Cross-selling and Up-selling: Learning from Purchases

Professor Song Yao
Olin Business School

Customer Analytics

Whom vs. What? A Different Setting for Targeting

So far we have been interested in the response to a given product offering

COMPARISON OF QUESTIONS

- Given a certain product we want to offer, which customers are the most likely to respond to the offer?
  - Whom to sell?

- Given that we would like to target a certain customer, which product is the most likely to lead to a positive response (or maximize profits)?
  - What to sell?

So far we have been interested in the response to a given product offering

COMPARISON OF QUESTIONS

- Given a certain product we want to offer, which customers are the most likely to respond to the offer?
  - Whom to sell?
  - Expected return needs to exceed cost of marketing

- Given that we would like to target a certain customer, which product is the most likely to lead to a positive response (or maximize profits)?
  - What to sell?
  - Expected return from Offer 1 needs to exceed return from Offer 2, 3, …

Model of choice when marketing costs are not the "limiting resource"

Such models have many applications

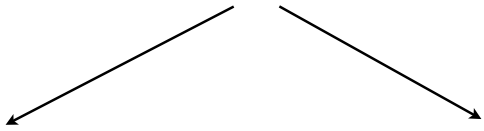
APPLICATIONS

- Banks:
  - Which financial product to offer next?
- Call centers:
  - Which additional product to offer during an inbound call?
- Online retailers:
  - Which product to promote in a weekly e-mail
  - Which product recommendation to offer on the home page
- B2B:
  - Which product to push in the next sales call
- ...

To choose a model, we need to first determine our setting

APPLICATIONS OF MODELS

What is the scope of the prediction?



Situational:

"If a consumer is currently considering product X, what product Y should we offer?"

General:

"Considering a consumers entire purchase history, what is the next product we should offer to the consumer?"

To choose a model, we need to first determine our setting

APPLICATIONS OF MODELS

What is the scope of the prediction?

Situational:

"If a consumer is currently considering product X, what product Y should we offer?"

General:

"Considering a consumers entire purchase history, what is the next product we should offer to the consumer?"

amazon

Hello Song,

Thank you for shopping with us. We'll send a confirmation when your item ships.

Order Confirmation

Arriving:
Monday, January 13

Ship to:
Song
OLIVETTE, MO

Order #
111-2705944-8501837

View or manage order

Nature Made Vitamin D3 2000...
Qty : 1

Order Total: **$8.84**

Keep shopping for

Petmate 2 Cup Pet Food Scoop With...
$3.82 prime

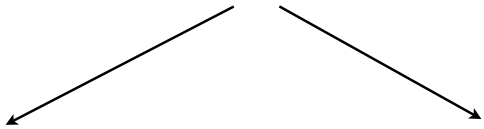
Kleenex® Professional Facial Tissue,...
$27.99 prime

Situational Cross-Selling/ Upselling

To decide on a model we need to first determine our setting

APPLICATIONS OF MODELS

What is the scope of the prediction?



Situational:

"If a consumer is currently considering product X, what product Y should we offer?"

General:

"Considering a consumers entire purchase history, what is the next product we should offer to the consumer?"

The most common method for generating situational recommendations is a "Market Basket Analysis" (MBA)

MARKET BASKET ANALYSIS

Basket ID	Purchases
1	pizza, soda
2	milk, cleaner
3	soda, pizza, detergent
4	pizza, detergent
5	cleaner, soda
6	pizza, cleaner, soda
7	pizza, soda
8	cleaner, detergent
9	soda, pizza

basket	cleaner	deter-gent	milk	pizza	soda
1	0	0	0	1	1
2	1	0	1	0	0
3	0	1	0	1	1
4	0	1	0	1	0
5	1	0	0	0	1
6	1	0	0	1	1
7	0	0	0	1	1
8	1	1	0	0	0
9	0	0	0	1	1

Goal: Implement a rule "If buy/consider product A then offer product B "

For each pair of products we calculate three key measures

MARKET BASKET ANALYSIS

1. How likely is this rule to apply?

$$Support(A, B) = \frac{\#\ (A \text{ and } B)}{\#\ Orders}$$

2. How likely is product B purchased after a consumer has purchased A?

$$Confidence(A, B) = \frac{\#\ (A \text{ and } B)}{\#\ A}$$

3. Are the relationship between A and B coincidental (e.g., A and B are both very popular but unrelated)

$$Lift(A, B) = \frac{Support(A \text{ and } B)}{Support(A) \times Support(B)}$$

For each pair of products we calculate three key measures

MARKET BASKET ANALYSIS

- Assume there are 100 customers
- 10 of them bought milk, 8 bought butter and 6 bought both of them.
- Rule: bought milk => recommend butter?
- support = #(Milk & Butter)/#(Orders) = 6/100 = 0.06
- confidence = #(Milk & Butter)/#(Milk) = 6/10 = 0.60
- lift = support(Milk, Butter)/[support(Milk) * support(Butter)]

