



A Logistic Regression Model for Consumer Responses to Email Campaigns

October 16th, 2013

Toronto Data mining Forum

Jiarui (Carrie) Dang

QUADS Solutions, LoyaltyOne

Agenda

1. Who we are
2. Background
3. Data preparation
4. Logistic regression model
5. Solutions
6. Summary
7. Questions



Who we are

Loyalty programs are being leveraged across industries

- A loyalty program is
...an invitation to begin a long and profitable relationship
- Three loyalty formats

Proprietary



Partnership



Coalition



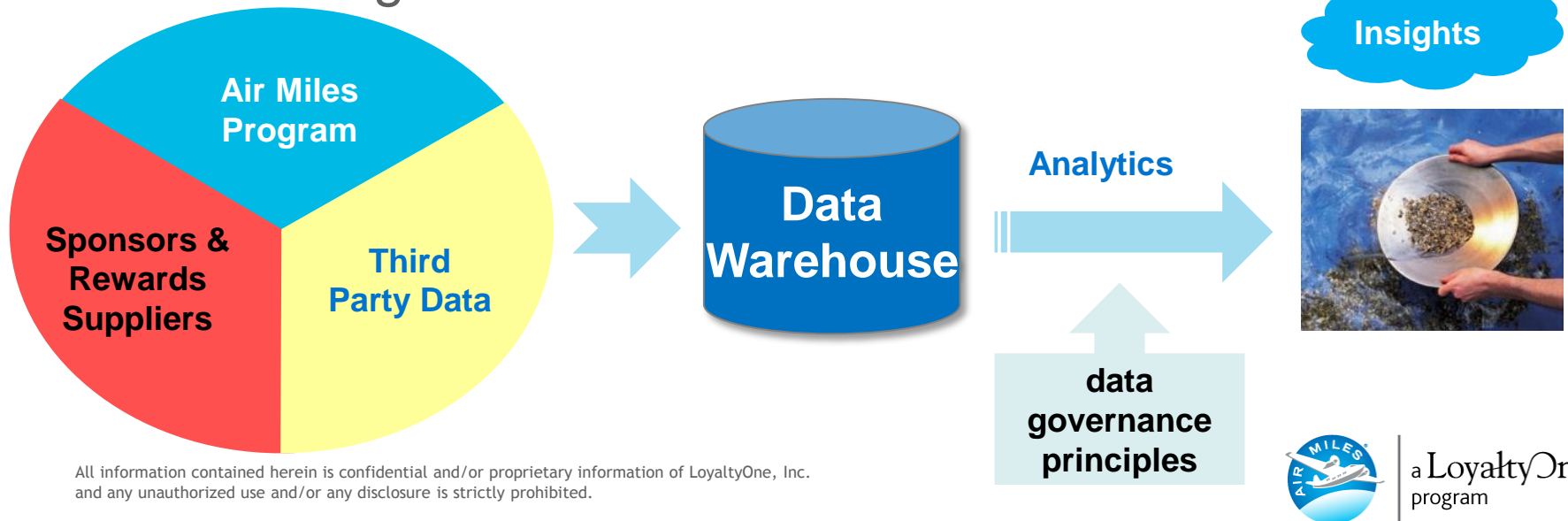
- Coalition benefits to collectors and sponsors



Who we are

Quantitative Analytics and Decision Sciences

- Goal
 - Understand business
 - Change consumer behavior
- Mission
 - Apply our skill set to support internal/external clients
 - Involves profiling, analyzing, modeling, and researching
- Data to insights



All information contained herein is confidential and/or proprietary information of LoyaltyOne, Inc. and any unauthorized use and/or any disclosure is strictly prohibited.



