Increasing Profitability of MMS Activation Campaigns – Traditional Modelling Methods vs. Two-Stage Modelling
SAS Forum International 2004 - Copenhagen 15.-17.6.2004

Riku Mäkeläinen & Sakari Forslund
TeliaSonera Finland / Consumer Marketing
Contents

1. Introduction
2. Business Case
3. Two-Stage Modelling
4. Campaign Execution
5. Summary
The Nordic Baltic Telecommunications Leader

**TeliaSonera 2003**
- 22 million customers
- Sales SEK 81,772 million

**TeliaSonera Finland 2003**
- 2.4 million mobile customers, 0.8 million fixed, 0.3 million internet
- Sales SEK 17,697 million
- 815 million SMS messages sent 2003
- 645000 MMS messages sent 2004Q1
Business Case

• Business problem: Activate MMS usage profitably
• Solution: Targeted direct mail campaign

- Collecting and storing information
  - Oracle
  - SAS Enterprise Miner

- Analysing and reporting results
  - Business Objects
  - Base SAS

- Campaign execution
  - SAS Marketing Automation

- Updating data
1. Introduction
2. Business Case
3. Two-Stage Modelling
4. Campaign Execution
5. Summary
Why Two-Stage Modelling

- to build better models
- another challenge
- better business

Traditional activation models may overlook customers who are
  - hard to activate - activation probability small
  - but if activated yield high revenues - predicted revenue high

=> it makes sense to try to activate several hard to activate but potentially high value customers

**GOAL**: Show how two-stage modelling increases revenues
Overview of Analysis Chain

1) Preliminary Analysis
2) Data Preparation
3) Traditional Modelling
4) Two Stage Model
5) Stage 1 of Manual Two-Stage Modelling
6) Stage 2 of Manual Two-Stage Modelling
Preliminary Analysis

Analysis chain - 1/8

- **Input Data Source**: Input analysis data
  - target segment: 329 000 subscriptions with 1.77% target events
  - target: subscription that is actively using MMS – one-time users and random users excluded
  - profit matrix:
    - profit: successful contact yields 3.98 eur
    - cost: contact cost 0.70 eur/contact
- **Insight**: make preliminary data analyses
Data Preparation

Analysis chain - 2/8

- **Sampling**: Make stratified sample containing 25% target events and 75% non-target events
- **Transform Variables**: Bucket interval target to enhance data partition
- **Data Partition**: Split data to 60% training and 40% validation, test data not needed
  - use bucketed interval target to make sure both training and validation partitions have equal distribution on interval target
More Data Preparation

Analysis chain - 3/8

- **Filter Outliers**: Filter outliers based on statistics and business knowledge
  - knowing your data is mandatory!
- **Transform Variables**: Maximize normality and standardize variables to enhance modelling

- At this point we have dataset
  - we are familiar with
  - that has no outliers – we have the behaviour of the mass
  - which includes variables with ‘nice’ distributions

=> We are ready to do some serious modelling!
Regression Model

Analysis chain - 4/8
• **Regression**: Logistic regression
  • Stepwise selection method
  • Profit / Loss evaluation criteria
• **Assessment**: Results
  • Impressive Lift Value
  • Average profit 0,016386
  • Total profit 212,52

However, we can improve!
Two Stage Model Node

Analysis chain - 5/8
• Two Stage Model:
  • Stage 1 – activation probability: Decision Tree
  • Stage 2 – revenue from usage: Regression
• SAS Code: Model assessment
  • Adjust probability and profit calculations for separate sampling
  • Select subscriptions which satisfy
    \[ \text{activation\_probability} \times \text{predicted\_revenue} > \text{contact\_cost} \]

• Results
  • Average profit 0.035734
  • Total profit 463.43

=> Increase of 118% in profits by using two stage modelling!

However, we can improve!
Manual Two Stage Modelling – Stage 1

Analysis chain - 6/8

• **Variable Selection**: use R-square selection method to pre-select variables for modelling - override some rejections based on business knowledge

• **Regression**:
  • logistic regression with stepwise selection method
  • goal has shifted from making good decision to making an unbiased probability prediction - use **validation error** criteria to evaluate model fit
  • enhance with significance levels, forced selection, optimisation criteria…
  • tip: Neural Networks or Tree-models may be useful if relation between inputs and target is non-linear
  • tip: avoid using minimize resource usage –option!

---

SAS Forum International 2004 | Riku Mäkeläinen & Sakari Forslund TeliaSonera Finland
Preparation for Stage 2

Analysis chain - 7/8

• **Data Set Attributes**: modify metadata
  - model rejects variables not used, change status to input
  - coupling: add binary target predicted probability to list of inputs

• **Transform Variables**: log interval target
  - predicting negative usage amounts is not logical
  - by using log transformation we can correct error distribution of interval target and thus enhance model

• **Variable Selection**: use R-square selection method to pre select variables for modelling - override some selections based on business knowledge
Manual Two Stage Modelling – Stage 2

Analysis chain – 8/8
• Regression:
  • linear regression with stepwise selection method
  • log transformation helps to correct error distribution
  • tip: try Poisson or Gamma distribution
• SAS Code: Assessment
  • Average profit 0,038320
  • Total profit 496,97

⇒ increase of 7,24% over Two Stage Model node
⇒ increase of 134% over regular regression
Summary of Modeling Results

<table>
<thead>
<tr>
<th>Method</th>
<th>Total Revenue</th>
<th>Average Revenue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>212,52</td>
<td>0,016386</td>
</tr>
<tr>
<td>Two Stage Model</td>
<td>463,43</td>
<td>0,035734</td>
</tr>
<tr>
<td>Manual Two Stage Model</td>
<td>496,97</td>
<td>0,038320</td>
</tr>
</tbody>
</table>

Note: revenues shown represent only 4.68% of total due to sampling; magnitude of real profits are around 22 times higher

Evaluation

1. Two Stage Model node
   - easy and fast to use – model stability?
   - limitations - reduced control over models
   - increases profits by 118%

2. Manual two-stage model
   - modelling takes time
   - extensive control over models and data
   - Increases profits by 134%

⇒ Does increase in profits justify time required for manual modelling?
1. Introduction
2. Business Case
3. Two-Stage Modelling
4. Campaign Execution
5. Summary
• **Target Audience:** Customers who have a possibility to use MMS (handset & services)
• **Cell Splits:** Control model and message
• **Target Groups:** Measure either channels or messages OR personalise messages
Define Selection

- Use Two-Stage Model
  - data warehouse
  - score history
- Apply Contact Policies
  - restrict amount of campaigns e.g. no marketing in last two months
- Benefits of Contact Policies
  - make marketing evaluation easier and minimize overlaps in marketing programmes
  - minimize costs of campaigns and maximize profits
- Customer Satisfaction
Campaign Execution

• **Schedule**
  • load data & refresh score monthly
  • not event triggered

• **Results**
  • inferred responses
  • target & control group differences
  • level of usage - short term & long term

• **Programme**
  • one shot campaign not enough
  • repeating improves performance
Summary

• Definite improvement in target groups
• May require extra time - Balancing between results and requirements
• Automated campaigns – The way to efficiency
The Nordic and Baltic telecommunications leader