  = 0.06/(0.10 * 0.08) = 7.5

The three measures suggest whether to use a rule and if so, what rule to use

IMPLEMENTING RULES IN MBA

Goal: Implement a rule "If buy/consider product A then offer product B "

- Statistical significance (and if relevant, sufficient "support") is a necessary condition for implementing this rule
- If "Lift" and "Confidence" are "high enough," then implement rule.
- If multiple rules pass the hurdle (if A then B, if A then C),
  - Recommend multiple products to buy next (e.g., Amazon recommends 2)
  - Profit comparison (i.e., see which recommendation results in higher profits)

There are various other considerations when using MBA

CONSIDERATION IN MARKET BASKED ANALYSIS (2)

- Market Basket Analysis can be easily extended to more complicated condition clauses
  - If A1 and A2 then B –> form a super-product that is A1 and A2 = A1*A2
  - If A not bought then B –> form an anti-product "anti-A"=1-A
  - If A bought, then not B –> form an anti-product "anti-B"=1-B

basket	cleaner	detergent	milk	pizza	soda	cl_det	anti_det	anti_soda
1	0	0	0	1	1	0	1	0
2	1	0	1	0	0	0	1	1
3	0	1	0	1	1	0	0	0
4	0	1	0	1	0	0	0	1
5	1	0	0	0	1	0	1	0
6	1	0	0	1	1	0	1	0
7	0	0	0	1	1	0	1	0
8	1	1	0	0	0	1	0	1
9	0	0	0	1	1	0	1	0

There is another crucial consideration before using MBA

CONSIDERATION IN MARKET BASKED ANALYSIS

The data generating process differs from the prediction setting

- Targeting:
  - Offer first sent to a randomly selected sample
  - Based on observed responses, we use RFM/Logistic/Machine Learning to predict response rate of the other customers outside of the sample
  - The same offer is then sent to those people with high response rates
- Market Basket Analysis
  - There is no data collected with recommendations
  - We use existing market data (without recommendation) to predict what would happen if we started to recommend
  - The assumption is that the behavior of consuming A and B together does not change due to the recommendation

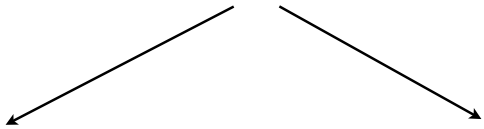
--> **Needs testing and adjusting**

General Cross-selling/Upselling

To decide on a model we need to first determine our setting

APPLICATIONS OF MODELS

What is the scope of the prediction?



Situational:

"If a consumer is currently considering product X, what product Y should we offer?"

General:

"Considering a consumers entire purchase history, what is the next product we should offer to the consumer?"

How about using a Market Basket Analysis?

USING MBA WITH FULL PURCHASE HISTORIES

- Market Basket Analysis makes recommendations looking at only one purchase (or look) at a time
  - **"If customer has purchased/looked at product A then offer customer product B"**
- How do we make recommendations based on a collection of purchases?
- Need:
  - **"If customer has purchased products A1, A2, A3, then offer customer product B"**
- Need to have observed enough people who purchased products A1, A2, A3, and B

One key approach is to Market Basket Analysis on product pairs and combine/sort recommendations

AMAZON'S (*BASIC*) APPROACH (PATENT 6,266,649)

$$s_{A,B} = \frac{n_{A,B}}{\sqrt{n_A * n_B}}$$

$$n_{A,B} = 10, n_A = 100, n_B = 100 \Rightarrow s_{A,B} = 0.1$$
$$n_{A,B} = 20, n_A = 25, n_B = 25 \Rightarrow s_{A,B} = 0.8$$
$$n_{A,B} = 20, n_A = 20, n_B = 20 \Rightarrow s_{A,B} = 1$$

- *Key idea: Reduce a multi-product problem into a sequence of pairwise problems*
- *Define **similarity** between two product as:*
  - *$n_{A,B}$ is number of times that products A and B are purchased by a customer*
  - *n_A is total number of times product A is purchased; similar for n_B*
- *Suppose Bob has purchased product A, B, C, and D previously.*
- *Place product E on a "short-list" if the similarity between E and any of product A, B, C, and D is sufficiently high. Repeat for products F, G, H, etc.*
- *Sort products on the short list by the highest similarity score an item has with any of the items in the purchase history*
- *Combine and sort short lists*

Consider an example that uses this approach

EXAMPLE OF MBA WITH FULL PURCHASE HISTORIES

- Customer has purchased photo products in the past
- Digital photography department wants to make e-mail/website offer for camera accessory

Purchase history	Potential accessories
• Canon S95 digital camera • 4 GB SD card • Nikon D80 digital SLR camera	• Eye-Fi Wifi SD card • Lens cleaning kit • Camera case (universal) • Battery (S95) • External flash • Lens cap • Book: "Understanding Close-ups"