a LoyaltyOne
program

Background

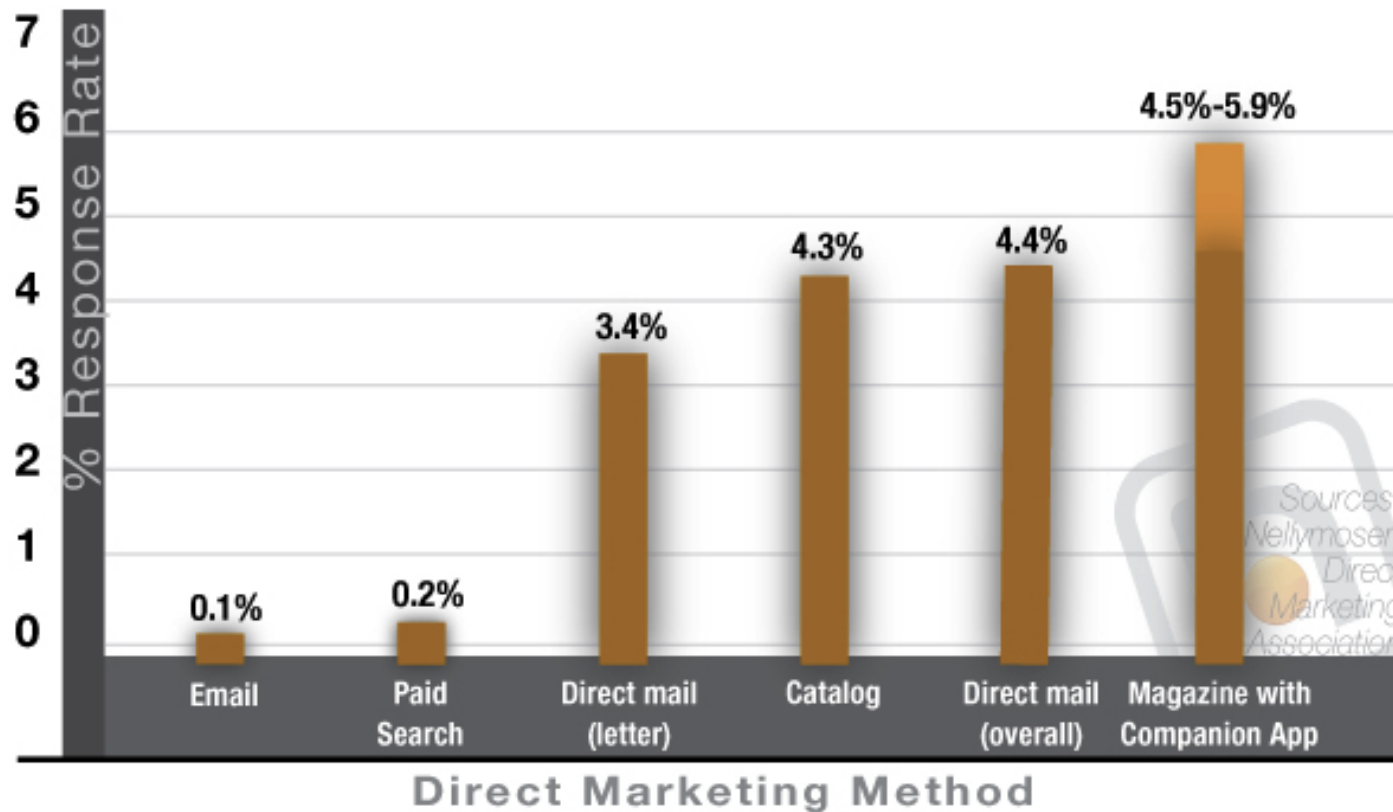
- E - Marketing campaigns
 - Display ads, Facebook or LinkedIn ads, Paid search, Email, etc.
 - Enhance the relationship with current or previous customers
 - Encourage customer loyalty and repeat business
- Compared to traditional mail
 - Advantages
 - Reduces delays in communication
 - Cheaper and faster
 - Can reach substantial numbers of email subscribers
 - Disadvantages
 - Low response rate
 - Messages being rejected or filtered
 - Risk of subscribers opting-out



