Data Mining in CRM & Direct Marketing

Jun Du

The University of Western Ontario

jdu43@uwo.ca

Outline

- Why CRM & Marketing
- Goals in CRM & Marketing
- Models and Methodologies
- Case Study: Response Model
- Case Study: Attrition Model
- Case Study: Solicitation Channel Model

Why CRM & Marketing

- KDnuggets polls: Industries/Fields for Analytics/Data Mining
  - 2011 , 2010 , 2009 , 2008 , 2007 , 2006
  - CRM has been the # 1 industry/field in each of the past 6 years
  - Direct Marketing/Fundraising has been one of the top 10 industries/fields in each of the past 6 years.
- Rexer Analytics surveys: In what fields do you typically apply data mining?
  - 2011 , 2010, 2009, 2008, 2007
  - CRM/Marketing has been the # 1 field in each of the past 5 years.
- News
  - Feb, 2012: How Companies Learn Your Secrets
  - Aug, 2012: Romney Uses Secretive Data-Mining

Outline

- Why CRM & Marketing
- Goals in CRM & Marketing
- Models and Methodologies
- Case Study: Response Model
- Case Study: Attrition Model
- Case Study: Solicitation Channel Model

Goals

- What goal(s) do you expect to achieve by applying data mining to CRM/Marketing?
  - Do you want to attract new customers?
  - Do you want those new customers to be profitable?
  - Do you want to understand the characteristics of your current customers?
  - Do you want to make your unprofitable customers more profitable?
  - Do you want to retain your profitable customers?
  - Do you want to win back your lost customers?
  - Do you want to improve customer satisfaction?
  - Do you want to increase sales?
  - Do you want to reduce expenses?
  - ……

Outline

- Why CRM & Marketing
- Goals in CRM & Marketing
- Models and Methodologies
- Case Study: Response Model
- Case Study: Attrition Model
- Case Study: Solicitation Channel Model

Models and Methodologies

- Response Model
  - Improving response rate of market campaign (attracting new customers)
  - E.g., sending camera advertisement to people who plan to buy it
  - DM technique: classification
- Attrition Model
  - Preventing customer churn (keeping existing customers)
  - E.g., identifying customers who plan to switch cellphone plan to other companies
  - DM technique: classification

Models and Methodologies

- Response Value Model
  - Improving total profit of market campaign (attracting most profitable customers)
  - E.g., sending camera advertisement to people who plan to spend much money on it
  - DM technique: Regression, Ranking
- Solicitation Channel Model
  - Identify the best channel to reach customers
  - E.g., direct mail is the best way to send promotions to some customers
  - DM techniques: Classification

Models and Methodologies

- Segmentation Analysis
  - Segmenting customers by market potential (based on buying behaviour, etc.)
  - E.g., identifying a group of people who might be interested in buying dog food
  - DM technique: Clustering
- Cross-Sell & Up-Sell Model
  - Predicting that customers would buy other (or more) products
  - E.g., predicting people who buy gas lawn mower might also buy gas can
  - DM technique: Association Rules, Classification
- ……

Outline

- Why CRM & Marketing
- Goals in CRM & Marketing
- Models and Methodologies
- Case Study: Response Model
- Case Study: Attrition Model
- Case Study: Solicitation Channel Model

Case Study: Response Model

- Suppose we have an alumni campaign coming in January, 2013.
- We have about 200,000 alumni as candidates, but can solicit only 20,000 of them given the budget.
- How can we apply data mining to achieve the best performance in the campaign?
  - How exactly can we select 20,000 out of 200,000 (i.e., 10%)?
  - What task is it? Classification? Regression? Clustering? …
  - What learning algorithm should we apply?
  - What is the training / test data?
  - How to evaluate the model?
  - ……

Selecting 10%

- Basic idea: rank all the candidates, and select the top 10%.
- Problems:
  - Rank all the candidates, according to ???
  - How can we do ranking? We have never learnt it.
- Solution:
  - We can achieve ranking by doing classification or regression (i.e., use the classification / regression results as the metrics to rank all the candidates)
  - But, how?

Task (1)

- What task is it?
  - Supervised learning: Classification? Regression?

- What is your answer? And why?

