

Personalized Marketing

Optimizing *Who, How, and When* to Market Any Product at Target



MBA Students: Charles Haocheng Bi (hbi@mit.edu), Raphael Chew (raphael1@mit.edu)
Unilever Sponsors: Matt Algar, Kevin Casey
MIT Advisor: Amr Farahat



IMPACT

An estimated **5M+ dollar uplift** for Dove Deodorant annual sales through one-on-one **personalized marketing** at Target stores

An ML framework that is **easily scalable** to over **5000 Unilever** products at Target and other retailers

A long-term opportunity to identify **category growth**, **personalize shopper experience**, and drive product **household penetration**

PROBLEM STATEMENT

For each Unilever marketing campaign,

- Who are the best **shoppers** to market a particular Unilever product to?
- How to personalize **marketing content** based on shopper segments?
- When is the best **time** to engage shoppers for a refill?

DATASET

- Demographics (26):** Location, Gender, Income, Household Info
- Item Feature (66):** Price, Category, Time Introduced, Brand, Size
- Transaction History (35):** Date, Promotion Used, Item Purchased
- Features Engineered (61):** RFMV, Re-order Characteristics, Segments

TIMELINE

Business Scoping & Exploratory Data Analysis

Product Affinity Scoring Model

Automated Shopper Segmentation

Re-purchase Prediction Model

Dashboarding & Presentations

February March April May June July August

PRODUCT AFFINITY SCORING

Key Problem: can we prescribe a product affinity score for each shopper given any Unilever product at Target, based on shopper demographics, transaction history, and item features?

Key Modeling Requirements

- Interpretability:** model insights will need to be **explainable** and **transparent** to inform future marketing efforts
- Real-life Implication:** incorporate **false positive cost** and **false negative cost** into the binary classification problem
- Ranked-ordered Scoring:** **optimize marketing ROI** under restrained marketing budgets

Approach: ensemble of 6 gradient boosted trees (GBT) of different hyperparameters

Data



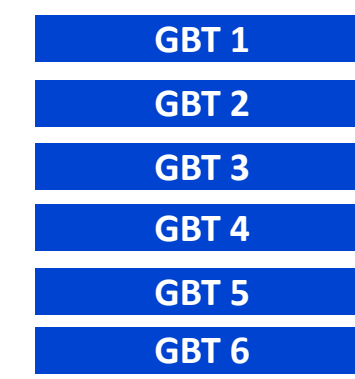
Model Selection

GBT has superior **predictive power**, **balanced class predictions**, and **interpretability** using Shapley values.

Model	AUC
Logistic Regression	0.73
Decision Tree Classifier	0.69
Random Forest	0.72
Gradient Boosted Tree (GBT)	0.79
Multi-layer Perceptron (MLP)	0.77

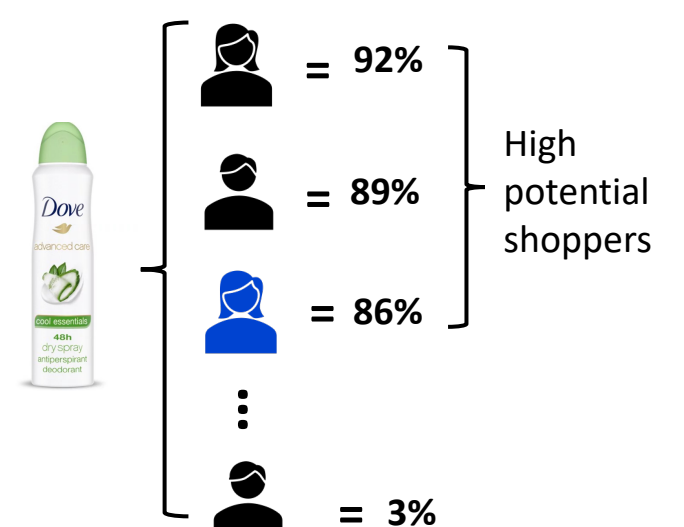
GBT Ensemble

Each GBT has a slightly different set of **hyperparameters and seed**. Shoppers' final affinity score is an **unweighted average** of the 6 predictions.



Affinity Scoring

The team considers shoppers with an affinity score **above 50%** to be high-potential shoppers.