Background

Marketing campaign response rates

- A research on the U.S. market by Nellymoser



Background

2013 Email statistics for UK SME

- Compiled by Dave Chaffey
- Across all industries the average results for UK SME email marketing campaigns
 - Open rate: 21.47%
 - Click-through rate: 3.16%
 - Unsubscription rate: 0.47%
 - Click-to-open rate: 14.72%
 - Unsubscribe-to-open rate: 2.29%



Logistic Regression Model

- Introduction
 - Predict the outcome of a categorical dependent variable
 - Maximum likelihood
- Types of logistic regression by response level
 - Binary
e.g., event vs. non-event
 - Ordinal
e.g., small, medium, large
 - Nominal
e.g., Sprite, Coke, Pepsi



Data Preparation

- Predictive vs. Descriptive
 - Quick turnaround
- Aggregated data
 - Low response rate
- Predictors
 - Air Miles
 - Demographics
 - Sponsors
 - Point of sale
 - Redemptions
 - Third party
 - Additional variables

# Responses	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	xxx	54.55	xxx	54.55
2	xxx	20.24	xxx	74.79
3	xxx	10.1	xxx	84.89
4	xxx	5.82	xxx	90.71
5	xxx	3.5	xxx	94.21
6	xxx	2.24	xxx	96.44
7	xxx	1.43	xxx	97.87
8	xxx	0.88	xxx	98.75
9	xxx	0.47	xxx	99.22
10	xxx	0.27	xxx	99.49
...
159	xxx	0	xxx	100



Building Model - Basics

- Study period
 - Cutoff date
 - Prior one year
- Sample
 - All customers who were contacted during the study period
 - Active customers
 - Divide into two sets
 - Training
 - Validation
- Event
 - Responded: 1
 - Not responded: 0



Building Model – Two Approaches

- Response level
 - Binary approach

Target	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	xxx	91.04	xxx	91.04
1	xxx	8.96	xxx	100

- Ordinal approach

Target	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	xxx	91.04	xxx	91.04
1	xxx	6.52	xxx	97.57
2+	xxx	2.43	xxx	100



Building Model – SAS Code

- Binary

```
proc logistic data = training desc outest = param  
    outmodel=model1;  
    model dependent = independent / scale=none  
    lackfit stb selection=stepwise;  
    output out=results reschi=Pearson;  
run;
```

