Banking Prediction Dataset

Introduction

This is a dataset from a Bank with approximately 100.000 of non credit card holders in a tab delimited text file. Product balances & transactions are the main attributes of this dataset.

The dependent variable is the spontaneous activation of a new credit card. All customer attributes have been collected before the activation of the new credit card.

It can be used for deriving new variables, descriptive analysis, association analysis and propensity modeling for finding new target groups for new credit cards promotion,

Fields Description

THIS IS A TAB DELIMITED TEXT FILE

Tenure is the oldness of customer in months

ounts are the average monthly balances

<channel> transactions are the average monthly transactions

New credit card flag refers to a spontaneous activation of a new credit card by previously non credit cards holders. (1: Yes, 0: No)

Analysis Guidelines

Convert cproduct> amounts into flags in order to derive new variables such number of products and perform association analysis

Divide cproduct> amounts to total amount in order to derive cproduct> percentages

Divide <channel> transactions to total transactions in order to derive <channel> percentages

Use multiple propensity modeling techniques such logistic regression and classification trees to estimate the propensity for a non credit card holder to activate a new credit card

You can use the book '<u>Data Mining Techniques in CRM</u>' for more tips and tricks

Send us your conclusions to info@customers-dna.com



Applications

- Simple Data Management
- Descriptive Statistics
- Association Analysis
- Propensity Modeling



