

IOE 373 F23 Final Project

Due Tuesday, December 5, 11:59 PM

Universal Bank is a relatively young bank that is growing rapidly in terms of overall customer acquisition. The bank is interested in growing its customer base rapidly to bring in more loan business. In particular, it wants to explore ways of converting customers to personal loan customers.

A campaign the bank ran for customers showed a 9% successes. This has encouraged the retail marketing department to devise smarter campaigns with better target marketing. The goal of your 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 bank's dataset includes data on 5000 customers. The data include customer demographic information (age, zip code, etc.), customer financial information (income, mortgage, etc.) and a list of customer accounts with the bank (credit cards, personal loan, etc.) The customer response to the last personal loan campaign (Personal Loan) is also included in the customer account line table as a “**Personal Loan**” account type.

The data is included in the attached Access Data Base. Field descriptions for each of the Tables in the database is as follows:

Table 1. Customer Demographics

CustomerID	Customer ID
Age	Customer's age in completed years
ZIPCode	Home Address ZIP code.
Family	Family size of the customer
Experience	#years of professional experience
Education	Education Level. 1: Undergrad; 2: Graduate; 3: Advanced/Professional

Table 2. Customer Financials

CustomerID	Customer ID
Income	Annual income of the customer (\$000)
CCAvg	Avg. spending on credit cards per month (\$000)
Mortgage	Value of house mortgage if any. (\$000)

Table 3. Customer Bank Accounts

AccountID	AccountID
CustomerID	Customer ID
AccountType	Four Types of Account a customer can have: Personal Loan , Securities Account, CD Account, Online, CreditCard

The goal of your 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.

- a) Organize your data into a Final Data Analysis table so that you can run a prediction model (e.g. logistic regression). You will need to use SQL to create indicator variables for some of the variables. For example:

$$\begin{aligned}\text{EducProf} &= \begin{cases} 1 & \text{if education is Professional} \\ 0 & \text{otherwise} \end{cases} \\ \text{EducGrad} &= \begin{cases} 1 & \text{if education is at Graduate level} \\ 0 & \text{otherwise} \end{cases} \\ \text{Securities} &= \begin{cases} 1 & \text{if customer has securities account in bank} \\ 0 & \text{otherwise} \end{cases} \\ \text{CD} &= \begin{cases} 1 & \text{if customer has CD account in bank} \\ 0 & \text{otherwise} \end{cases} \\ \text{Online} &= \begin{cases} 1 & \text{if customer uses online banking} \\ 0 & \text{otherwise} \end{cases} \\ \text{CreditCard} &= \begin{cases} 1 & \text{if customer holds Universal Bank credit card} \\ 0 & \text{otherwise} \end{cases}\end{aligned}$$

Hint: You should have a total of 5000 records in your Final Data Analysis table, corresponding to the 5000 customers in the Customer Demographics and Customer Financials Tables. Note that not all customers have current accounts and a customer may have multiple accounts as shown in the table Customer Bank Accounts.

- b) Export your Final Data Analysis table to python and run at least one classification model (e.g. logistics regression). Partition your data set into 2 sets (training and validation sets) at a 50-50 split (e.g. 50% of random observations for the training set and 50% for the test set). Evaluate your model and identify significant factors. Validate using the test set.

Deliverables:

1. Brief written report (1-2 pages max) including description of your model (what factors seem the most important?), validation results (how well does your model predict the target variable?) and conclusions (what type of customers should be the focus of the marketing campaign?). You can include tables and/or figures in your report (no appendices)
2. Supporting files showing your queries and analyses.