**Bank Loan Modelling**

Context

Bank: Thera Bank

Most depositors with varying size of deposits.

The number of customers who are also borrowers (asset customers) is quite small.

the bank is interested in expanding this base rapidly to bring in more loan business and in the process, earn more through the interest on loans.

In particular, the management 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 healthy conversion rate of over 9% success.

This has encouraged the retail marketing department to devise campaigns to better target marketing to increase the success ratio with a minimal budget.

The department wants to build a model that will help them identify the potential customers who have a higher probability of purchasing the loan. This will increase the success ratio while at the same time reducing the cost of the campaign.

Content

Column descriptions

ID

Customer ID

Age Customer's age in completed years’

Experience #years of professional experience

Income Annual income of the customer ($000)

ZIPCode Home Address ZIP code.

Family Family size of the customer

CCAvg : Avg. spending on credit cards per month ($000)

Education Level. 1: Undergrad; 2: Graduate; 3: Advanced/Professional

Mortgage Value of house mortgage if any.

($000) Personal Loan Did this customer accept the personal loan offered in the last campaign?

Securities Account Does the customer have a securities account with the bank?

CD Account Does the customer have a certificate of deposit (CD) account with the bank?

Online Does the customer use internet banking facilities?

Credit Card Does the customer use a credit card issued by Universal Bank?

The data set includes 5000 observations with fourteen variables divided into four different measurement categories.

The binary category has five variables, including the target variable personal loan, also securities account, CD account, online banking, and credit card.

The interval category contains five variables: age, experience, income, CC avg and mortgage.

The ordinal category includes the variables family and education.

The last category is nominal with ID and Zip code.

The variable ID does not add any interesting information e.g., individual association between a person (indicated by ID) and loan does not provide any general conclusion for future potential loan customers.

Therefore, it will be neglected in the examination.

Problem Statement.

Study the data distribution in each attribute, share your findings.

Use a classification model to predict the likelihood of a liability.

customer buying personal loans.