Improve Marketing Campaign ROI using Uplift Modeling

Ryan Zhao
http://www.analyticsresourcing.com

Objective

- To introduce how uplift model improve ROI
- To explore advanced modeling techniques for uplift modeling
- To recommend a few things must be paid attention to
- To introduce applications of uplift models

ROI as a platform for marketing stakeholders

Marketing manager
- Defend or get new budgets
- Design events to improve ROI
- Comparable to other marketing initiatives
- Concentrate on ROI

CMO
- Prove the values of marketing programs
- Improve the efficiency and effectiveness
- Improve the allocation of marketing spend
- Increase credibility, influence and perceived value with senior management

ROI

Modeling manager
- Speak common language as business does
- Prove materialized values of modeling
- Align analytic to business goal
- Offer ROI services to business
- Support optimizing marketing events

Customers
- Receive more quality service
- Feel relevance of communication
- Improve loyalty
- Gain confidence on the brand

Campaign ROI is driven by incremental response

- Three ways to calculate profit per sale ($profit_{sale}$)
  - Profit of product or service (Margin)
  - Net present value (NPV)
  - Life time value (LTV)

- Incremental response equals to difference of response rates between test group and control group

$$r_{test} - r_{control}$$

- ROI of DM campaign is calculated:

$$ROI = (N * profit_{sale} * (r_{test} - r_{control}) - cost_{campaign}) / cost_{campaign}$$

- In order to improve ROI, incremental response is required to be maximized, equivalent to ***Increasing*** $\boldsymbol{r_{test}}$ and meanwhile ***Reducing*** $\boldsymbol{r_{control}}$

Current situations in most marketing organization from a modeler's eyes

- Direct marketing
  - A contact is delivered through DM channel such as DM, DEM or OBTM
  - Customers who are most likely to respond or churn will be selected for the targeting

- Most popular modeling solution
  - Ordinary logistic model was build to score customers' propensities of product acquisition or service activation
  - Survival model was build to score how likely and when a customer is going to churn
  - Most often, we name it either **Propensity model** or **Response model**, or **Churn model**

- A big assumption was made
  - Direct Marketing campaign will achieve maximal **Incremental Response** when a group of the highest scored customers is targeted

Propensity/Response model is NOT necessary to drive neither campaign lift nor ROI

It is not promising that the incremental campaign response will be maximized if customers with the highest propensities of response are targeted.

