Data based Targeting collaboration Retailor X Manufacturer

- Insight based Category Look-alike
- Propensity Modeling

Junyoung Lee
DaeKyung Lim

P&G

ULTRA Downy

Pampers

BRAUN

Gillette

head & shoulders

Oral-B

febreze

SK-II

Gillette Venus

PANTENE

Junyoung Lee

P&G Analytic & Insights Manager

Carlson Funds Enterprise

Fixed Income Analyst

Pfizer

Corporate & Marketing Strategy

P&G

Analytic & Insights

DaeKyung Lim

P&G Data Science Team Leader

SK C&C

IBM **GBS Consultant**

coupang Color Your Days **Product Reco. Team**

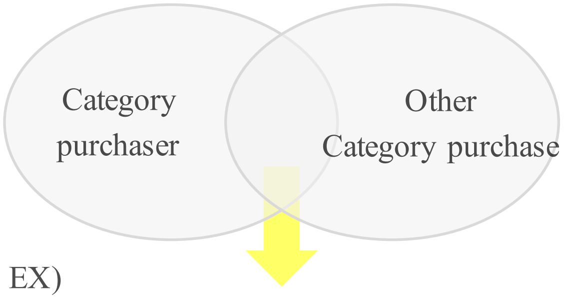
P&G **Data Science Team**

Project overview

Data analysis & targeting projects

Drive Category growth & recruit new users

Phase 1) Insight based shopper clustering



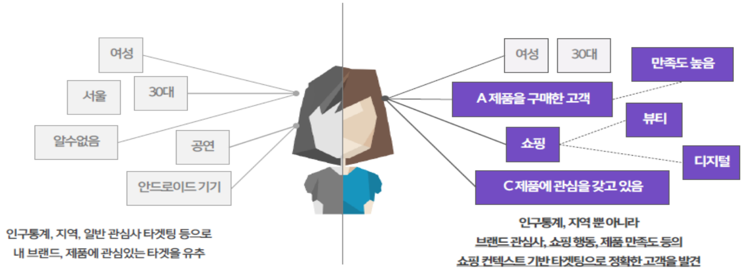
Understand the category growth shoppers characteristic

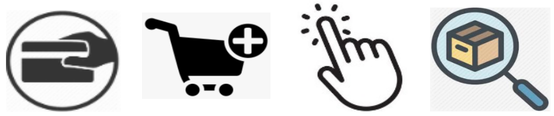
EX)

[Sports LOVER] [New hired] [Grooming]

Develop the promotion theme & target to wide rage of shoppers

Phase 2) Propensity modeling





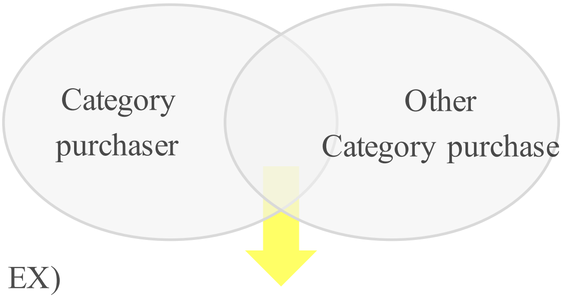
Sorting out shoppers who have high probability to become new users

Project overview

Data analysis & targeting projects

Drive Category growth & recruit new users

Phase 1) Insight based shopper clustering



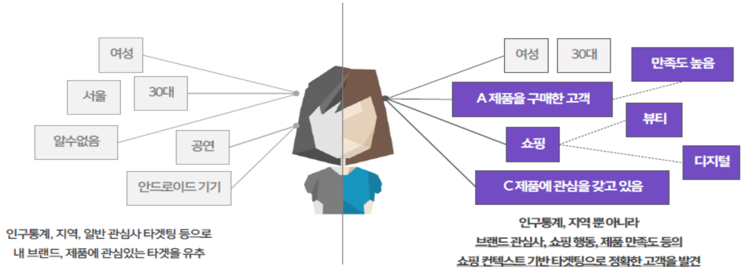
Understand the category growth shoppers characteristic

