Incremental Response Modeling: Uses, Algorithms and Comparisons

Steve Gallant

Director, Data Mining Services

Steve.Gallant at kxen.com

kxen
KNOWLEDGE EXTRACTION ENGINES

May 12, 2010

THE DATA MINING AUTOMATION COMPANY™

Agenda

- Why Incremental Response Modeling?
- IRM Algorithms – and Estimating Results
- Experiments
- Results
- Automation
- Conclusions

- Q&A

We want to model and optimize the right measure!

Example: Coupon for 20% off

- Retailer sends out 20% off coupons. Who should receive coupons?
- Classic Response modeling: Send coupons to those who responded to the last promotion.
- “Result” is that model performs well, and lots of coupon recipients make purchases.
- But this includes people who shop every week!
- Actual result: For these people, we are reducing margin and hurting the bottom line.

What is Incremental Response Modeling?

- **In the 20% coupon example, we only want to send a coupon to <u>those who shop more if they receive a coupon</u>.**
- **Typically we want to maximize the probability of shopping, times the expected amount for that individual if they shop.**
- **Incremental Response Modeling (also called “Uplift Modeling”) seeks to maximize the increased profitability from a decision.**
- **There can be a distinction between shopping more often and the spend amount.**

Example 2: Offers to Existing Customers

Example: Credit Line Increase

- Bank wants to know which customers should get a credit line increase.

- **In the credit line increase example, we want to model who would increase profitability, not who would take advantage of the increase.**

Variation: Retention Offer by a Wireless Provider

- Who should get a special offer to keep them from churning?

Agenda

- Why Incremental Response Modeling?
- IRM Algorithms – and Estimating Results
- Experiments
- Results
- Automation
- Conclusions

- Q&A

“Standard Approach”

- **Terminology:**
  - **Treated Customers receive the promotion**
  - **Untreated Customers don’t receive the promotion**
- **Basic idea is to subtract probabilities of model for untreated customers from model for treated customers**
- **“Standard Approach” requires twice as many models: probability of purchase with Treatment, and probability of purchase without Treatment**

Standard Approach for Incremental Lift

Customer	Probabilities for Mail Model P(response)$_{mailed}$	Probabilities for Control model P(response)$_{not\ mailed}$	Probability Difference (Incremental Response)
1	0.7536	0.7502	0.0034
2	0.5824	0.4795	0.1029
3	0.0579	0.0285	0.0294
4	0.3754	0.3841	-0.0087
5	0.0102	0.0101	0.0001
6	0.6421	0.6954	-0.0533
.	.	.	
.	.	.	
.	.	.	
.	.	.	
.	.	.	
.	.	.	
.	.	.	
.	.	.	
1000	0.1873	0.1654	0.0219

Customer	Incremental response probability
2	0.1029
.	.
.	.
3	0.0294
.	.
1000	0.0219
.	.
.	.
1	0.0034
.	.
5	0.0001
.	.
4	-0.0087
6	-0.0533

