PROJECT SMART HOME PRESENTATION

Project presentation

Data Mining

Objective of the document

The present document summarizes the guidelines for the Datamining projects.

Smart Home Company

Welcome to the Smart Home Company. This organization is a well-established company operating in the Smart Home sector. Presently they have around 100,000 registered customers and serve more than 200,000 consumers a year.

They sell products from 5 major categories: Lighting, Door Locks, Cameras, Thermostats and Security Systems. These five categories can further be divided in Premium products and Regular products.

The Customers can order and acquire those products through 3 channel groups: Physical stores, yearly catalogs and the companies' website.

Globally, the company had solid revenues and a healthy bottom line in the past 3 years, but the profit growth perspectives for the next 3 years are fickle. A few strategic initiatives are being considered to invert the situation. One of those is a Marketing efficiency program to improve marketing activities with a special focus on boosting tremendously the efficiency of the marketing campaigns.

Marketing Department

The marketing department is under pressure to spend more wisely its annual budget. The CMO knows the importance of having a more quantitative approach to marketing decisions. A small team of 3 data scientists (your group) was setup with a clear objectives in mind: build a first predictive model to support direct marketing initiatives. Desirably, the success of these activities will prove the value of the approach and convince the more skeptical within the company.

Project “Predictive Model”

Objective

The objective of the team is to build a predictive model that will produce the highest profit for the next direct marketing campaign of the company – the sixth campaign this year that is scheduled for next month. The campaign aims at selling a new gadget to the Customer database (potential of 100,000 Customers).

To be able to build the predictive model, a pilot campaign was carried out. A sample of 2.500 Customers were contacted by mail regarding the acquisition of the gadget. During the next 3 Months, the Customer who bought the offer were tagged with a 1 whereas the non-respondents were tagged with a 0.

The total cost of the sample campaign was 2.500 contacts * 4€, for a total of 10.000€. 312 Customers accepted the offer, each contributing with 14€ of revenue. Overall, the success rate of the campaign was around 12.5% (which is quite good) but with a negative profit of -5.625€.

The idea is to develop a model that predicts customer behavior and apply this model to the rest of the customer base. Hopefully the model will allow the company to cherry pick the customers that are most likely to purchase the offer, while leaving out the non-responders, making the next campaign highly profitable.

Datasets

The dataset used for predictive model building will be “Group_xx_SmartHome_campaign.xlsx”. This dataset includes *a priori* information about the 2.500 customers contacted but also *a posteriori* information about who responded positively (1) or not (0) to the offer (DepVar variable).

Another dataset is also made available for the predictive model project: the “Group_xx_SmartHome_score.xlsx”. These are the Customers to score. Afterwards a decision must be made, for each one, whether to contact them or not. The file contains 5.000 records - acting as the 97.500 remaining customers that were not included in the trial campaign – from which a list of customers to contact must be picked.

Below you can find a list of variables present in the Datatsets

Variable	Description
AcceptedCmp1	Flag indicating customer accepted offer in campaign 1
AcceptedCmp2	Flag indicating customer accepted offer in campaign 2
AcceptedCmp3	Flag indicating customer accepted offer in campaign 3
AcceptedCmp4	Flag indicating customer accepted offer in campaign 4
AcceptedCmp5	Flag indicating customer accepted offer in campaign 5
Complain	Flag indicating if customer has complained (last 18 months)
Dep Var	Target. Binary variable indicating if the Client accepted (1) or rejected (0) the marketing offer offer
Custid	Customer ID
Dt_Customer	Date of customer's enrolment with the company
Education	Level of education of Customer
Income	Yearly Income of household of Customer
Kidhome	Number of kids in household
Marital_Status	Marital Status of Customer
MntLighting	Amount spent on Lighting related products (last 18 months)
MntDoor_Locks	Amount spent on Door_Locks and related products (last 18 months)
MntCameras	Amount spent on Cameras and related products (last 18 months)
MntThermostats	Amount spent on Thermostats and related products (last 18 months)
MntSecurity_Systems	Amount spent on Security_Systems and related products (last 18 months)
MntPremiumProds	Amount spent on premium products (last 18 months)
NumCatalogPurchases	Number of purchases made through catalog
NumDealsPurchases	Number of purchases made with discounts
NumStorePurchases	Number of purchases made through store
NumWebPurchases	Number of purchases made through web
NumWebVisitsMonth	Number of web visits a month to companies site
Recency	Days since last purchase
Teenhome	Number of teenagers in household
Year_Birth	Customer's Year of birth
Z_CostContact	Campaign's Cost per Contact
Z_Revenue	Campaign's positive answer revenue
ElementXX	Group Element. Reject in alaysis
Group	Group ID. Reject in analysis

Evaluation Criteria

Please refer to “Project Evaluation Criteria” document available on Moodle