WARGAMES STORE

Project presentation

Data Mining

Objective of the document

The present document summarizes the guidelines for the Data mining project.

War Games Company

Welcome to the War Games Stores. This organization is a well-established company operating in the Wargames and Miniatures sector. Presently they have around 250,000 registered customers and serve more than 800,000 consumers a year.

They sell products from 5 major categories: Miniatures, card games, painting material, specialized magazines and playing Scenario (forests, houses, mountains,…). These five categories can be also be divided in "Brand A" material and other suppliers.

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

Globally, the company had solid revenues and a healthy bottom line in the past 4 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 250,000 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 proposing the acquisition of the gadget. During the next 2 months, the customers 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€. Around 12.5% of Customers accepted the offer (which is quite good), each contributing with 20€ of revenue. Overall, the campaign had a negative profit of around -3.750€. (This can vary slightly by group dataset)

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 development will be “Group_xx_WarGames_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_WarGames_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 247.500 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	Role	Level
AcceptedCmp1	Flag indicating customer accepted offer in campaign 1	Input	Binary
AcceptedCmp2	Flag indicating customer accepted offer in campaign 2	Input	Binary
AcceptedCmp3	Flag indicating customer accepted offer in campaign 3	Input	Binary
AcceptedCmp4	Flag indicating customer accepted offer in campaign 4	Input	Binary
AcceptedCmp5	Flag indicating customer accepted offer in campaign 5	Input	Binary
Complain	Flag indicating if customer has complained (last 18 months)	Input	Binary
Custid	Customer ID	ID	Interval
DepVar	Binary variable indicating if customer accepted (1) or not (0) a marketing offer from current campaing. Dependent variable of the problem.	Target	Binary
Dt_Customer	Date of customer's enrolment with the company	Input	Interval
Education	Level of education of Customer	Input	Nominal
Income	Yearly Income of household of Customer	Input	Interval
Kidhome	Number of kids in household	Input	Interval
Marital_Status	Marital Status of Customer	Input	Nominal
MntMiniatures	Amount spent on Miniatures (last 18 months)	Input	Interval
MntCard_Games	Amount spent on Card Games (last 18 months)	Input	Interval
MntPainting_Material	Amount spent on Painting Materials (last 18 months)	Input	Interval
MntMagazines	Amount spent on Magazines (last 18 months)	Input	Interval
MntScenario	Amount spent on Scenario (last 18 months)	Input	Interval
MntBrandA _Material	Amount spent on BrandA Material (last 18 months)	Input	Interval
NumCatalogPurchases	Number of purchases made through catalog (last 18 Months)	Input	Interval
NumStorePurchases	Number of purchases made through store (last 18 Months)	Input	Interval
NumDealsPurchases	Number of purchases made with discounts (last 18 Months)	Input	Interval
NumWebPurchases	Number of purchases made through web (last 18 Months)	Input	Interval
NumWebVisitsMonth	Average number of web visits a month to the company site (last 18 Months)	Input	Interval
Recency	# days since last purchase	Input	Interval
Teenhome	Number of teenagers in household	Input	Interval
Year_Birth	Customer's Year of birth	Input	Interval
Z_CostContact	Campaign's Cost per contact	Rejected	Interval
Z_Revenue	Campaign's positive answer expected revenue	Rejected	Interval
Element1	Group Element 1. Unique Identifier of dataset - reject in analysis	Rejected	Nominal
Element2	Group Element 2. Unique Identifier of dataset - reject in analysis	Rejected	Nominal
Element3	Group Element 3. Unique Identifier of dataset - reject in analysis	Rejected	Nominal
Element4	Group Element 4. Unique Identifier of dataset - reject in analysis	Rejected	Nominal
Element5	Group Element 5. Unique Identifier of dataset - reject in analysis	Rejected	Nominal
Group	Group ID. Unique Identifier of dataset - reject in analysis	Rejected	Nominal

Evaluation Criteria

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