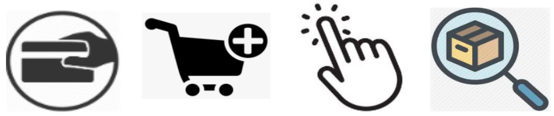
EX)

[Sports LOVER] [New hired] [Grooming]

Develop the promotion theme & target to wide rage of shoppers

Phase 2) Propensity modeling





Sorting out shoppers who have high probability to become new users

WHAT THEY PURCHASE – Category adjacency

How to analyze

*Top selling category = Mean (Log(Purchase)) >Mean (Log(Purchase)) +(Stad (Log(Purchase)) * 2)

Men shopper's purchase category

	Category
1	A
2	B
3	C
4	D
5	E
6	.
7	.
8	.
9	.
10	.
11	.
12	.
13	.
14	.
15	.
.	.
.	.
.	.
.	.
.	.
.	.
.	.
.	.
.	.

1 Identifying Top selling category* among men

2 Identifying unique characteristic of BnR category users.

Exclude out categories that show high rank in TTL Men shoppers purchase category from BnR Shoppers Top selling category

Shave Shopper's Purchase category

	Category
1	X
2	1
3	Z
4	A
5	C
6	Y
7	ㄱ
8	D
9	2
10	ㄴ
11	B
12	.
13	.
14	.
15	.
.	.
.	.
.	.
.	.
.	.
.	.
.	.
.	.
.	.

	Category
1	X
3	Z
6	Y

	Category
6	1
7	2

	Category
9	ㄱ
10	ㄴ

3 Grouping the categories that shows similarity & commonality

4 Target 'Look-alike' shoppers with customize communication

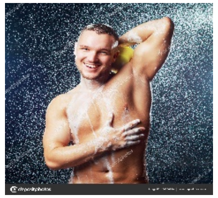
Target users = Category X/Z/Y buyers – BnR purchaser

Target users = Category 1/2 buyers – BnR purchaser

Target users = Category 1/2 buyers – BnR purchaser

WHAT THEY PURCHASE – Correlation

Ever purchaser 1 year – 6 month Non-shave purchaser

Target Characteristic	 Cost effective Fashion Seeker	 Scent Sensitive	 Extreme Skincare	 Clean & Neat	 Convenient Food / No cook	 Health Conscious
Purchase Category	티셔츠 점퍼 자켓 셔츠/남방 바지/팬츠 청바지 맨투맨 양말	디퓨저 탈취제 섬유유연제 세차/관리용품	올인원 스킨케어 남성화장품세트 팩/마스크 스킨케어 세트 선크림/선블록	샴푸 바디워시 바디로션 칫솔 치약 구강청정제/가글	냉장/냉동식품 푸드/배달 쿠폰 통조림/캔 베이커리/도넛 과자	수입쇠고기 닭고기 영양제 탄산수 주스/과즙음료 과일 스포츠 의류 반팔 운동화/스니커즈
Category Relevancy (% among 11st system purchaser)	36.9%	25.4%	24.2%	34.2%	23.9%	12.3%
Gillette Preference (Gillette Over/under-development)	→	↑	↑	↑	↓	↓

Potential Actions

1) <u>**Expand to external media & develop commercial ideas for boarder use**</u>
2) **Develop Customized promotion page with targeted app-push**

Ⅲ. 조사 결과 상세

1. 남성의 외모 관리 및 콘텐츠 이용 경험 | 4) 구독 중인 외모 관리 관련 인플루언서

구독하는 외모 관리 인플루언서: 디렉터 짱구대디 〉 디렉터 파이 〉 깡스타일리스트·씬님

- 전체 응답자 10명 중 1명(15.8%)은 남성 외모 관리 관련 인플루언서를 구독하고 있음.
- 구독 중인 남성 외모 관리 관련 인플루언서는 디렉터 짱구대디(10.1%), 디렉터 파이(6.3%), 깡스타일리스트·씬님(5.1%) 순 임.
- 구독하는 인플루언서는 대부분 남성으로, 주로 같은 성별의 인플루언서에게 외모 관리 관련 정보를 얻는 편임.

