Net Lift Modeling

Kim Larsen, VP Analytical Insights, MarketShare

Introduction

- The true impact of a marketing campaign or promotion is measured by its incremental impact
- However, targeting criteria are often not designed to maximize the incremental impact
- Net lift models are designed to maximize the incremental impact by targeting the undecided “swing clients”

A case study

- Objective: Web-based campaign to sell a specific product
- Targeting: Contact clients visiting the product's web page

The campaign result

- Test design
  - Test group: Received an offer
  - Control group: Did not receive an offer

- The overall client 90-day purchase rate was 1.5%

Cell	Purchase Rate
Test	5.01%
Control	5.00%

- Net purchase rate = 5.01% - 5.00% = 0.01%

Why did we not see any campaign lift?

Not interested → Will never purchase the product. No point in marketing to them

Self-selectors → Likely to purchase the product on their own. Marketing could even have an adverse effect. **The campaign targeted too many of these clients**

Swing clients → Interested in the product, but need to be motivated to buy it. **Target more of these clients**

The solution: A net lift model

	Purchase Rate		
Net Score	Test (gross)	Control	Net
Top 20%	6.10%	3.9%	2.20%
Lowest 80%	4.75%	5.28%	-0.48%

Net lift models versus propensity models

Net Purchase Rate = **Test group purchase rate** – **Control group purchase rate**

(Gross purchase rate) (Self-selection purchase rate)

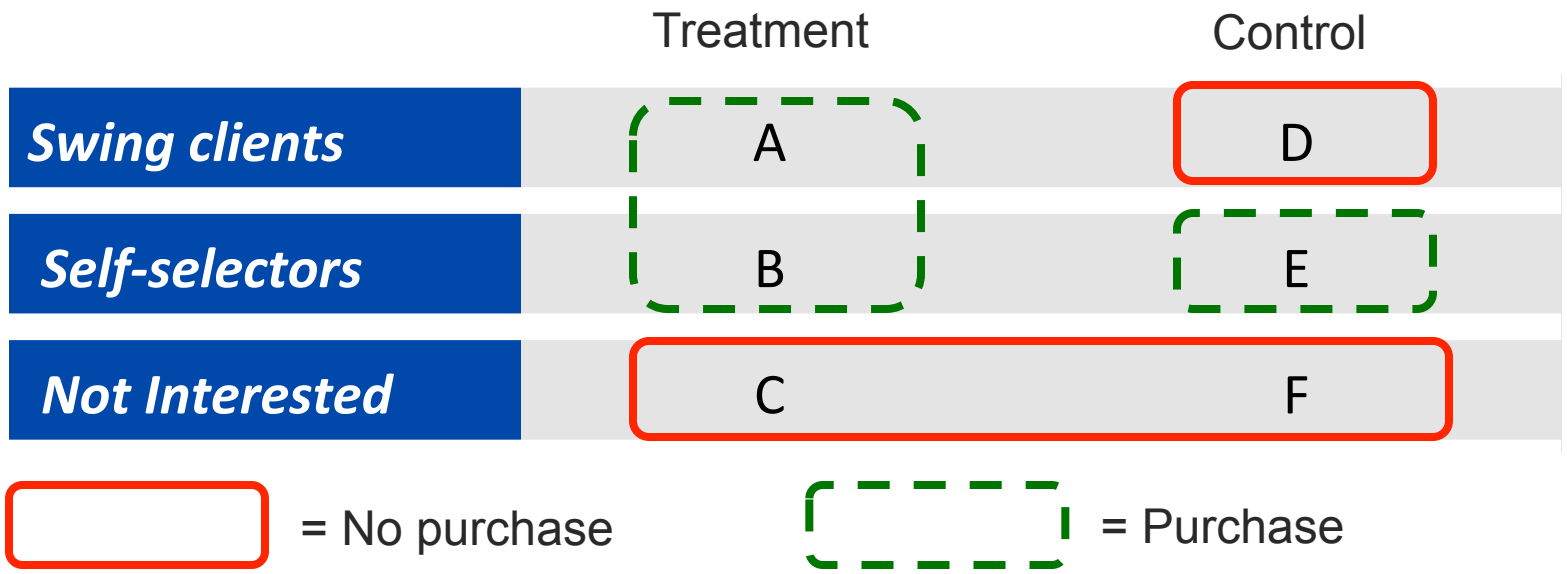
Propensity Model

- Most common approach
- Targets the clients with the highest probability of making a purchase following a marketing contact
- Maximizes the gross purchase rate

