

# Marketing Data Mining

**Victor S.Y. Lo**

*Fidelity Personal Investments, USA*

## INTRODUCTION

Data mining has been widely applied over the past two decades. In particular, marketing is an important application area. Many companies collect large amounts of customer data to understand their customers' needs and predict their future behavior. This article discusses selected data mining problems in marketing and provides solutions and research opportunities.

## BACKGROUND

Analytics are heavily used in two marketing areas: market research and database marketing. The former addresses strategic marketing decisions through the analysis of survey data, and the latter handles campaign decisions through the analysis of behavioral and demographic data. Due to the limited sample size of a survey, market research normally is not considered data mining. This article focuses on database marketing, where data mining is used extensively to maximize marketing return on investment by finding the optimal targets. A typical application is developing response models to identify likely campaign responders. As summarized in Figure 1, a previous campaign provides data on the dependent variable (responded or not), which is merged with individual characteristics, including behavioral and demographic variables, to form an analyzable data set. A response model is then developed to predict the response rate given the individual characteristics. The model is then used to score the population and predict response rates for all individuals. Finally, the best list of individuals will be targeted in the next campaign in order to maximize effectiveness and minimize expense.

Response modeling can be applied in the following activities (Peppers & Rogers, 1997, 1999):

- (1) **Acquisition:** Which prospects are most likely to become customers?
- (2) **Development:** Which customers are most likely to purchase additional products (cross-selling) or to add monetary value (up-selling)?
- (3) **Retention:** Which customers are most retainable? This can be relationship or value retention.

## MAIN THRUST

Standard database marketing problems have been described in literature (e.g., Hughes, 1994; Jackson & Wang, 1996; and Roberts & Berger, 1999). In this article, I describe the problems that are infrequently mentioned in the data mining literature as well as their solutions and research opportunities. These problems are embedded in various components of the campaign process, from campaign design to response modeling to campaign optimization; see Figure 2. Each problem is described in the Problem-Solution-Opportunity format.

## CAMPAIGN DESIGN

The design of a marketing campaign is the starting point of a campaign process. It often does not receive enough attention in data mining. If a campaign is not designed properly, postcampaign learning can be infeasible (e.g., insufficient sample size). On the other hand, if a campaign is scientifically designed, learning opportunities can be maximized.

The design process includes activities such as determining the sample size for treatment and control groups (both have to be sufficiently large such that measurement and modeling, when required, are feasible), deciding on sampling methods (pure or stratified random

*Figure 1. Response modeling process*

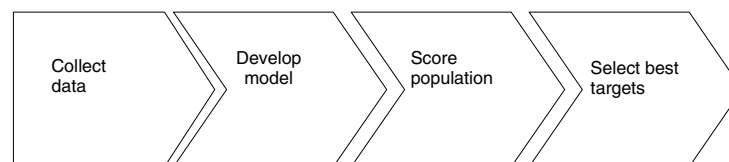
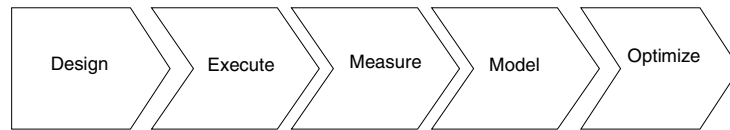


Figure 2. Database marketing campaign process



sampling), and creating a cell design structure (testing various offers and also by age, income, or other variables). I focus on the latter here.

### Problem 1

Classical designs often test one variable at a time. For example, in a cell phone direct mail campaign, you may test a few price levels of the phone. After launching the campaign and uncovering the price level that led to the highest revenue, another campaign is launched to test a monthly fee, a third campaign tests the direct mail message, and so forth. A more efficient way is to structure the cell design such that all these variables are testable in one campaign. Consider an example: A credit card company would like to determine the best combination of treatments for each prospect; the treatment attributes and attribute levels are summarized in Table 1. The number of all possible combinations =  $4^4 \times 2^2 = 1024$  cells, which is not practical to test.

