Riku Mäkeläinen, TeliaSonera Sweden
Two-Stage Modelling – Cross-Sell Broadband

SAS Xperience, Stockholm, 19.2.2009

TeliaSonera

Contents

- Basic idea behind two-stage modeling
- On selecting customers to target group
- XSell BroadBand case & EM 5.3 demo
- Combined results
- Summary

Limited data and methods used!

Presentation serves as an example only!

:)

Basic Idea Behind Two-Stage Modeling

- Target customers that are likely to give high revenue
- Including more customers with higher expected revenue should compensate lower response probability

  => Total ROI is probably better but response-rate may suffer

- Terms to look in literature
  - Two-stage models
  - Limited dependence models
  - Double hurdle models
  - Zero-inflated models
  - Bivariate Tobit models
  - Dual response models
  - Component models

Cross-Selling Broadband Subscriptions Setup

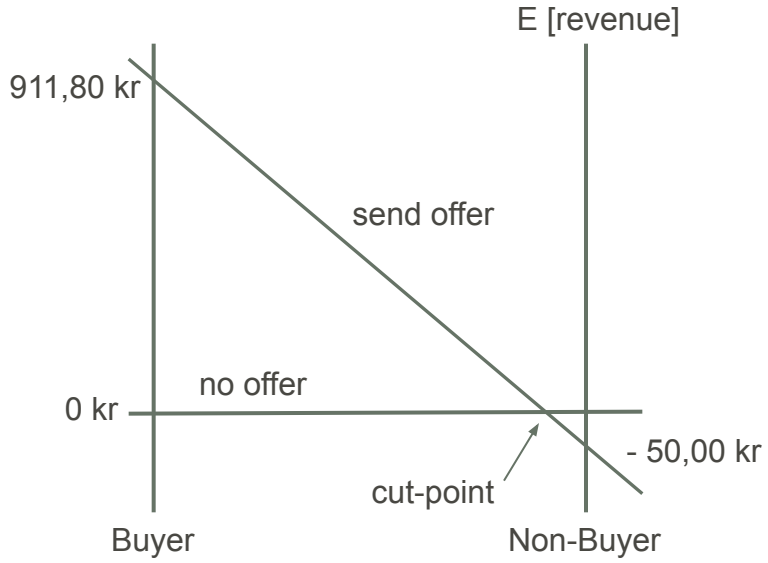
- Goal: offer broadband subscription to current customers who can get broadband but who do not have it today
- Modeling bases on responses to June 2008 BB cross-sell campaign with slight modifications
  - 319 960 customers
  - 42 variables (transformations not included)
- Two target variables available
  - Binary target: customer accepts offer, 12 155 (3,80%) buyers
  - Interval target: if customer accepts offer, what will be the chosen speed (= price)
    - 199 kr – 2 363 customers (19,44%)
    - 229 kr – 5 057 customers (41,60%)
    - 279 kr – 3 587 customers (29,51%)
    - 329 kr – 1148 customers (9,44%)
    - Mean 247,39 kr & median 229 kr
    - Missing value for non-buyers

Profit Matrix

- Simple profit calculus
  - Campaign cost = 50 kr per customer
  - Assume 229 kr/month median revenue with 12 months binding and 35% margin and we get 229 kr/month * 12 months * 0,35 – 50 kr = 911,80 kr income per customers
- Profit matrix (note unbalance):

Outcome	Decision 1 (send offer)	Decision 2 (do not send offer)
1 (Buyer)	911,80	0
0 (Non-Buyer)	- 50,00	0

=> cut-point = 0,05198



On Selecting Customers to Target Group

Some possible selection criteria

- P_BINARY_TARGET > 0.5
- P_BINARY_TARGET > cut-point
- P_BINARY_TARGET * MEAN_REVENUE > 0
- P_BINARY_TARGET * MEDIAN_REVENUE > 0
- P_BINARY_TARGET * MEDIAN_REVENUE – CAMPAIGN_COST > 0
- P_BINARY_TARGET * MEDIAN_REVENUE * MARGIN – CAMPAIGN_COST > 0
- P_BINARY_TARGET * P_REVENUE * MARGIN – CAMPAIGN_COST > 0

- Expected profit if we decide to send offer:
  - E(PROFIT | SEND_OFFER) = P_BINARY_TARGET=1 * REVENUE + P_BINARY_TARGET=0 * (-COST)
- Expected profit if we decide not to send offer:
  - E(PROFIT | NO_OFFER) = P_BINARY_TARGET=1 * 0 + P_BINARY_TARGET=0 * 0 = 0
- Profit matrix can always be transposed so that no-offer column contains zeroes

Selection Criteria Used This Time

- Expected profit if we decide to send offer:
  - E(PROFIT | SEND_OFFER) = P_BINARY_TARGET=1 * REVENUE + P_BINARY_TARGET=0 * (-COST)
- Single stage models:
  - if P_target_binary1 * (229 * 12 * 0.35 - 50) - (1 - P_target_binary1) * 50 > 0;
- Two-Stage models:
  - if P_target_binary1 * (P_target_interval_price * 12 * 0.35 - 50) - (1 - P_target_binary1) * 50 > 0;