Net Lift Model

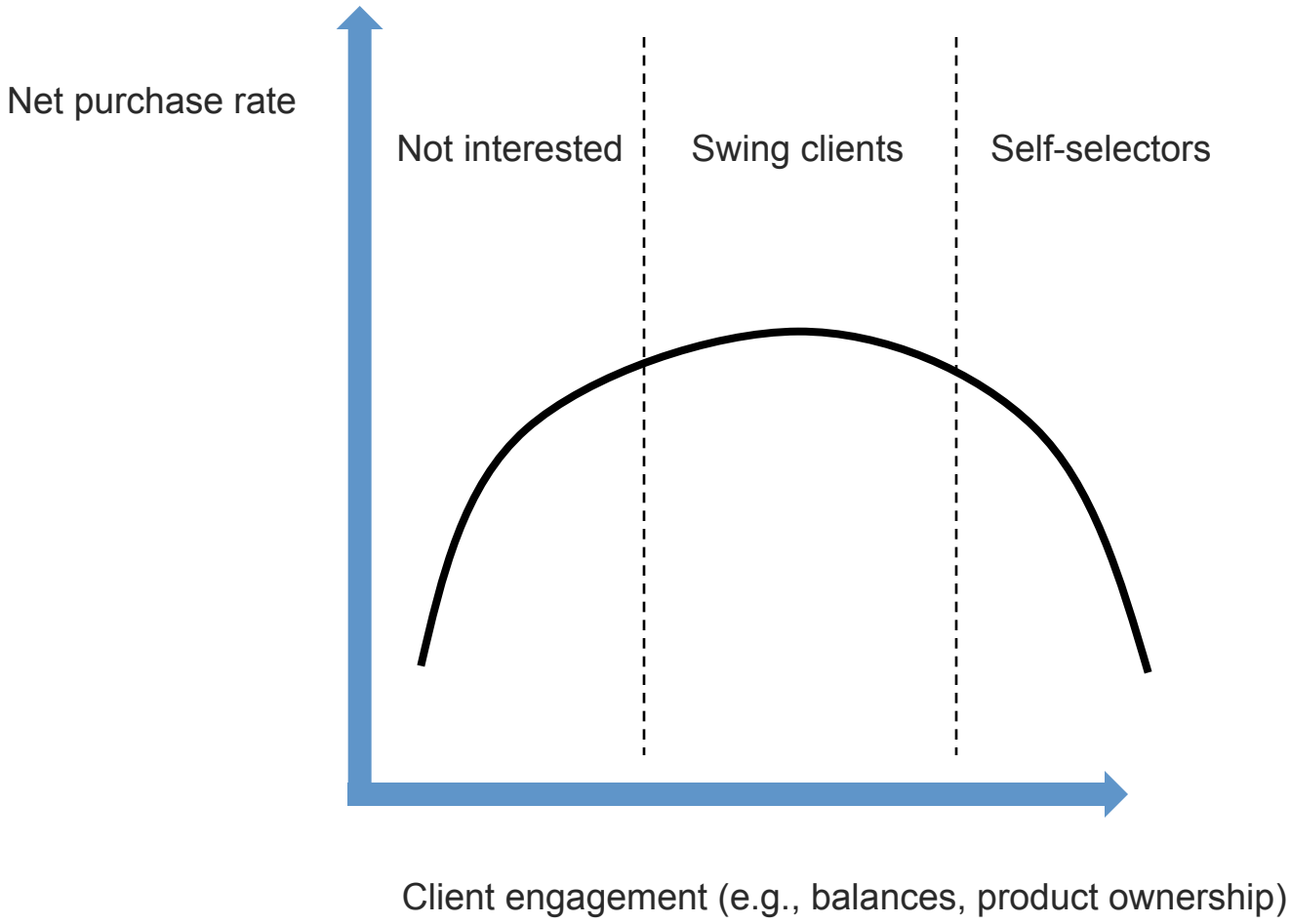
- Targets the undecided clients that can be motivated by marketing
- Maximizes the incremental purchase rate

Challenges when building net lift models



- We cannot observe cell “A” directly from the data
- The objective function is a difference of two rates – double variance problem

Nonlinearity is common in net lift modeling



Overview of key net lift modeling techniques

Regression-based techniques	**Difference score models** **Probability decomposition models** **Bifurcated logistic regression**
Non-regression techniques	**KNN classifiers** **Naïve Bayes** **Classification trees**