구독 중인 외모 관리 관련 인플루언서

[Base: 미응답자 제외, n=79, 복수, 주관식]
*없음, 모름 응답 제외

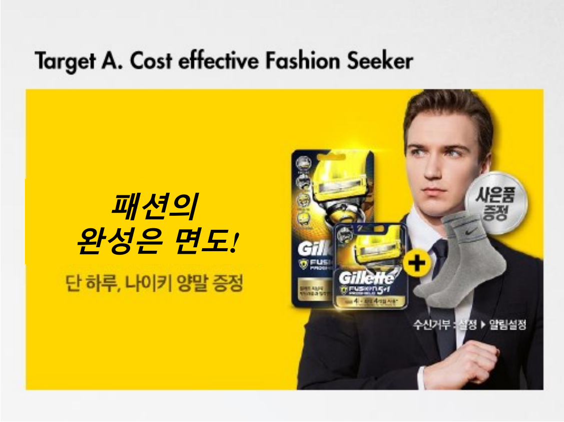
1위	2위	3위		5위		
디렉터 짱구대디	디렉터 파이	깡스타일리스트	씬님	금강연화	신쿡	스타일가이드 최겨울
패션	뷰티	패션	뷰티	헤어	뷰티	패션
						
10.1%	6.3%	5.1%	5.1%	3.8%	3.8%	3.8%

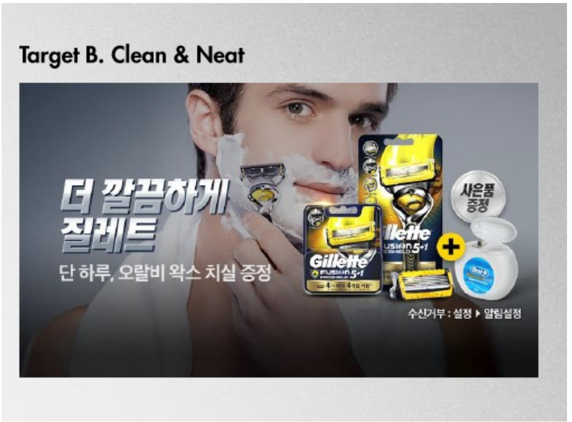
*그 외 : 김인호TV, 무신사TV, 아우라M, 이사배, 패션TV쩡대, 포맨트, 피부는민동성 등

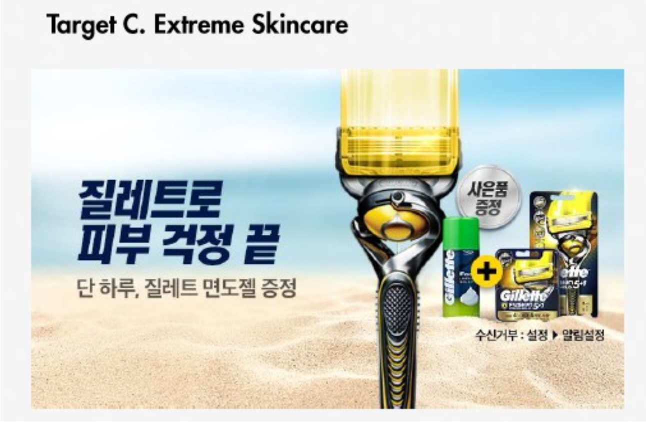
*이미지 출처: Youtube 공식계정

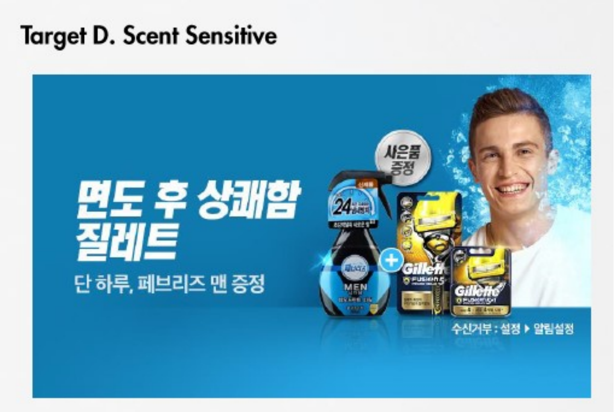
Potential Actions

1) Expand to external media & develop commercial ideas for boarder use
2) <u>Develop Customized promotion benefit with targeted app-push</u>







