

Modeling loan acceptance

Miguel Ángel Canela

Associate Professor, IESE Business School

Introduction

The objective of this project is to develop a model for predicting the acceptance of an offer of a personal loan by the customer of a bank. Universal Bank is a relatively young bank growing rapidly in terms of overall **customer acquisition**. The majority of these customers are liability customers (depositors) with varying sizes of relationship with the bank. The customer base of asset customers (borrowers) is quite small, and the bank is interested in expanding this base rapidly to bring in more loan business. In particular, it wants to explore ways of converting its liability customers to **personal loan customers** (while retaining them as depositors).

A campaign that the bank ran last year for liability customers showed a **conversion rate** of over 9% success. This has encouraged the retail marketing department to devise smarter campaigns with better target marketing. The goal of our analysis is to model the previous campaign's customer behavior to analyze what combination of factors make a customer more likely to accept a personal loan. This will serve as the basis for the design of a new campaign.

The data set

The data set (file `loan.csv`) contains data on 5,000 customers, including customer demographic information, the customer's relationship with the bank and the customer response to the last personal loan campaign. Among these 5,000 customers, only 480 (9.6%) accepted the personal loan that was offered to them in the earlier campaign.

The variables are:

- The customer's age (**age**).
- The customer's family size (**family**).
- The income of the customer in thousand GBP per year (**income**).
- The educational level (**education**): 1 = Undergraduate, 2 = Graduate, 3 = Professional.
- The average monthly credit card spending in thousand GBP per year (**ccavg**).
- The size of the (eventual) mortgage in thousand GBP per year (**mortgage**).
- A dummy for having a securities account (**secacc**).
- A dummy for having a certificate of deposit account (**cdacc**).
- A dummy for operating on-line with the bank (**online**).
- A dummy for having a credit card (**ccard**).
- A dummy for having accepted the personal loan offer (**loan**).

Source (slightly edited): G Shmueli, PC Bruce, I Yahav, NR Patel & KC Lichtendahl (2018), *Data Mining for Business Analytics*, Wiley.

Questions

- Q1.** Develop a model for predicting loan acceptance, based on a logistic regression equation.
- Q2.** How is the distribution of acceptance scores? Is it different for the actual positives and the actual negatives?
- Q3.** Set an adequate cutoff for the acceptance scores and apply it to decide which customers are potential borrowers. What is the true positive rate? And the false positive rate?
- Q4.** The bank is planning to continue its campaign by sending its offer to 1000 additional customers. Suppose that the cost of preparing and sending the offer is \$2 and the profit from an accepted offer is \$50. What is the expected profitability of this campaign?
- Q5.** To maximize the expected profitability, should the cutoff be increased or decreased?