AUTOMATIC SEGMENTATION

Key Problem: create marketing content that resonate the most
Approach: item description parsing + spend analysis

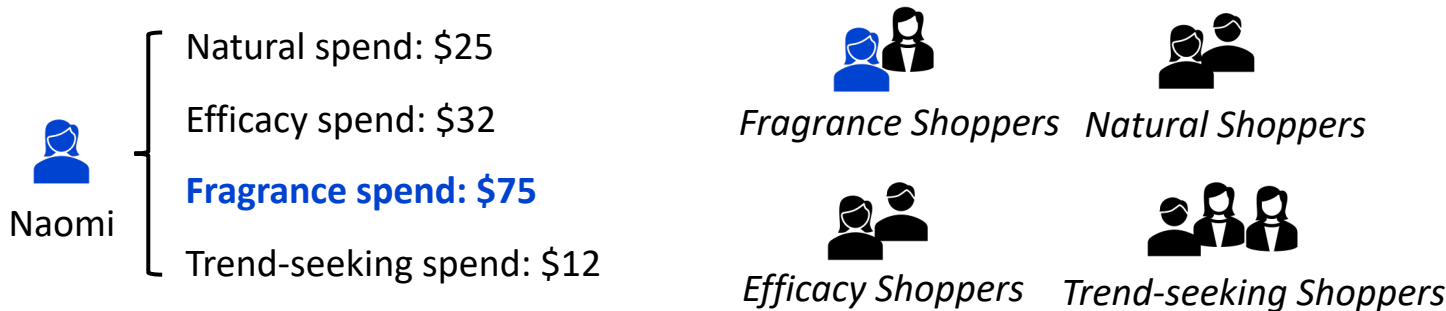
1 Tag Items

Item Description: "cucumber 48h cool aluminum-free antiperspirant"
Segments: fragrance, efficacy, skin-feel, natural



2 Segment Shoppers

Approach: track segment purchase and classify shoppers into the segment of highest normalized spend in order to personalize content



RE-PURCHASE PREDICTION

Key Problem: optimize marketing cycle by predicting refills and pinpointing high demand
Key Modeling Requirement: time-series feature engineering to avoid data leakage

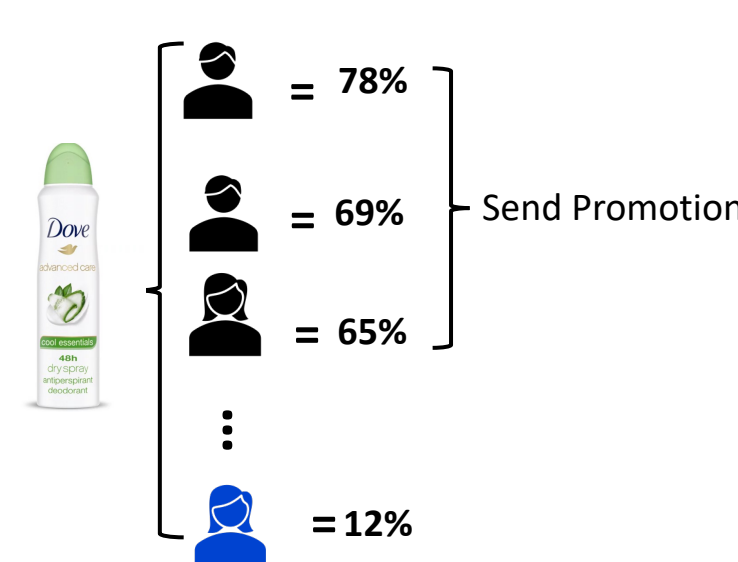
Model Selection

GBT has superior predictive power, balanced class predictions while taking only **20% training time** of MLP's.

Model	AUC
Logistic Regression	0.79
Decision Tree Classifier	0.61
Random Forest	0.82
Gradient Boosted Tree (GBT)	0.87
Multi-layer Perceptron (MLP)	0.86

Refill Needs

Refill needs depend heavily on shoppers' **purchasing habits** and **recent orders**.
 Optimized marketing time = best marketing ROI



Optimize Marketing Cycle

If the budget can reach 100k shoppers, then we take the top 100k **most ready shoppers** and **average** re-purchase likelihood to **pinpoint the high-demand** week to market.



DELIVERABLES

Power BI Dashboard



10+ ML Models



3 Presentations

Audience include 3 VPs across 4 teams



RESULTS

\$5M+

Annual POS Uplift for Dove Deo*

ML Pipeline of **5K+** Unilever Products at Target

1

upcoming campaign integration in Oct. 2021

* Estimated based on 2019 data

NEXT STEPS

Short-term

- Validate model results
- Identify growth drivers
- Unify cross-retailer data
- Add price sensitivity inputs

Long-term

- Complete an eco-system of "who", "how", "when", "how much", and "where"
- Build experimentation capability to test live strategies