Sort by descending score difference

Select from top to bottom for targeting

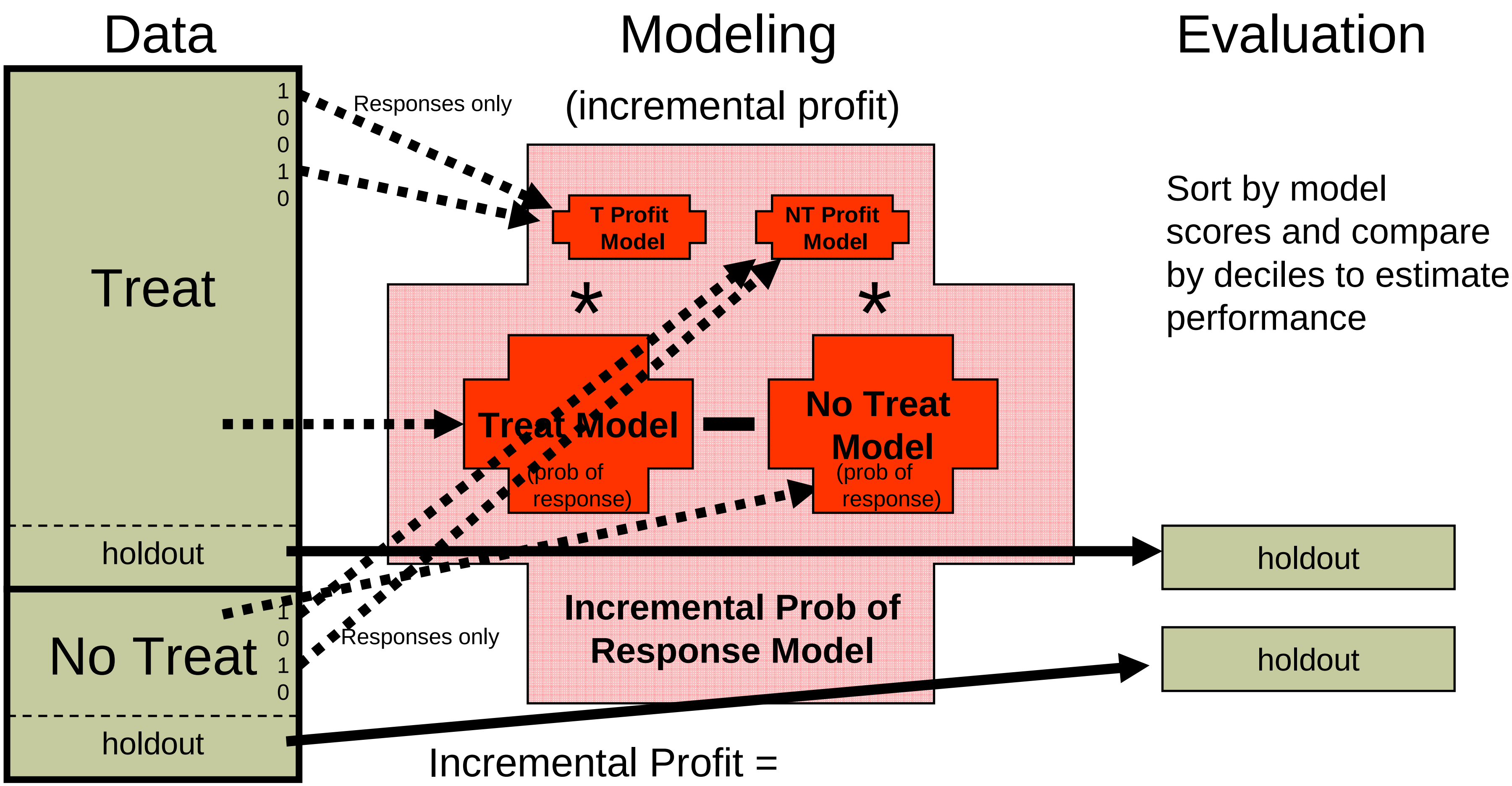
Source: Laura

Evaluating a Model

1. **Apply model to Treat holdout and take top 10%.**
   - Gives response probability of those Treated
2. **Apply same model to Untreated holdout and take top 10%.**
   - Gives response probability of **similar people** who weren't Treated
3. **Subtract response probabilities**
   - Gives estimated increase in probability for top 10% due to the Treatment

Can also do this with *profitability*, rather than *probability*.

Incremental Response Modeling and Evaluation of Models



Random Splits!

Incremental Profit =
Treat Resp Prob * Treat Profit –
No Treat Resp Prob * No Treat Profit

Advantages and Disadvantages

- **Advantage is that this method is “natural” from the definitions**
- **Disadvantage is that we are building many models, so errors are propagating**

Method 2: Large Bank Effect Modeling and Portrait/Quadstone (??)

Age **Marital Status** **…**

[0-17] [18-23] [24-28] […]

Treat  0 0 1 0 1 0 0 1 1 0 0 0 1 0 0 1 0 1 0 0

No Treat  0 0 0 1 0 1 0 0 1 1 0 0 0 1 0 0 1 0 1 0

Target Values

Step 1: Get differences in target percentages for Treat and No Treat

For Age [0-17], Treat is 60% and No Treat is 20%

Make sure there are enough entries in both groups.

Repeat for other Age ranges.

Repeat for other Variables.

Step 2: Take the best candidate as the initial tree splitting node.

Step 3: Repeat, to build a decision tree.

"Decision Tree Riding 2 Horses"

2-Horse Decision Tree Advantages and Disadvantages

- **Advantage: Directly finds simple good combinations and directly models continuous variables**
- **Disadvantage: Top split looks only at one node, as does each of the other splits. This is a standard issue with decision trees.**
- **With noisy data, decision trees usually are not as good as regression approaches.**
- **Must be careful that a node contains enough cases for both "Treat" and "No Treat". Otherwise comparison statistic is not reliable.**

Victor Lo's "True Lift" Approach

- **Basic idea:**
  1. Reduce the number of variables for modeling "Treat" and "No Treat" groups.
  2. Model "Treat" and "No Treat" data together. Make an additional copy of "Treat" variables, setting these variables to 0 for "No Treat" customers. ("Treat" customers use all variables, including duplicates of values.)
  3. To score, apply the model twice. Once with all variables, and again with "Treat" variables set to 0. Then subtract the scores

Agenda

- Why Incremental Response Modeling?
- IRM Algorithms – and Estimating Results
- Experiments
- Results
- Automation
- Conclusions

- Q&A

Data Sets

- **Large Bank: 2 sets for different promotional offers to existing customers**
- **Specialty Retailer: 3 sets for different customer/prospect groupings (eg., active customers, inactive customers, acquisition possibilities)**
- **Financial data from Fidelity: artificial data**
- **Other sets are being received.**

Predictive Models

Models	Binary (shop or not)	Profit if they shop (only shoppers)	Model on Profit directly (all customers, including non-shoppers)
"Response" – Model on Treat data	X	X	X
"Look Alike" – Model on Untreated Data	X	X	X
"True Lift" Models	X	TBD	TBD
Tree	TBD	N/A	X

Scoring

- **Compute incremental profit when promote top 10% of list, and top 20%**
- **Examples (comparisons computed by SQL queries on model outputs):**
  - Order treat and untreated hold out sets by shop probability model built on treated data.
  - Order treat and untreated hold out sets by difference in shop probabilities for models built on treated data and built on untreated data (Standard IRM, looking only at shop probabilities and ignoring profit modeling).
- **For all except Tree models, use KXEN “out of the box”, without any model tuning, variable reduction, etc.**

What I expected:

One approach would dominate, probably the “Standard” approach.