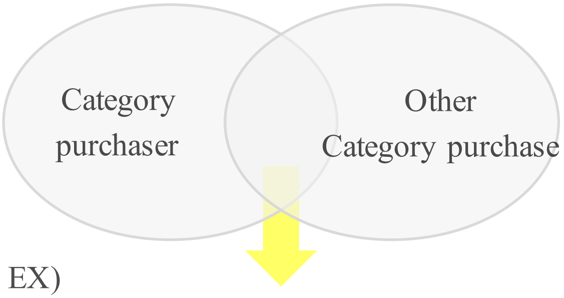


Project overview

Data analysis & targeting projects

Drive Category growth & recruit new users

Phase 1) Insight based shopper clustering



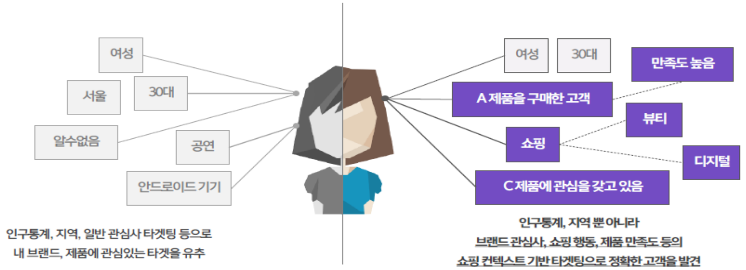
Understand the category growth shoppers characteristic

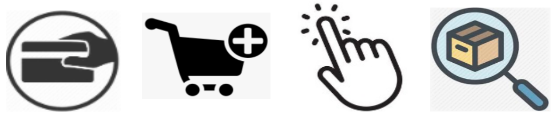
EX)

[Sports LOVER] [New hired] [Grooming]

Develop the promotion theme & target to wide rage of shoppers

Phase 2) Propensity modeling

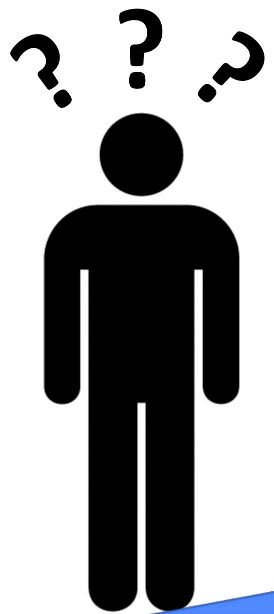




Sorting out shoppers who have high probability to become Gillette new users

Big Data – Propensity Modeling

How to identify “**WHO**”?



“**WHAT**” to sell?

…can answer the question!

P&G Knowledge Database

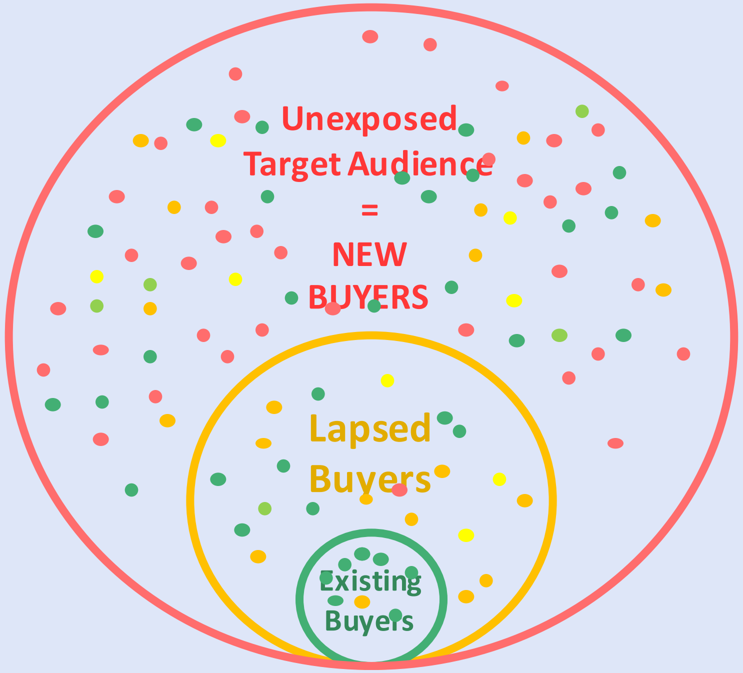
P&G Analytics Capability

eCommerce Data

Mock-up Example – Targeting

Propensity modeling starts with the all members in the specific eCommerce, the model results in “likelihood to ” by each member level.