- Ordinal

- Automatic

- Nominal

- Option: link = glogit



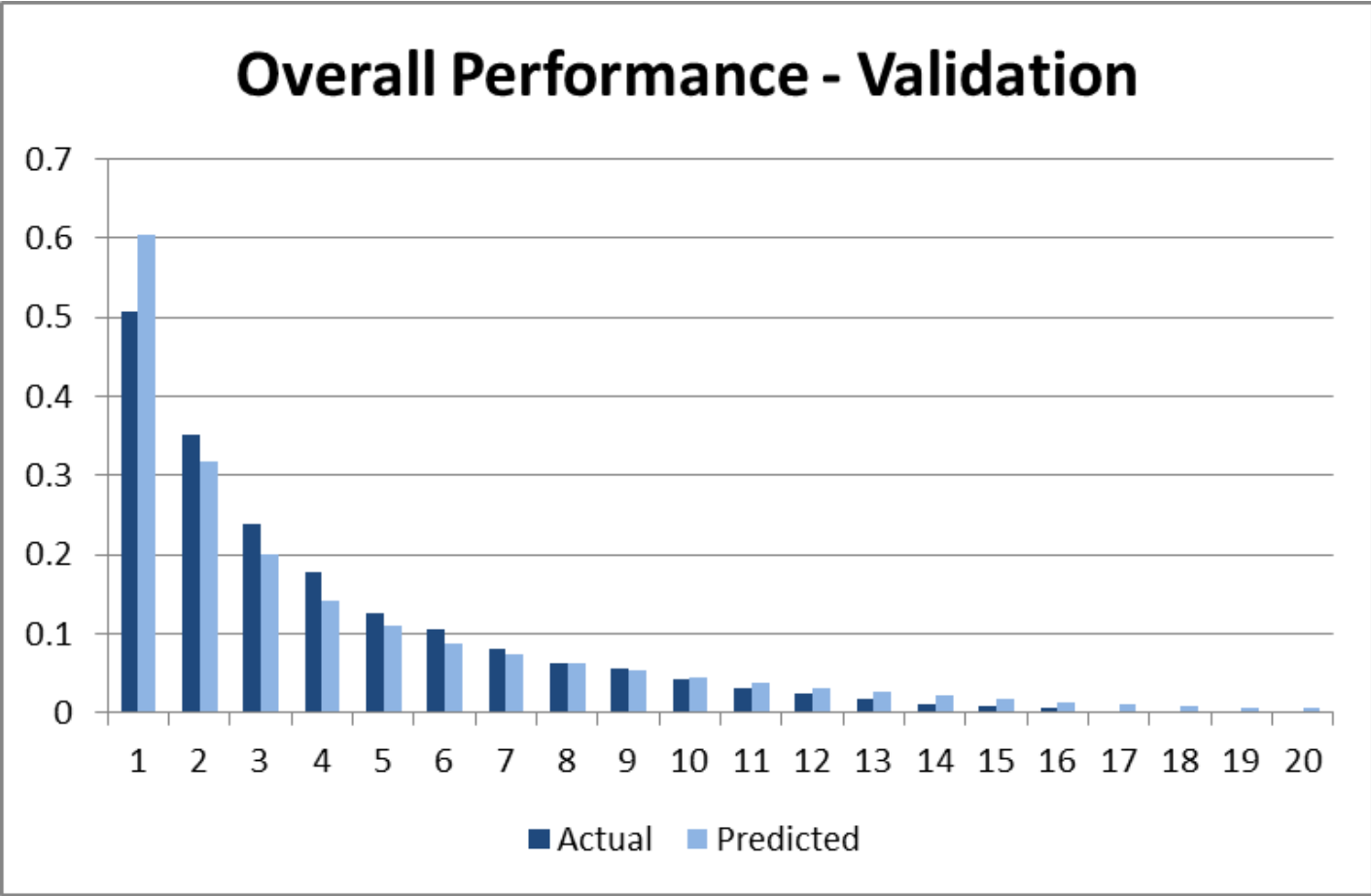
Solutions - Lift

Decile	Num. of Collectors	Num. of Responses	Actual	Predicted	Cum. % Captured	Lift	Cum. Lift
1	xxx	xxx	0.50832	0.60449	27.52%	5.50	5.50
2	xxx	xxx	0.35083	0.31672	46.51%	3.80	4.65
3	xxx	xxx	0.23902	0.19921	59.45%	2.59	3.96
4	xxx	xxx	0.17735	0.14198	69.05%	1.92	3.45
5	xxx	xxx	0.12565	0.10954	75.85%	1.36	3.03
6	xxx	xxx	0.10564	0.08823	81.57%	1.14	2.72
...
19	xxx	xxx	0.00016	0.00673	99.99%	0.00	1.05
20	xxx	xxx	0.00016	0.0049	100.00%	0.00	1.00



