Who to Call, When to Call, What’s the Objective?

Mike Howard

www.AustinLogistics.com
Collections Environment
Overview

- Large credit banks have 50-100 million accounts
- 1% to 3% late on payment means large volume of decisions
- Action options: do nothing (self-cure), send letter, call, do letter and call, send account to a collections agency
Action Selection Approach

- Segment population by delinquency level
- Develop action-specific models to predict $\text{Prob}[\text{Cure} | \text{Action}]$ for each action
- Predict net revenue result if account cure results from a given action
- Use heuristics or optimization to assign actions to accounts
Action Specific Modeling

• Performance definition: make payment within 30 days = cure
• Models must provide cure probability estimates per account, action
• Independent samples, each treated entirely with a single action
• Methodology: Stepwise Binary Logistic Regression, Continuous Variables
Action Specific Modeling

- Modeling variables: credit bureau, call and delinquency history, demographics
- Model usage requirements: unbiased probability estimates
- Rebuild models every few months - harder estimation job than rank-ordering
- Typically 1-10 models
Rules-Based Action Selection

- Exclusions: pre-assign accounts with certain business conditions or probability values
- Select action on per account basis, based on comparing expected net revenue:
  - Prob[Cure|Action]*(Revenue from Cure due to that Action)
Optimal Action Selection

- Apply exclusions: business conditions, probabilities
- Present accounts set to optimizer, with expected net revenue per account per action
- Constraints: minimum and maximum number of accounts assigned to each action
Action Selection Example

CallSelect Optimization based on Net Revenue

Expected Total Net Revenue

- $1,650,000
- $1,675,000
- $1,700,000
- $1,725,000
- $1,750,000
- $1,775,000
- $1,800,000
- $1,825,000
- $1,850,000

% of Accounts Treated Internally

- 10%
- 20%
- 30%
- 40%
- 50%
- 60%
- 70%
- 80%
- 90%

- Final Objective
- Rules-Based

Final Objective

Rules-Based
Action Selection Method Trade

- Optimization is more effective, allows simulation, does not have to be customized for each installation, and can be used to evaluate efficiency of rules-based method.
- Rules-based approach is more familiar and comprehensible to many customers.
- Austin Logistics CallSelect (TM)
Calling Action Selection

How to Treat Accounts Selected for “Calling” - CallTech (TM)
Call Center Environment

• Automated dialing machines: continuous dialing and call queuing to operators
• Large number of operators - Insufficient
• Call results vary with time of day, day of week, nature of customer, state of account
• Legal restrictions on calling hours
Calling Action Options

- Do not call (insufficient calling resources)
- Time of week, time of day
- Strategy selection (segmentation) - typically by delinquency level, loan type
- Strategy resource allocation - divide calling resources up between strategies
Call Classification Modeling

- Performance classes: right party contact, wrong party contact, non-contact
- Models must provide probability estimates for each of the three outcomes
- Independent samples at lowest segmentation level
- Methodology: forward ternary logistic regression, continuous variables
Call Classification Modeling

- Modeling variables: call, payment, and delinquency history; phone and account data, demographics
- Rebuilt models every 1-2 months - dynamic environment, probability estimation
- Often > 100 models
- Automated Model Development
Promise-To-Pay Modeling

- Performance definition: given a right party contact, was a promise-to-pay made?
- Model usage: estimate probability of promise-to-pay for right party contacts
- Develop at higher level (sample size limits)
- Methodology: forward binary logistic regression, continuous variables
Daily Calling Action Selection

- Allocate calling resources to strategies
- Do optimal assignment over a set of time periods, per strategy
- Objective function based on \( \text{Prob}[\text{right party contact}] \)
- Compute expected consumption per time period and account using call outcome probabilities and average resource use
Daily Call Action Selection

Resources per Time Period

Accounts Data

Models per Time Period
TP1  TP2  ............. TPN

Resource Consumption Estimation

Assignment Process

Probability Estimates

File to Dialer
Daily Calling Action Selection

• Baseline objective: maximize expected right party contacts, sum of Prob[RPC] terms
• Possible weights on objective function terms: net revenue, delinquency level, default risk level
• Generalized linear network problem, integer programming problem solved with modified simplex method
Daily Selection Behavior

• Runs during night to set up dialer files
• Sensitive to accuracy of resource estimates
• Optimizer likes high Prob[RPC] accounts; outliers can severely bias the process
• Symptoms: overestimated right party contact rates
Daily Selection Behavior

• Result: Overestimate of resource consumption rate, disruption of calling schedule, loss of efficiency
• Other sources of same symptom: data quality deterioration, erroneous mixing of subpopulations, or significant shift in subpopulation behavior
Custom Modeling Limitations

- Usual custom modeling approach geared to rank-ordering of accounts, not probability estimation
- Causes for outliers: using zeros for missing values, overbuilding of models, spurious variables in models, data behavior shifts
- Outliers can increase dramatically when only one variable changes behavior
Modeling Methodology

- Compute means for use in application
- Automatically adjust significance thresholds based on sample size (smallest group)
- Truncate modeling candidate variables at 1% and 99% values
- Use information values to downsize variable set (eliminate spurious variables)
Modeling Methodology

• Compute development sample probability distributions
  – 95%, 99% and worst case values
  – Probability deciles and local fit
  – Identify outlier problems before models are deployed
Modeling Methodology

Development Sample Decile Fit

Prob[Right Party Contact]

Decile

Estimate
Actual
Model Application

- Use mean values when modeling variables are missing
- Truncate modeling variables at 1% and 99% values
- Truncate probability estimates at 99% value
- Problem diagnostics: collect data at model sample level, compute fit to development deciles
Model Application

• Use summary measures of performance

• Diagnostics when deciles match poorly:
  – Review subpopulation behavior
  – Compare frequency distribution of modeling variables to development sample
  – Most problems due to data deterioration, improper mixing of subpopulations, or population shifts
Conclusion

• Complex modeling problems can be automated, but unique application requirements must be identified and addressed

• Higher degrees of automation are needed to scale application products to a large customer base

• The process can never be fully “hands-off”