The entire member pool



Existing Buyers: Fabric Enhancer purchasers in P6M
Lapsed Buyers: Fabric Enhancer non-purchasers in P6M
Unexposed: members who never purchased Fabric Enhancer in JBPs

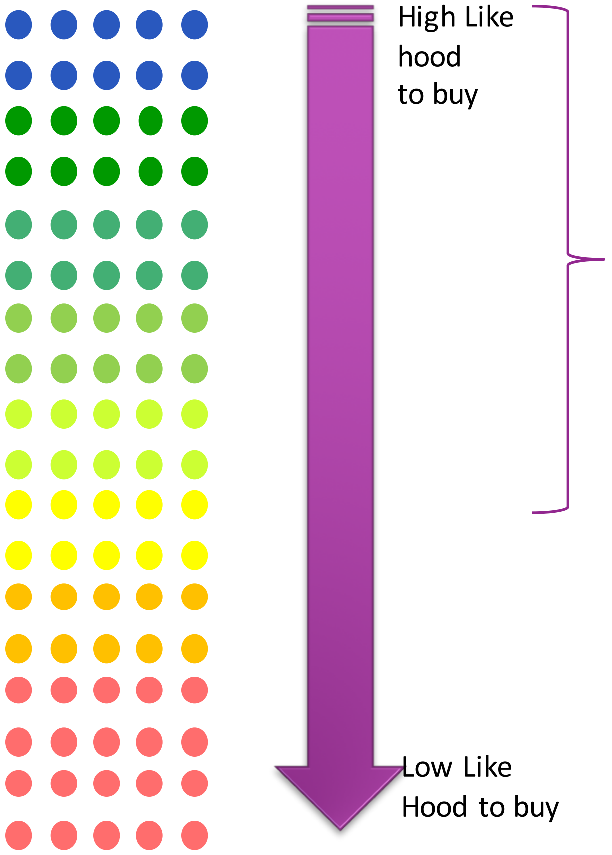
Propensity Modeling via Machine Learning