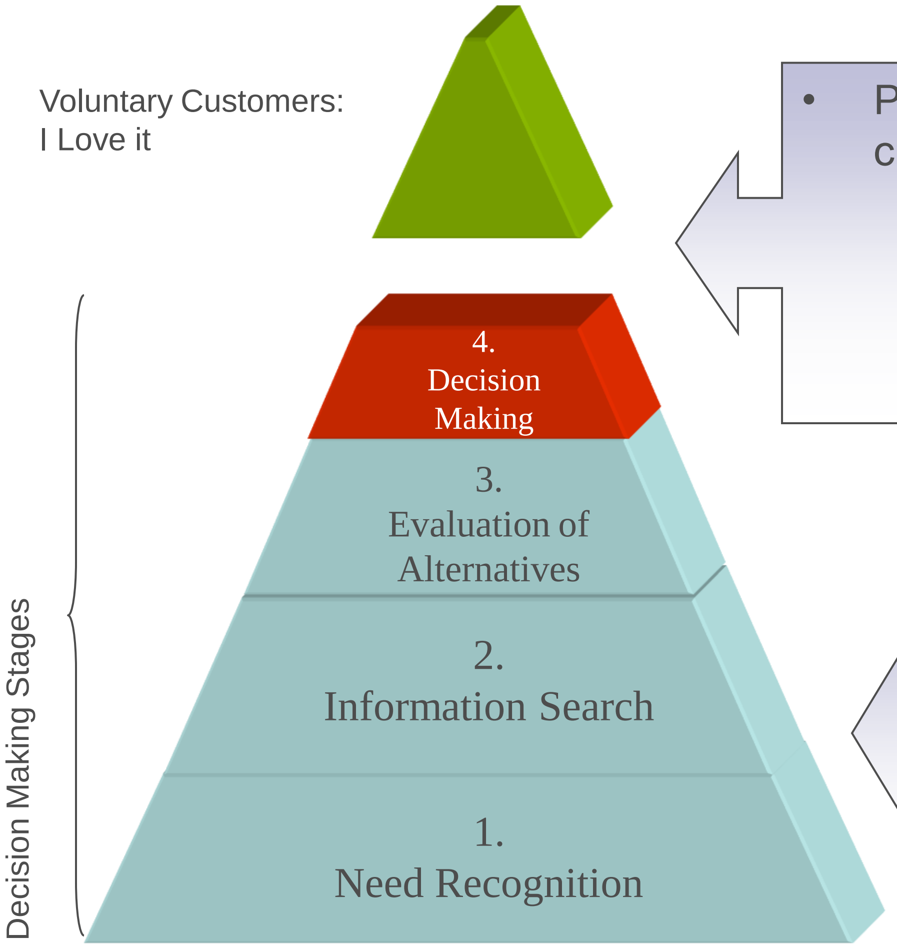
Model Deciles	Campaign resp	Model control resp	Incrmntl resp	ROI	
1	19.4%	18.0%	1.4%	**-0.7189**	
2	18.2%	17.1%	1.1%	**-0.7667**	
3	16.9%	14.5%	**2.4%**	**-0.4576**	
4	15.4%	11.6%	**3.8%**	**-0.0713**	
5	14.0%	8.7%	**5.3%**	**0.4020**	☑
6	10.8%	6.6%	**4.2%**	**0.3687**	☑
7	7.7%	5.5%	2.2%	**-0.0754**	
8	5.6%	4.2%	1.5%	**-0.2679**	
9	3.9%	3.3%	0.7%	**-0.6090**	
10	2.7%	2.3%	0.4%	**-0.6844**	

Random resp	Random control resp	Incrmntl resp	ROI
11.4%	9.2%	2.3%	-0.3130

- Propensity/Response model itself is not going to tell marketers which customers are most likely to contribute to the incremental campaign response

➢ An alternative statistical model is needed, targeting the customers whose propensities of response are dramatically driven by promotion

Customers' Decision Making Stages



Only customers in stage 1, 2, 3 are likely to be influenced

- Propensity model gives higher scores to customers who are most likely to "love it"
  - Voluntary customers are targeted

- Usually, Marketing wants to Influence customers to "love it"
  - Customers in stage 1, 2, 3 are targeted

Churners

High

6.0

Sleeping Dog

- *I don't know it's the time for shopping around*
- *I don't know I paid too much for what I got*
- *I don't know there is a L&R deal*

5.0

4.0

…Now, I knew it!

Lost Cause

- *I hate your products and services*
- *I find a product or service much better than yours*
- *I want something new*

…Huh, I don't care!

Churn If Treated

Sure Thing

3.0

- *I love your products and services*
- *I find your products and services are the best*
- *I don't bother to change*

2.0

…Hey, trust me!

Savable

- *I am thinking but I am not sure if worth to leave*
- *I am looking for a deal to renew my service*
- *I am tempting on an offer from other brand*

…Aha, You read my mind!

Low 1.0 2.0 3.0 4.0 5.0 6.0

Low Churn If Not Treated High

Ultimate modeling goal from ROI perspective

- Ultimate modeling goal is to find customers who are most likely influenced by marketing campaign
  - Drive incremental campaign response (r_{test} - $r_{control}$)
  - Increase ROI
  - Increase Overall Market Response Rate

- Thus, model should help avoid marketing money to be spent on :
  - Purchaser
    - Customers who will naturally respond
    - Customers who are die-hard non-responders
  - Churner
    - Sure thing
    - Lost cause
    - Sleeping dog

What is uplift model (WIKIPEDIA)

- Uplift modeling, also known as incremental modeling, true lift modeling, or net modeling is a predictive modeling technique that directly models the incremental impact of a treatment (such as a direct marketing action) on an individual's behavior

- Uplift modeling has applications in customer relationship management for up-sell, cross-sell and retention modeling. It has also been applied to personalized medicine.