Solutions

Overall Performance



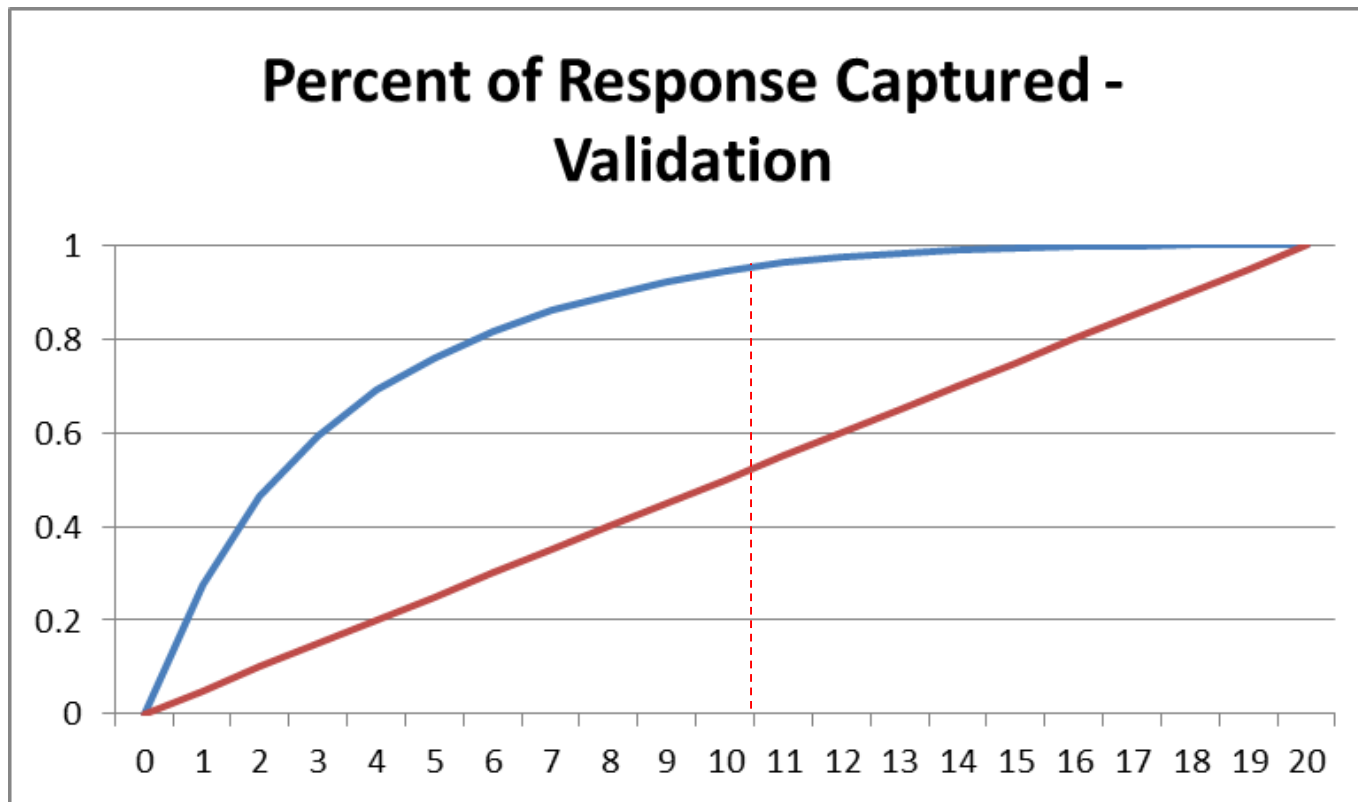
All information contained herein is confidential and/or proprietary information of LoyaltyOne, Inc. and any unauthorized use and/or any disclosure is strictly prohibited.



Solutions

Percent of Responses Captured

- Percent Concordant: 86.8



Summary

- Binary logistic regression
 - Straightforward
- Multiple response level
 - Helps us better understand multiple vs. one-time responders
 - Proportional odds assumption
 - Nominal logistic regression
- Next steps
 - Validation by future campaigns
 - Periodical re-calibration
 - Fine tune model in ever-changing marketplace



Questions?



All information contained herein is confidential and/or proprietary information of LoyaltyOne, Inc. and any unauthorized use and/or any disclosure is strictly prohibited.



a LoyaltyOne
program



AllianceData.

LoyaltyOne

© 2012 LoyaltyOne, Inc. All rights reserved.

No part of this material may be produced in any form without the written permission of the owner. No representation or warranty of any kind is made with respect to the content hereof and no liability is assumed for errors contained herein or for incidental or consequential damages in connection with the furnishing, performance or use of this publication or its contents.