Agenda

- Why Incremental Response Modeling?
- IRM Algorithms – and Estimating Results
- Experiments
- Results
- Automation
- Conclusions

- Q&A

Summary — Results top 10%

	ads_04_07	ads_05	ads_03	Large Bank Pricing	Large Bank CLI	Fidelity (artificial)
				18.40	16.05	0.0128
Baseline "Look Alike" using Untreated model ('CTL_ID' = 1) order by Prob * Price	23.02	19.25	38.64	5.00	5.69	0.0128
Baseline "Response" using Treated model ('CTL_ID' = "")	9.83	20.46	45.78	13.62	40.33	0.0130
Baseline "Look Alike" using Untreated model ('CTL_ID' = 1) order by Prob only	-1.96	21.37	49.22			
Baseline "Response" using Treated model ('CTL_ID' = "") order by Prob only	-12.66	17.93	50.62			
Standard Incremental Response order by Prob only	6.91	7.70	5.01			
"Standard" Incremental Response order by Prob * Price	23.74	1.31	29.51	17.27	99.21	0.0117
Direct Model on Continuous Target	12.21	31.75	44.23			
True Lift Models				39.67	75.44	0.0125
True Lift Models (NULLs, not 0.0)				31.73		
Tree (Effect Modeling)				13.84	93.20	0.0131

Within 5% of highest

Within 5% of lowest

Summary	Results top 20% ads_04_07	ads_05	ads_03	Large Bank Pricing	Large Bank CLI	Fidelity (artificial)
				13.03	23.84	0.0083
Baseline "Look Alike" using Untreated model ('CTL_ID' = 1) order by Prob * Price	5.03	8.47	30.10	9.34	3.88	0.0084
Baseline "Response" using Treated model ('CTL_ID' = "")	5.29	4.68	27.12	7.84	30.65	0.0083
Baseline "Look Alike" using Untreated model ('CTL_ID' = 1) order by Prob only	-5.72	7.83	27.88			
Baseline "Response" using Treated model ('CTL_ID' = "") order by Prob only	-3.04	7.71	28.90			
Standard Incremental Response order by Prob only	2.96	2.78	10.02			
"Standard" Incremental Response order by Prob * Price	4.52	5.88	20.30	10.84	63.22	0.0079
Direct Model on Continuous Target	11.14	7.42	28.63			
True Lift Models				13.32	50.49	0.0084
True Lift Models (NULLs, not 0.0)				15.44		
Tree (Effect Modeling)				13.84	60.07	0.0080

What I Expected

- ~~One approach would dominate, probably the "Standard" approach.~~
- We can find conditions that indicate which models will do best.

Complexity

- **If good models for <u>probability</u> of shop and <u>profit</u> if they shop for shoppers, then expect "Standard" approach to do well.**
- **If difficult to model <u>profit</u>, then just using shop <u>probability</u> should work better.**
- **If <u>profit</u> dominates everything (ie., looking for big fish), it is possible the profitability model alone will work better.**
- **Tree models are also faced with whether to model probability of shop or overall profit.**

What I Expected

- ~~One approach would dominate, probably the "Standard" approach.~~
- ~~We can find conditions that indicate which models will do best.~~

What I Expected

- ~~One approach would dominate, probably the "Standard" approach.~~
- ~~We can find conditions that indicate which models will do best.~~
- Incremental Response Modeling isn't just Algorithms – it's also an Automation story!

For best results, we need to try a set of models.

Agenda

- Why Incremental Response Modeling?
- IRM Algorithms – and Estimating Results
- Experiments
- Results
- Automation
- Conclusions

- Q&A

Steps to Generate Models and Get Performance Figures

1. **Run a batch script that:**
   - builds 6 KXEN models,
   - scores a Treat and notTreat holdout for each model as a column in a results table, and
   - writes a report file with key information.
2. **Cut and paste the report file into an Excel spreadsheet.**
3. **Run a SQL query to get 10% and 20% performance for various models, and paste result into the Excel spreadsheet.**

One batch job, one SQL query, and 2 cut-and-pastes.

Agenda

- Why Incremental Response Modeling?
- IRM Algorithms – and Estimating Results
- Experiments
- Results
- Automation
- Conclusions

- Q&A

Conclusions

- **No single IRM method dominates all others.**
- **Best to try various approaches for the data and problem at hand, and select the best approach from performance on held-out data.**

Incremental Response Modeling isn't just Algorithms – it's also an Automation story!