Use uplift model to drive ROI

Unit price	$14.99
Product margin	30%

Email cost	$0.02
Offer cost	$1.00

	Top Decile (Response Model)			Top Decile (Uplift Model)		
	Test	Control	Incremental	Test	Control	Incremental
Total customers	100,000	5,000	n/a	100,000	5,000	n/a
Take up rate	19.4%	18.0%	**1.4%**	8%	3%	**5%**
Total sales	$290,806	$269,820	$20,986	$119,920	$44,970	$74,950
Total profit	$87,242	$80,946	**$6,296**	$35,976	$13,491	**$22,485**
Total email cost	$2,000			$2,000		
Marketing overhead	$1,000			$1,000		
Offer cost	$19,400			$8,000		
Total cost			**$22,400**			**$11,000**
Dollar return			**($16,104)**			**$11,485**
ROI			**-0.719**			**1.044**

Uplift models – Differential response (two model)

- Build two logistic models

$$Logit(P_{test}(response|X,treatment=1)) = \alpha + \beta^*X + \gamma^*treatment$$
$$Logit(P_{control}(response|X,treatment=0)) = \alpha + \beta^*X$$

- Calculate the uplift score by taking difference of two scores

$$Score = P_{test}(response|X,treatment=1) - P_{control}(response|X,treatment=0)$$

Pros

- *Uses standard logistic regression modeling techniques*
- *Easy to implement and maintain*

Cons

- *Does not fit he target directly (i.e. incremental response)*
- *Introduces modeling errors twice*
- *Sensitive to predictive variable selections and parameter estimations*

Uplift models – Differential response (One model)

- Build two logistic models

$$Logit(P(reponse|X) = \alpha + \beta*X + \gamma*treatment + \lambda* treatment *X$$

- Calculate the uplift score by taking difference of two scores

$$Score = P(response|X,treatment =1) - P(response|X,treatment =0)$$

Pros

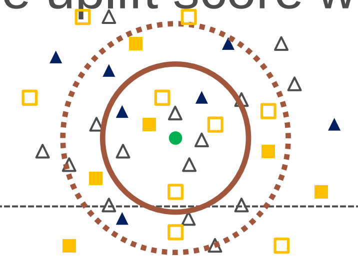
- *Uses standard logistic regression modeling techniques*
- *Better robustness comparing to two model approach*
- *Effect modifications due to treatment*

Cons

- *Does not fit the target directly (i.e. Lift)*
- *Increases modeling complexity due to assumptions of Non-linearity*
- *Needs trade-off between significances and sizes of parameter estimations due to turning treatment on/off*

Uplift models – KNN approach

- Step one: Conduct neighborhood components analysis to find the optimal distance measurement
- Step two: Find k nearest neighbors for a candidate customer
- Step three: Calculate the uplift score within the neighbors



Pros

- *Fits the business objective directly*
- *No assumption on linearity of predictive variables*
- *Straightforward concept and can easily to be implemented in SAS*

Cons

- *Increased computing complexity*
- *Difficult to be product ionized*
- *Need to do extra work to understand what the model tells*

Uplift models – Naïve Bayes approach

- Step one: Variable selection using Semi-Naïve Bayes Classifier

$$\log\left(\frac{P(response = 1 \mid X_1, \ldots, X_p)}{P(response = 0 \mid X_1, \ldots, X_p)}\right) = \alpha + \sum_{j=1}^{p} b_i g_j(x_j)$$

