CASE STUDY - BANKING
MARKETING CASE STUDY (RESPONSE MODELS)

ANALYTIXLABS

Website: www.analytixlabs.co.in
Email: info@analytixlabsl.co.in

OVERVIEW OF RESPONSE MODELS:

Targeting the Right Prospects: What are Response Models?
Response models use data mining to find similarities between responders from previous marketing campaigns to predict who is likely or not likely to respond to a future campaign. The model is then scored against the prospects of the new campaign and a marketer can choose to mail only those people that are most likely to purchase. This increases conversions and decreases costs by only mailing to those most likely to respond.

Direct Marketing Models: Good, Better, Best
Not all models are created equal. Here's a quick summary of different types of direct marketing models:

1. GOOD. *Recency, Frequency, Monetary (RFM) models:* **simple, better than not modeling**

Though very basic, many marketers still rely on RFM models. Technically RFM models aren't actually response models since they are descriptive but not truly predictive. This method emphasizes customer behavior and segments by how recently a customer purchased, how often they purchase, and how much they spend.

RFM can identify good customers and provide a lift in response rates versus not using any targeting. Another benefit is that it is both simple and descriptive, so it is easily understood by business people.

Unfortunately, RFM doesn't take into account life stage and assumes that customers are likely to continue to respond the same way. If RFM is the only targeting method, the most attractive segments are likely to be over-marketed to at the expense of other segments that could be invested in.

2. BETTER. *Traditional Response or Regression Models:* **more sophisticated and predictive than RFM**

Regression models determine the correlation between variables. Unlike RFM models, regression takes into account that scores can quickly change when combined with other variables.

The model is developed specifically to predict a desired behavior, such as response. Response models require both responder and non-responder data to identify patterns of those likely to respond to a marketing campaign.

This is by far the most widely used approach for marketers and has been a mainstay of predictive analytics for decades.

3. BEST. *Multi-Channel Customer Level Response Models:* **A New Approach that Outperforms Traditional**

This innovative approach identifies not only those prospects most likely to purchase, but also which marketing channel they are most likely to respond to. This allows marketers to optimize their marketing budgets most effectively by contacting the prospect in the channel(s) they prefer and are most likely to be moved by.

Multi-Channel Customer Level Response Models are different from traditional response models in that all of a prospect's known activity is taken into account – email opens, web browsing, display ad click-throughs, mobile, purchase behavior – and not just direct mail behavior. With a more holistic view of the customer, a marketer can create the ideal customer contact strategy for each customer.

BUSINESS PROBLEM:

This example uses data related with direct marketing campaigns of a Portuguese banking institution. The marketing campaigns were based on phone calls. Often, more than one contact to the same client was required, in order to assess if the product (bank term deposit) would be subscribed ('yes') or not ('no').

DATA AVAILABLE:

- bank-additional.csv (for this case study)
- bank-additional-full.csv (this data set may take more time to run some of the machine learning models. You may use it to build the model on complete data)

Data Description:

1. Bank client data:

- ✓ age (numeric)
- ✓ job: type of job (categorical: 'admin.', 'blue-collar', 'entrepreneur', 'housemaid', 'management', 'retired', 'self-employed', 'services', 'student', 'technician', 'unemployed', 'unknown')
- ✓ marital: marital status (categorical: 'divorced', 'married', 'single', 'unknown'; note: 'divorced' means divorced or widowed)
- ✓ education (categorical: 'basic.4y', 'basic.6y', 'basic.9y', 'high.school', 'illiterate', 'professional.course', 'university.degree', 'unknown')
- ✓ default: has credit in default? (categorical: 'no', 'yes', 'unknown')
- ✓ housing: has housing loan? (categorical: 'no', 'yes', 'unknown')
- ✓ loan: has personal loan? (categorical: 'no', 'yes', 'unknown')

2. Related with the last contact of the current campaign:

- ✓ contact: contact communication type (categorical: 'cellular', 'telephone')
- ✓ month: last contact month of year (categorical: 'jan', 'feb', 'mar', ., 'nov', 'dec')

- ✓ day_of_week: last contact day of the week (categorical: 'mon', 'tue', 'wed', 'thu', 'fri')
- ✓ duration: last contact duration, in seconds (numeric). Important note: this attribute highly affects the output target (e.g., if duration=0 then y='no'). Yet, the duration is not known before a call is performed. Also, after the end of the call y is obviously known. Thus, this input should only be included for benchmark purposes and should be discarded if the intention is to have a realistic predictive model.

3. Past Campaign Information:

- ✓ campaign: number of contacts performed during this campaign and for this client (numeric, includes last contact)
- ✓ pdays: number of days that passed by after the client was last contacted from a previous campaign (numeric; 999 means client was not previously contacted)
- ✓ previous: number of contacts performed before this campaign and for this client (numeric)
- ✓ poutcome: outcome of the previous marketing campaign (categorical: 'failure','nonexistent','success')

4. Social and economic context attributes:

- ✓ emp.var.rate: employment variation rate - quarterly indicator (numeric)
- ✓ cons.price.idx: consumer price index - monthly indicator (numeric)
- ✓ cons.conf.idx: consumer confidence index - monthly indicator (numeric)
- ✓ euribor3m: euribor 3 month rate - daily indicator (numeric)
- ✓ nr.employed: number of employees - quarterly indicator (numeric)

5. Output variable (desired target):

y - has the client subscribed a term deposit? (binary: "yes","no")

Missing Attribute Values: There are several missing values in some categorical attributes, all coded with the "unknown" label. These missing values can be treated as a possible class label or using deletion or imputation techniques.

Data Source: https://archive.ics.uci.edu/ml/datasets/Bank+Marketing