$$Y = Sigmoid(aX_1 + bX_2 + cX_3 + \cdots)$$

Propensity Score

Weight

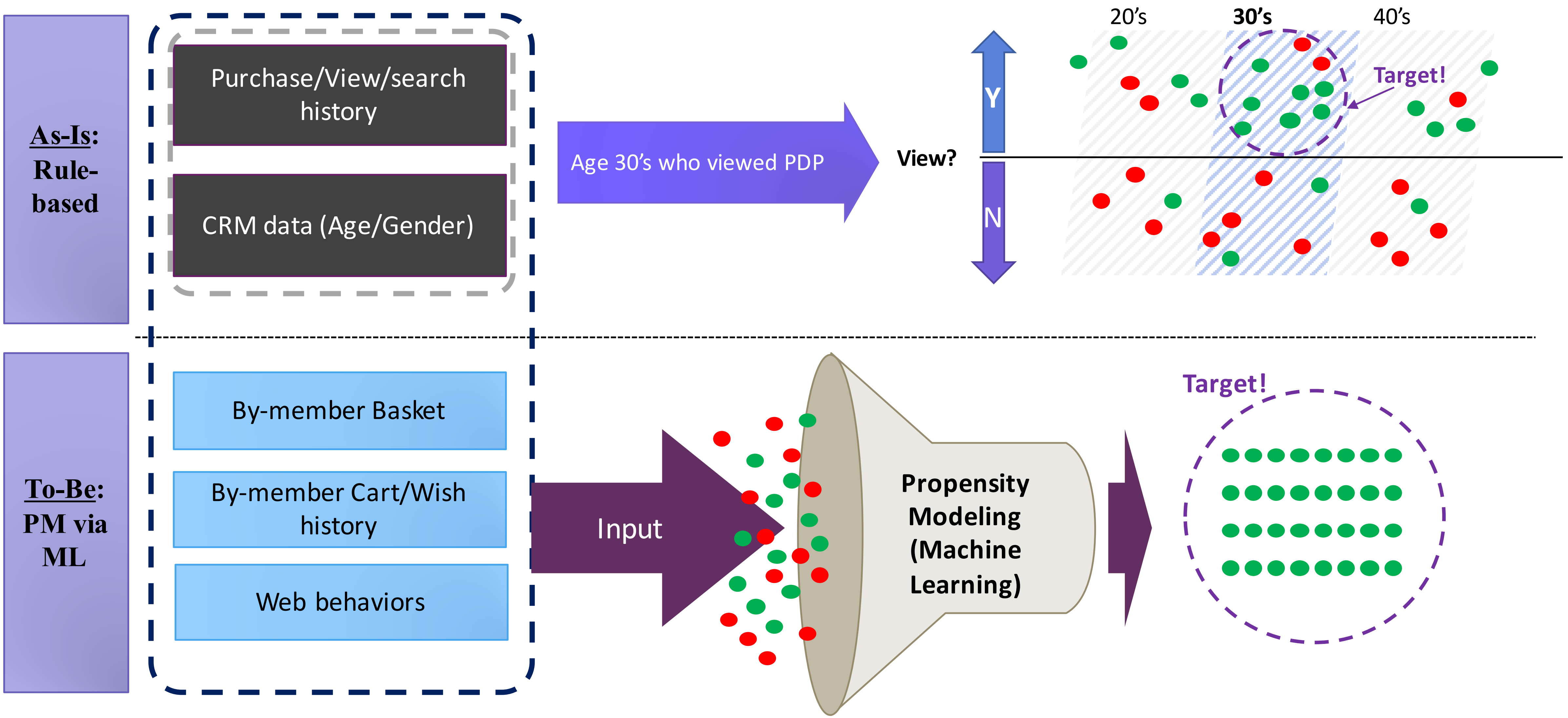
Predict **all consumers** propensity score to buy Category



Target Range: determined based on Available Funding

Mock-up Example – Targeting

Currently, the way to target is rule-based method but this method cannot capture all the “high-propensity” to buy FE category (green-colored) members. However, propensity modeling via Machine Learning can capture “green-colored” members.



Mock-up Example – Inputs for Propensity Modeling

Data should be wrangled distinct-by-member level; Model trains on positive & negative samples and learns patterns for users.