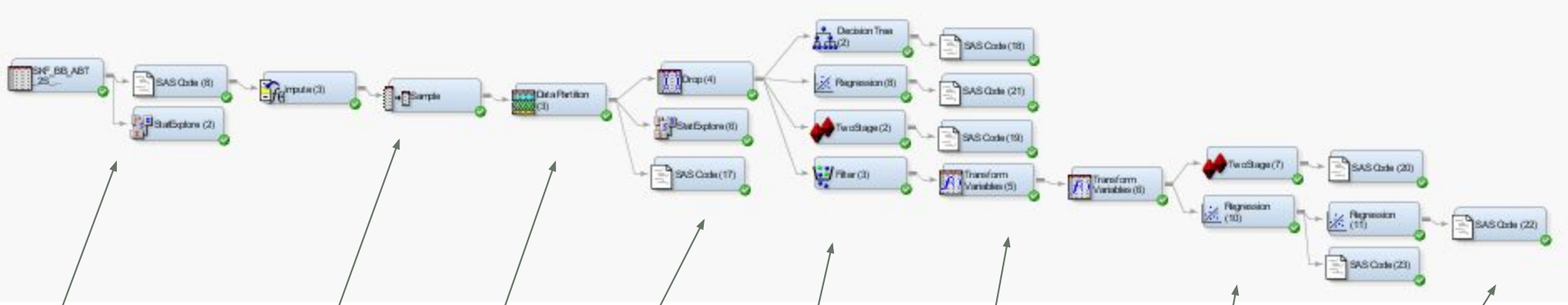
Some Evaluation Criteria For Model Selection

- Model Complexity => pick one you understand or one that performs best
- Statistical measures => misclassification rate, average squared error, …
- Selection-% (= target group size)
- Response-% (= traditional performance measure)
- Expected Average Return On Investment
- Expected Total Return On Investment
- Actual Average Return On Investment in validation data
- Actual Total Return On Investment in validation data

- Something else, for example percentage of ROI captured out of total maximum

- Often no single good measure, in practice combination of little bit everything chosen

Enterprise Miner 5.3 Example Flow



Create variable for stratification and fix data

10% sample

Split to train & validation

Find out optimal performance

Model:
- Tree
- Regression
- Two-Stage

Enhance data:
- remove outliers
- normalize
- standardize

Model:
- Two-Stage
- Regression + Regression

Evaluate results
- Selection-%
- Response-%
- AVG ROI
- SUM ROI

Potential improvements

- Biased sampling (50% / 50% - positive / negative ?)
- Improved modeling methods (bagging / boosting ?)
- Variable selection (variable clustering ?)
- Variable transformations (principal components, binning ?)
- Model log of interval target
- More data …

Combined Results

Model	Data	Target Group Size	Selection-%	Expected ROI	Avg Expected Roi	Actual ROI	Avg Actual ROI	Response-%
Tree	Train	2 676	16,73	79 719,60	29,79	217 917,60	81,43	8,3
Tree	Validation	**2 798**	**17,49**	**83 380,53**	29,80	232 303,20	**83,02**	**8,4**
Regression	Train	3 279	20,50	79 900,93	24,37	251 276,20	76,63	7,9
Regression	Validation	3 281	20,51	85 385,06	**26,02**	270 191,00	82,35	**8,4**
Two-Stage	Train	3 461	21,63	120 530,30	34,83	275 798,60	79,69	8,3
Two-Stage	Validation	3 367	21,05	122 537,00	36,39	214 641,60	**63,75**	**6,6**
Regression ®	Train	3 210	20,07	90 063,52	28,06	268 973,00	83,79	8,6
Regression ®	Validation	3 532	22,08	111 668,40	31,62	286 346,00	81,07	8,2
Two-Stage ®	Train	2 080	13,00	79 758,67	38,35	181 127,20	87,08	8,8
Two-Stage ®	Validation	2 248	14,05	99 708,89	**44,35**	**170 023,00**	75,63	7,8
Reg + Reg ®	Train	3 730	23,32	113 970,40	30,56	299 267,00	80,23	8,2
Reg + Reg ®	Validation	**4 055**	**25,35**	**135 788,30**	33,49	**315 548,00**	77,82	7,9

- Highest actual ROI, highest expected ROI and largest selection-%: Reg + Reg on refined data
- Highest response-%, highest average actual ROI and lowest selection-%: Decision Tree

- Reg + Reg on refined data captures 52,5% of maximum potential ROI
- Decision Tree captures only 38,7% of maximum potential ROI

=> Reg+Reg targets 45% more customers, loses 5,6% response-% (0,5 %-units) and gives 35,8% higher ROI (and Reg+Reg actually gives 26,9% more respondents)

Predicting History…

- Interval target with few distinct values is perhaps not that suitable for two-stage modeling
- Sales goal might not be compatible with higher ROI and lower response-%
  - However bigger target group with little lower response-% gives more respondents…
- Availability: some customers can't get top broadband speed => buyer profiles overlap
- Certain offers may be bundled with certain speed
- Some customers may want to buy multiple broadband subscriptions

=> it could be more useful to predict best offer and not use two-stage modeling here

Summary

- Choosing final model not easy
- Response rate probably goes down but total ROI is likely to be higher => marketing manager may be unhappy…

- Two-Stage Modeling is a good idea
- Takes bit more time
- Overall result is often better

More Material

- Increasing Profitability of MMS Activation Campaigns – Traditional Modelling Methods vs. Two-Stage Modelling, presentation at SAS Forum International 2004 - Copenhagen 15.-17.6.2004 by Riku Mäkeläinen & Sakari Forslund TeliaSonera Finland / Consumer Marketing
- SAS training course ‘Advanced Predictive Modeling Using SAS Enterprise Miner 5.1’
- Two-Stage node documentation


THANK YOU!
riku.makelainen@teliasonera.com
TeliaSonera