greatlearning
Power Ahead

Project on Supervised Learning

Learning Objectives

- Data Description
- Attribute information
- Steps to follow
- Conclusion

Data Description

Campaign for selling personal loans.

This case is about a bank (Thera Bank) which has a growing customer base. Majority of these customers are liability customers (depositors) with varying size of deposits. The number of customers who are also borrowers (asset customers) is quite small, and 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.

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

The file Bank.xls contains data on 5000 customers. The data include customer demographic information (age, income, etc.), the customer's relationship with the bank (mortgage, securities account, etc.), and the customer response to the last personal loan campaign (Personal Loan). Among these 5000 customers, only 480 (= 9.6%) accepted the personal loan that was offered to them in the earlier campaign.

Body here

Attribute information

- ID : Customer ID
- Age : Customer's age in completed years
- Experience : #years of professional experience
- Income : Annual income of the customer ($000)
- ZIP Code : Home Address ZIP code.
- Family : Family size of the customer
- CCAvg : Avg. spending on credit cards per month ($000)
- Education : 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?

Steps to follow

- Import the libraries
- Get the data
- Find the types of those variables
- Calculate summary
- If any missing values or negative values, remove them
- Compare all attributes visually to check for relationships that can be exploited by making plots
- Check for correlated variables
- Separate the independent attributes
- Prepare the data for modelling
- Build the model using logistic regression and choose the best metric to evaluate it.
- Summarize the fit of the model and give business insights.

greatlearning
Power Ahead

Happy Learning !