Features(Tags) including purchasing history/demographic/cart/wishlist/viewership

Target Variable

Positive samples; members purchased category in P6M

Negative samples; members not purchased category in P6M

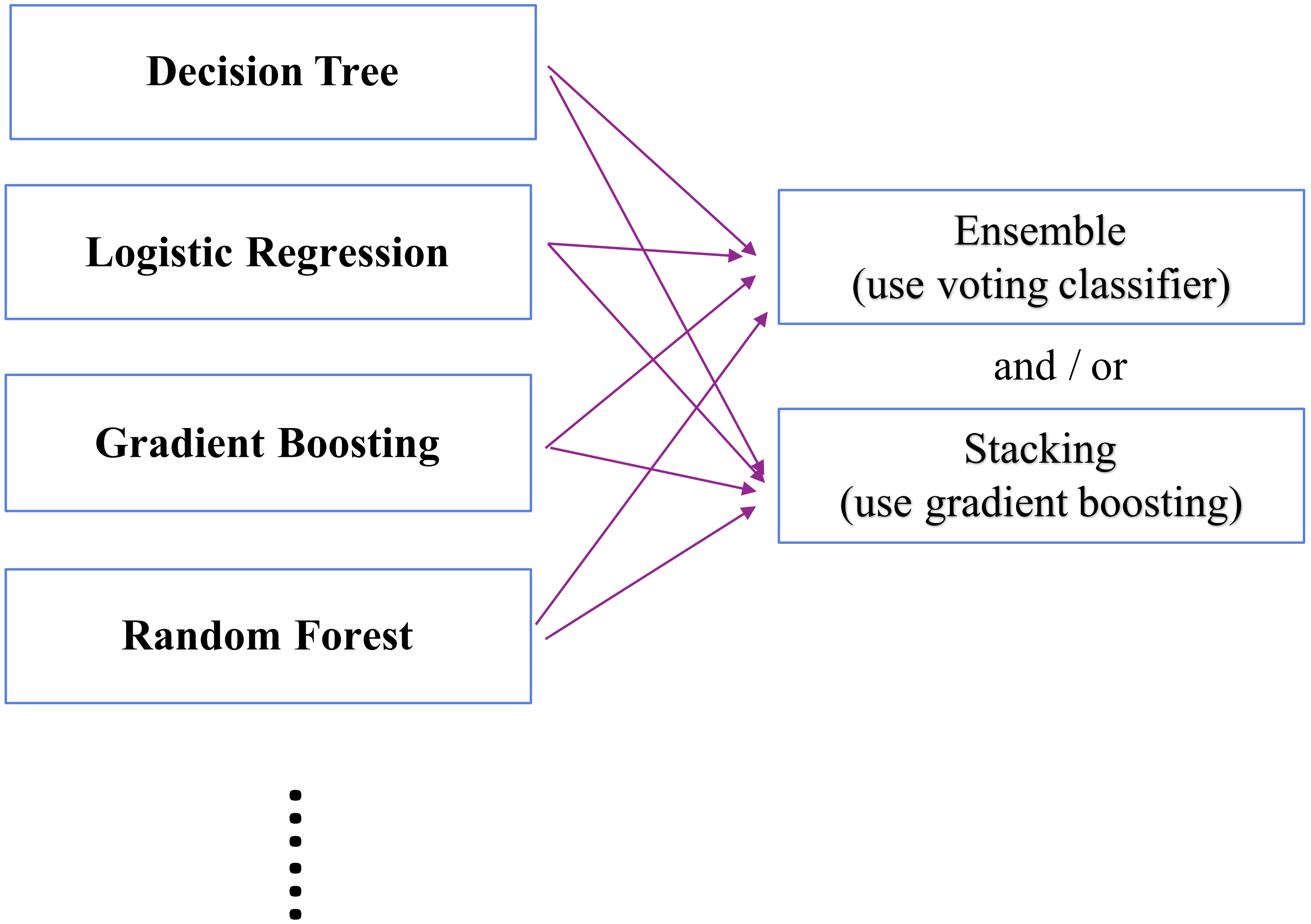
ID	Downy User?	Age Group	Gender	Purchased Fashion	Purchased Electronics	Viewed Kids	Viewed Beauty
A	1	1	0	13	0	5	20
B	1	2	1	59	5	0	7
C	1	3	0	0	0	55	0
D	1	1	0	2	0	3	0
E	0	2	1	0	9	21	8
F	0	3	0	1	1	4	22
G	0	4	0	0	0	2	36
H	0	6	1	0	10	0	6

……

⋮

Mock-up Example – Models for Propensity Modeling

There are so many algorithms to classify users. You can use only one algorithm, you can use ensemble or stacking with many algorithms



Use random forest algorithms as a Baseline model.

Use ensemble and stacking to beat baseline model.

Another choice can be user-based collaborative filtering.

Mock-up Example – Result for Propensity Modeling

How to evaluate Propensity Model

Confusion matrix

	P(predicted)	N(predicted)
P(Actual)	TP	FN
N(Actual)	FP	TN

Watch out

Our customers

No Interest

Potential new customers
They didn’t buy our product in this mall.
But model said they have very similar shopping interest to our existing customers → Persuadables

Mock-up Example – Result for Propensity Modeling

How to evaluate Propensity Model

Accuracy = TP+TN/TP+TN+FP+FN

Precision = TP / TP+FP

Recall = TP/TP+FN

Accuracy is not meaningful, most of case our data set is imbalanced.(Non Downy user >>>> Downy user)

Focus on the recall score. High Precision is not useless, our goal is find new users(Persuadables – FP) that has similar behavior and interest to our existing users

Make to higher recall, you have to adjust threshold.
Like
y_test_pred_prob[:,1] > 0.5
y_test_pred_prob[:,1] > 0.45
y_test_pred_prob[:,1] > 0.40 …


Thank you