Who will default on their credit card payments?

...

Joe Schwehr

Introduction

Why do we care?

Our dataset

Exploratory Data Analysis

Feature Engineering

Predictive Modeling

Results

Which customers will default on their credit card payment?

Why is this useful?

- The company may be able to intervene with certain customers to help increase repayment
- The company will know which customers are unable to pay their bill regardless of extra attention by the company.
- Customers that will pay their account as agreed upon do not need to be actively targeted for repayment or contacted.
- Knowing this information will save the company money.

The Data

We will use a UCI credit card dataset to make our prediction.

- It consists of 30,000 records
- 23 explanatory variables
  - Age
  - Gender
  - Payment History
  - Education
- 1 target variable
  - Default payment next month
- 22% is the current rate of defaults (unbalanced data)

Exploratory Data Analysis

- Most customers pay on time
- Most customers have a balance under 100,000 (New Taiwanese Dollars)
- Marital status doesn't have an obvious effect on late payments
- Customers in their 20s have the most late payments
- Males have a slightly higher percent of late payments
- Late payments decrease as a customer's education increases
- Customers with a limit of under 100,000 have the most late payments
- Customers with a bill between 10,000 and 40,000 have the most late payments

Feature Engineering

Adding columns to the dataset

- Whether a customer has a late payment in the past 6 months
- Current balance as a proportion of credit limit

Predictive Modeling

This is a binary classification problem; will a customer default next month: yes/no

Good methods for modeling our problem include:

- Logistic Regression
- Etc
- ...
- ..
- .

Results

Coming soon

Data-Driven Decisions

Conclusion

- Defining the problem
- Our methods
- The results
- Data-Driven Decisions