- Step two: Estimate Naïve effect for both Tests and Controls

$$g(x_j) = \log \frac{f(x_j \mid response = 1)}{f(x_j \mid response = 0)}$$

- Step three: Calculate the uplift score by subtracting NBCs

$$\sum g_{test}(x_j) - \sum g_{control}(x_j)$$

Pros

- *Nonparametric*
- *Posterior probability to update briefs of differentiators*
- *Alleviate problems stemming from the curse of dimensionality*

Cons

- *Linearly separable is assumed that is may not achievable in real world*
- *Independences of predictive variables are assumed*

Variable selection is the key

- Select predictive variables to maximize the differential response
- The standard process to select the variables in predictive analysis may not be necessary to select the best differentiator variables
- Nonlinearity is assumed for continuous variables unless linearity is identified
- The regular way to build a logistic model may not work very well
  - Goodness of fit, Concordance, C-statistics, Pearson chi-square, Ward chi-square etc. only make the model robust
  - But not enough to maximally differentiate the responses

Pay attention when building a uplift model

- Differential response model is very sensitive

- Any invalid assumption ***may*** make the model ***fail*** and seek ***worse*** targets for marketing campaigns
  - Training sample represents universe

  - In training sample, contacted customers are basically the same as non-contacted

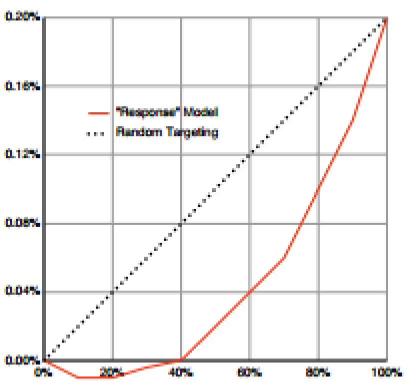
  - Customer’s response behaviors towards a product/service won’t change even though campaign message or channel is changed

Magic story – cross selling

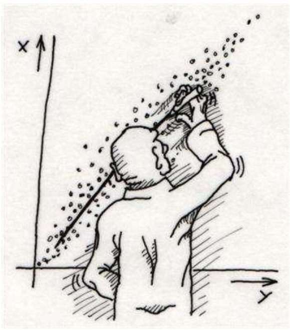
Challenge: A major North American bank repeatedly ran a cross-selling campaign for a high-value product to selected segments of its customer base. The first run of the campaign was a random trial, achieving 0.2 percentage points incremental from 0.9% of control group. Since the mailing cost was around $1.50, and the NPV of a sale was over $1,000, this was a highly profitable campaign.

The second run of the campaign targeted the best 30% of those identified by a standard “response” model. To assess the quality of the model, around 10% of the lower seven deciles were also mailed as a “targeting control”. The result shows that targeting the “best” 30% (as chosen by the “response” model) yields no incremental sales at all. Although insignificantly, in fact, the rate of sales is slightly lower than in the control group.

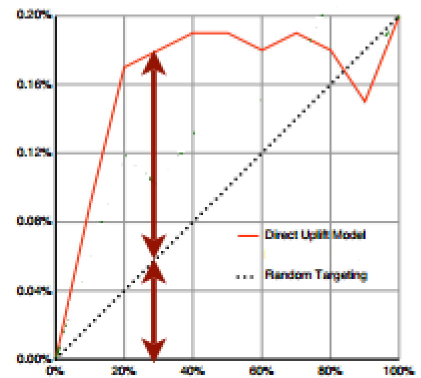
Used ordinary “response” model



Proposed Uplift Modelling approach



Initiative delivered significant incremental sales



Achievement: The uplift approach not only led to significant incremental sales at the chosen cut-off (30%), but significantly (unlike the “response” model) out-performed an untargeted approach, generating incremental sales about three times as high as with random targeting and almost two times more NPV from the campaign.