The most popular regression-based technique –*The Difference Score Model*

- Find a target group, T, such that:

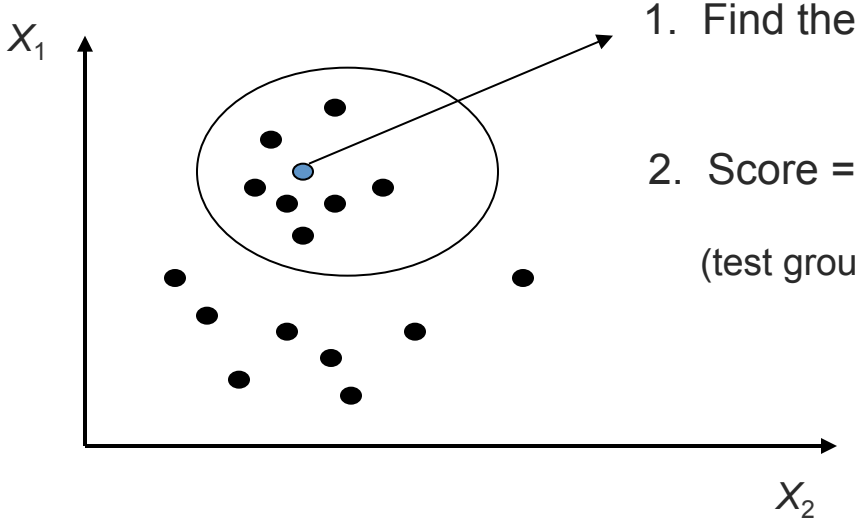
$$\max\left\{\sum_{i\in T}\left(E(Y_i \mid \text{Offer}) - E(Y_i \mid \text{No offer})\right)\right\}$$

Estimated through a logistic regression model:
P(Purchase | Offer)

Estimated through a logistic regression model:
P(Purchase | No offer)

- Score = P(Purchase | Offer) – P(Purchase | No offer)

A mathematically appealing non-regression approach to net lift modeling –*The KNN Classifier*



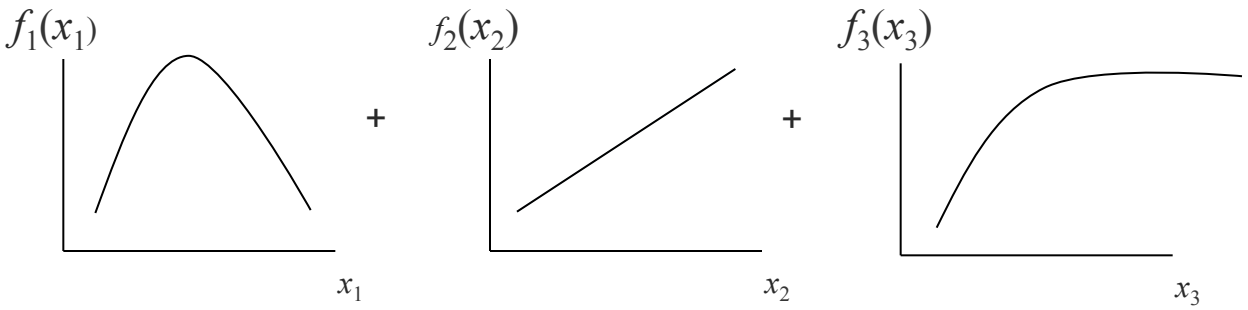
1. Find the K nearest neighbors for the client to be scored

2. Score = net purchase rate for the neighborhood

   (test group purchase rate – control group purchase rate)

- Fits the target directly and handles all types of nonlinearity
- KNN models are not transparent. Additional "post-model" analysis is needed to describe the models
- Implementation is computationally intensive

A cousin of the KNN Classifier – *the Net Naïve Bayes classifier*

Net Score = 

Smooth functions that reflect the *Net WOE* – i.e., the log-odds ratio of a purchase for test versus control

Estimated with *single dimension* KNN smoothers

Applying five different methods to the case study

Net model method	Net purchase rate (top 2 deciles)
Probability Decomposition Model (using adaptive[1] logistic regression)	4.8%
Difference Score Model (using adaptive[1] logistic regression)	4.6%
Generalized Net Naïve Bayes	4.4%
Net Naïve Bayes	3.5%
KNN Classifier (K=100)	2.3%
Linear net difference score	1.8%

[1] Using the Gains# software (www.infodecipher.com)