Task (2)

- It depends .... on the campaign goal!
- Different campaign goals:
  - To achieve as high participation rate as possible

    (Participation rate = # donors / # solicited-prospects)
    - From DM perspective, to predict giving likelihood
  - To raise as much money as possible
    - From DM perspective, to predict giving potential
- Can we achieve these two goals simultaneously?
  - E.g., two alumni, who should we solicit
    - A: 90% chance to give; give $100 if so
    - B: 10% chance to give; give $100,000 if so

Task (3)

- To achieve high participation rate --- Classification
  - Label: Donor vs. Non-donor (nominal)
  - Giving amount is not considered
  - DM model is expected to produce the probability (likelihood) of being donor for each candidate
  - Candidates with highest probabilities are chosen
- To achieve high raised money --- Regression
  - Label: Giving amount (numeric)
  - Giving amount being 0 for non-donors
  - DM model is expected to estimate the giving amount for each candidate
  - Candidates with highest estimated giving amount are chosen

A More Complex Model

- Recall an example: two alumni, whom should we solicit
  - A: 90% chance to give; give $100 if so
  - B: 10% chance to give; give $100,000 if so
    - Giving likelihood (90% for A, 10% for B) produced by classification; giving potential ($100 for A, $100,000 for B) produced by regression
  - What is the expected giving amount for A, B?
    - A: 100 * 90% = 90
    - B: 100,000 * 10% = 10,000
- Combining classification and regression
  - Produce expected giving amount for each candidate:
    - Expected giving amount = giving potential * giving likelihood
  - Candidates with highest expected giving amounts are chosen

Learning Algorithms

- For regression model, any regression learning algorithm can be applied:
  - Linear regression, KNN, ...
  - Ensemble models usually work better

- For classification model, the model is expected to produce good probability estimation.
  - Can decision tree, KNN, naive Bayes produce good probability estimation?
  - Logistic regression, random forest (and other ensemble approaches) are usually preferred.

Training / Test Data

- Ideally, both training and test data should be i.i.d. (independently drawn from identical distribution)

In our case,

- Test set is the 200,000 candidate alum in a 2013 campaign
- Preferred training data: all solicited prospects in the same (similar) campaign in 2012 (or 2011).
  - Why 2012 (or 2011)?
    - For each of these prospects, we know whether he/she has given and how much he/she has given.
  - Why same (similar) campaign?
    - Training and test data approximately satisfy the “identical distribution” assumption. (Can you explain?)

Evaluation

- Evaluate the model after it has been built
  - Traditional way: 10-fold cross-validation on training data
  - Another way (if sufficient historical data is available):
    - Build model on data from year 2009, and evaluate it on 2010 data;
    - Build model on data from year 2010, and evaluate it on 2011 data;
    - More consistent with how we apply the model in the future

- Evaluate the model after it has been applied
  - After the campaign is over, evaluate the model by using the true results (v.s. the predicted results)
  - This is the actual evaluation

Other Things We Can Do

- Apply ranking techniques to choose candidates
  - Recall: the goal is to select 20,000 prospects with the highest giving potential from total 200,000.
  - Do we really need estimate the specific giving potential for them?
  - Or do we just need rank them according to the giving potential?
  - Ranking techniques can be applied, and might further improve it.
- Instead of using the fixed “20,000”, can we determine the optimal # of alum to be solicited (given the cost of solicitation)?
  - If more alum (in addition to the 20,000) have higher expect giving amount (than solicitation cost), should we solicit them as well?
  - If some alum (within the 20,000) have even lower expected giving amount (than solicitation cost), should we still solicit them?

Demonstration

- Real world data / applications

Outline

- Why CRM & Marketing
- Goals in CRM & Marketing
- Models and Methodologies
- Case Study: Response Model
- Case Study: Attrition Model
- Case Study: Solicitation Channel Model

Case Study: Attrition Model

- Suppose we have 5,000 donors in 2012, how many of them will keep giving in 2013, and how many of them will leave?
- How can we apply data mining to predict whether they will keep giving or leave in 2013?
  - What task is it? Classification? Regression? Clustering? Association Rules? …
  - What learning algorithm we should apply?
  - What is the training / test data?
  - How to evaluate the model?
  - ……

Task

- A classification problem:
  - For each donor in 2012, model needs to make a prediction of whether he/she will stay or leave in 2013.
  - Label: stay as a donor in 2013 vs. leave in 2013

- In addition to the class label, it is preferred for the model to output the probability (likelihood) of leaving (or stay)
  - Priority can be given to the donors who are most likely to leave
  - Classification with probability estimation

Learning Algorithms & Evaluation

- Learning algorithms:
  - Classification algorithms with good probability estimation are preferred
  - Logistic regression, random forest (see previous slides)

- Evaluation:
  - Evaluation after model has been built
  - Evaluation after model has been applied (real evaluation)
  - Same principles as in “Response Model” part

Training / Test Data

- Test set is the 5,000 donors in 2012
  - Snapshot of database in the end of 2012 for these 5,000 donors
- Preferred training data: all donors in 2011 (and whether they made donations in 2012)
  - Snapshot of database in the end of 2011 for all donors in 2011
  - Labels being whether they made donations in 2012
- Question: if one alum made donations in both 2011 and 2012, will we have two identical examples in training and test sets?

Demonstration

- Real world data / applications

Outline

- Why CRM & Marketing
- Goals in CRM & Marketing
- Models and Methodologies
- Case Study: Response Model
- Case Study: Attrition Model
- Case Study: Solicitation Channel Model

Case Study: Solicitation Channel Model

- See demonstration for real-world application