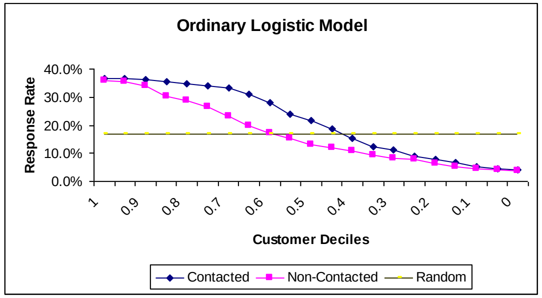
Source: http://stochasticsolutions.com/

Magic story – multi-channel campaign optimization

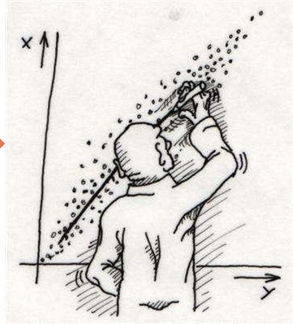
Challenge: A Canadian drug chain decided to run a integrated multi-channel campaign to promote its own private label products. The initiative included massive promotion through advertisements on newspaper, TV commercial and in-store signage's, and direct mailing of a catalogue with coupons for selected lead products as well. Previous experiences suggested that the direct mailer generated insignificant lift. Considering high cost of mailers and coupon offers, ROI of the direct mailer was very disappointed.

Due to the direct mailing being mixed with other promotion activities such as TV commercial and newspaper advertisement and in-store signage's, ordinary 'response' model is most likely to capture the customers who had great exposure to the massive promotion and hence already made mind.

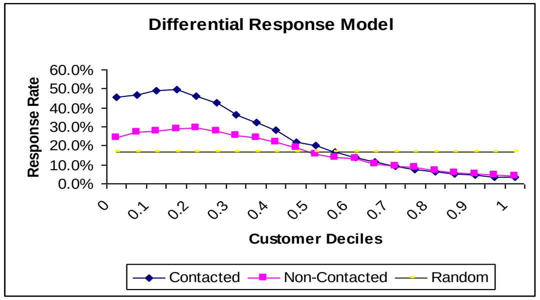
Used ordinary "response" model -- lift of over and above control in top decile is ***1.022***



Proposed Uplift Modelling approach



Used uplift model -- lift of over and above control in top decile is ***1.902***



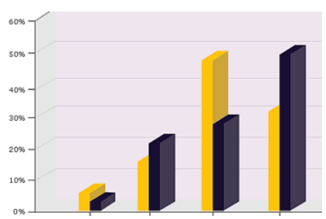
Achievement: The uplift approach led to a significant lift by four times more than expected response rate. New to PL customers became the major driver in the direct mailing campaign. As in an email to upper management, the uplift model impact was listed as the second important factor why the campaign was over budget significantly.

Magic story – savable churners

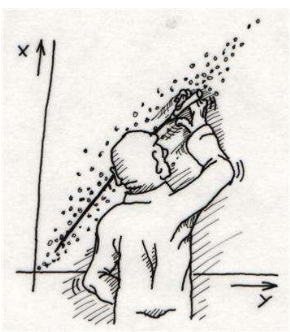
Challenge: A major North American Telecom was experiencing an average monthly churn of 2% among its high speed Internet client base. Loyalty and Retention team deployed retention strategies in synchronizing programs and processes to keep customers longer by providing them with tailored products and services. In order to support retention campaigns, modeling team developed a survival model not only to predict which customers are at high risk of churn, but also how soon these high-risk customers will churn.

When the model was applied to the churn prevention practice, a few questions were raised: **1).** Contacting all client base achieved negative impact, how to select best targets; **2).** When is the best time to contact clients; **3).**The survival model doesn't predict if a client is savable or not

Used ordinary "Survival" model



Proposed Uplift Modelling approach



Delivered significant incremental retentions

Achievement: Reduced churn by 36%, while reduced the target volume of the campaign by 40% and increased the LTV 1100%.

Analytics Resourcing

CENTRE

☑ Graduate Recruitment ☑ Analytics Coaching ☑ Professional Service