### Solution 1

To reduce the number of cells, a fractional factorial design can be applied (*full factorial* refers to the design that includes all possible combinations); see Montgomery (1991) and Almquist and Wyner (2001). Two types of fractional factorials are a) an orthogonal design where all attributes are made orthogonal (uncorrelated) with each other and b) an optimal design where a certain criterion related to the variance-covariance matrix of parameter estimates is optimized; see Kuhfeld (1997, 2004) for the applications of SAS PROC FACTEX and PROC OPTEX in market research. (Kuhfeld's market research applications are also applicable to database marketing).

For the preceding credit card problem, an orthogonal fractional factorial design using PROC FACTEX in

SAS with estimable main effects, all two-way interaction effects, and quadratic effects on quantitative variables generates a design of 256 cells, which may still be considered large. An optimal design using PROC OPTEX with the same estimable effects generates a design of only 37 cells (see Table 2; refer to Table 1 for attribute level definitions).

Fractional factorial design has been used in credit card acquisition but is not widely used in other industries for marketing. Two reasons are (a) a lack of experimental design knowledge and experience and (b) business process requirement — tight coordination with list selection and creative design professionals is required.

### Opportunity 1

Mayer and Sarkissien (2003) proposed using individual characteristics as attributes in the optimal design, where individuals are chosen “optimally.” Using both individual characteristics and treatment attributes as design attributes is theoretically interesting. In practice, we should compare this optimal selection of individuals with stratified random sampling. Simulation and theoretical and empirical studies are required to evaluate this idea. Additionally, if many individual variables (say, hundreds) are used in the design, then constructing an optimal design may be very computationally intensive due to the large design matrix and thus the design may require a unique optimization technique to solve.

## RESPONSE MODELING

### Problem 2

As stated in the introduction, response modeling uses data from a previous marketing campaign to identify

Table 1. Example of design attributes and attribute levels

Attribute	Attribute level
APR	5.9% (0), 7.9% (1), 10.9% (2), 12.9% (3)
Credit limit	\$2,000 (0), \$5,000 (1), \$7,000 (2), \$10,000 (3)
Color	Platinum (0), Titanium (1), Ruby (2), Diamond (3)
Rebate	None (0), 0.5% (1), 1% (2), 1.5% (3)
Brand	SuperCard (0), AdvantagePlus (1)
Creative	A (0), B (1)

5 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: [www.igi-global.com/chapter/marketing-data-mining/10687](http://www.igi-global.com/chapter/marketing-data-mining/10687)

## Related Content

---

### Learning Bayesian Networks

Marco F. Ramoni and Paola Sebastiani (2005). *Encyclopedia of Data Warehousing and Mining* (pp. 674-677).  
[www.irma-international.org/chapter/learning-bayesian-networks/10682](http://www.irma-international.org/chapter/learning-bayesian-networks/10682)

### The Development of Ordered SQL Packages to Support Data Warehousing

Wilfred Ng and Mark Levene (2002). *Data Warehousing and Web Engineering* (pp. 285-311).  
[www.irma-international.org/chapter/development-ordered-sql-packages-support/7876](http://www.irma-international.org/chapter/development-ordered-sql-packages-support/7876)

### Efficient and Robust Node-Partitioned Data Warehouses

Pedro Furtado (2007). *Data Warehouses and OLAP: Concepts, Architectures and Solutions* (pp. 203-229).  
[www.irma-international.org/chapter/efficient-robust-node-partitioned-data/7622](http://www.irma-international.org/chapter/efficient-robust-node-partitioned-data/7622)

### Advanced Ad Hoc Star Query Processing

Nikos Karayannidis, Aris Tsois and Timos Sellis (2007). *Data Warehouses and OLAP: Concepts, Architectures and Solutions* (pp. 136-256).  
[www.irma-international.org/chapter/advanced-hoc-star-query-processing/7619](http://www.irma-international.org/chapter/advanced-hoc-star-query-processing/7619)

### Enhancing Web Search through Web Structure Mining

Ji-Rong Wen (2005). *Encyclopedia of Data Warehousing and Mining* (pp. 443-447).  
[www.irma-international.org/chapter/enhancing-web-search-through-web/10638](http://www.irma-international.org/chapter/enhancing-web-search-through-web/10638)