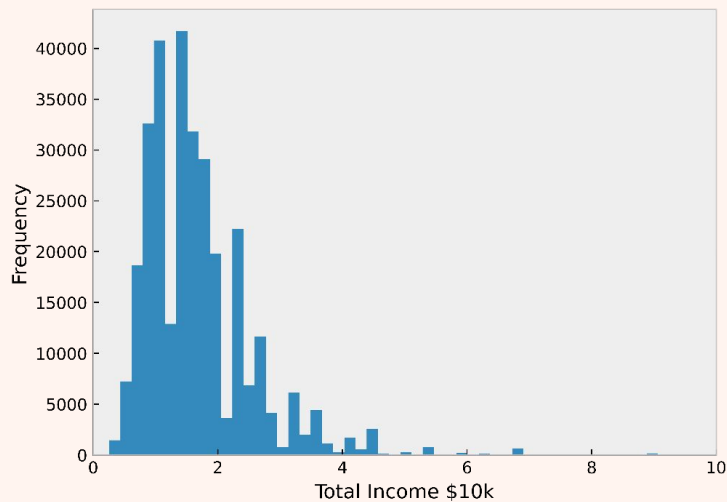
- ❑ Cash loans -more common
- ❑ Cash loans- higher default

Clients' Age



Clients' Income

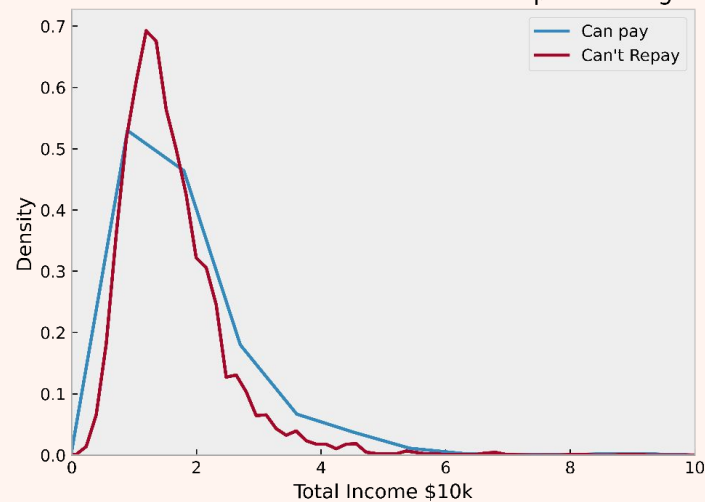
Distribution of Clients' Income



- **Lower income**

Apply more loans

Distribution of Clients' Income with Respect to Target

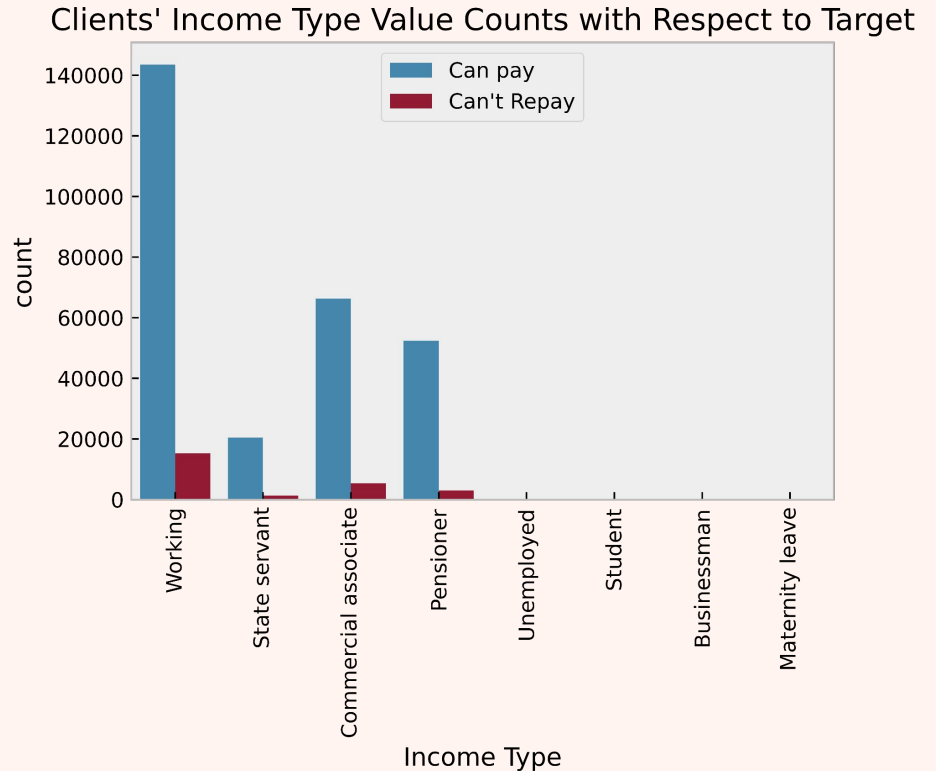


- **115k to 228k**

Higher Default

Employment Type

- ❑ Working clients - Most loan
- ❑ Businessmen and students - less count but always repay



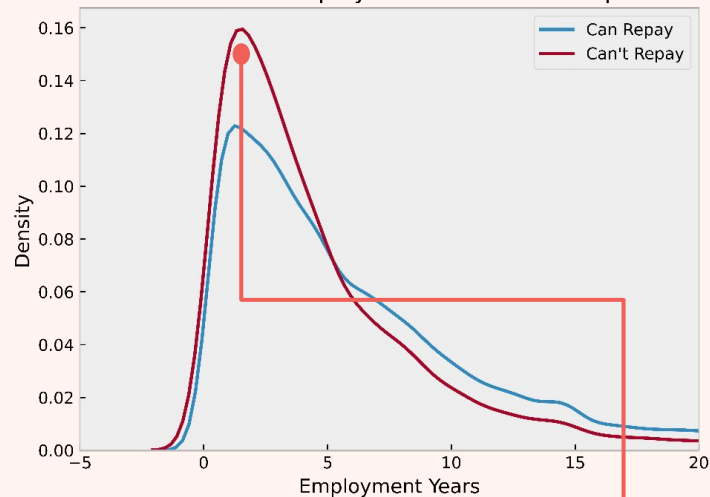
Employment Years



<5 years

More loans applications

Distribution of Clients' Employment Years with Respect to Target



<2 years

Higher Default

Model Results



Final Classifier method

Random Forest Classifier



AUC ROC

0.97



Generalized model

Model built for new and old clients

Recommendations



Revolving loans

Aim to increase the amount of revolving loans



Age

Target clients older than 40 years



Employment

Focus on clients-

- ☐ Working
- ☐ Businessmen
- ☐ Students

Thank you!

Any Questions?

Alice Agrawal | alice.agrawal3o@gmail.com