Consider an example that uses this approach

SIMILARITY RATINGS $s_{A,B}$ AND SHORT LISTS

Canon S95 digital camera		4 GB SD card		Nikon D80 digital SLR camera	
• **Eye-Fi Wifi SD card**	**0.03**	• Eye-Fi Wifi SD card	0.00	• **Eye-Fi Wifi SD card**	**0.06**
• Lens cleaning kit	0.001	• Lens cleaning kit	0.00	• **Lens cleaning kit**	**0.21**
• **Camera case (universal)**	**0.12**	• **Camera case (universal)**	**0.05**	• Camera case (universal)	0.004
• **Battery (S95)**	**0.08**	• **Battery (S95)**	**0.03**	• Battery (S95)	0.00
• External flash	0.00	• External flash	0.00	• **External flash**	**0.14**
• Lens cap	0.00	• Lens cap	0.02	• Lens cap	0.00
• "Understanding Close-ups"	0.02	• "Understanding Close-ups"	0.01	• "Understanding Close-ups"	0.02

COMBINED SHORT LISTS

- **Eye-Fi Wifi SD card** 0.03, **0.06**
- **Lens cleaning kit** **0.21**
- **Camera case (universal)** **0.12**, 0.05
- **Battery (S95)** **0.08**, 0.03
- **External flash** **0.14**

SORTED FINAL LIST

- **Lens cleaning kit** **0.21**
- **External flash** **0.14**
- **Camera case (universal)** **0.12**
- **Battery (S95)** **0.08**
- **Eye-Fi Wifi SD card** **0.06**

Individually Customized

Market Basket Analysis uses only data on basket IDs, products, and sometimes customer IDs

DATA USED FOR MBA

Basket ID	Product ID	Customer ID
1	Nikon D80	10045
1	Eye-Fi SD	10045
1	Lens Cap	10045
2	Canon A80	38930
2	Battery (A80)	38930
3	External Flash	10045
4	Canon S95	98543
4	4GB SD	98543
4	Cam. Case	98543
5	Lens Cleaning Kit	38930

Market Basket Analysis uses only data on basket IDs, products, and sometimes customer IDs

TYPICALLY AVAILABLE PURCHASE-RELATED DATA

			Transaction Details			Buyer Details		
Basket ID	Product ID	Customer ID	Time	Price	Other...	Age	Income	Other...
1	Nikon D80	10045	3/23/2007	$784	Searched	34	50-60K	10+ reviews
1	Eye-Fi SD	10045	3/23/2007	$49	Promotion	34	50-60K	10+ reviews
1	Lens Cap	10045	3/23/2007	$18	Front page	34	50-60K	10+ reviews
2	Canon A80	38930	10/4/2008	$199	...	23	10-20K	4 returns
2	Battery (A80)	38930	10/4/2008	$46	...	23	10-20K	4 returns
3	External Flash	10045	12/1/2010	$110	...	52	40-50K	...
4	Canon S95	98543	1/13/2011	$399	...	65	90-100K	...
4	4GB SD	98543	1/13/2011	$24	...	65	90-100K	...
4	Cam. Case	98543	1/13/2011	$35	...	65	90-100K	...
5	Lens Cleaning Kit	38930	2/2/2011	$5	...	23	10-20K	...

We can better exploit purchase data by building a model that relates current choices to rich descriptions of past behavior

DATA REQUIREMENTS FOR RICHER MODEL

		t-4	t-3	t-2	t-1	t	Time
Customer 10045:	Buyer descriptors	A		A	B	A	
Customer 38930:	Buyer descriptors						
Customer 10045:	Buyer descriptors	A	D			B	
Customer 98543:	Buyer descriptors	C		C	D	D	
Customer ... :	...	B	B				
Customer ... :	...						
Customer ... :	...						
Customer ... :	...	A		D		C	

Independent Variables **Dep. Variables**

We can use a variety of approaches to estimate a richer model

MODEL ALTERNATIVES FOR MODEL

- Binary Logit with different product offers
  - Different consumers are offered different products
  - Predict for each consumer the probability of choosing each product
  - Used in Pentathlon Part III e-mail customization case
- Multinomial Logit (and Nested Logit)
  - Like logistic regression but dependent variable is the chosen product (J values)
- Machine Learning models (multi-class)

Let's look at an example of a cross/upselling model using a binary logit model

BBB NEXT-PRODUCT-TO-BUY EXAMPLE

- Stan Lawton (marketing director) prepares for **e-mail marketing** and the problem!
- Sends out one of three offers to 10,000 consumers each:
  - Offering in the art category: "The Art History of Florence."
  - Offering in the do-it-yourself category: "Painting Like a Pro."
  - Offering in the cooking category: "Vegetarian Cooking for Everyone."
- Profit varies between books:
  - "The Art History of Florence" --> $6
  - "Painting Like a Pro" --> $4
  - "Vegetarian Cooking for Everyone" --> $7
- Cost of making the offer is irrelevant (e-mail and the frequency has been set)
- **Key problem: Which book offer is the best match for each customer?**

Demo Code