# A Scalable Dynamic Bayesian Mixture Model for Fine-grained Promotion Mix Analysis of Digital Coupons

Gourab Mukherjee

University of Southern California

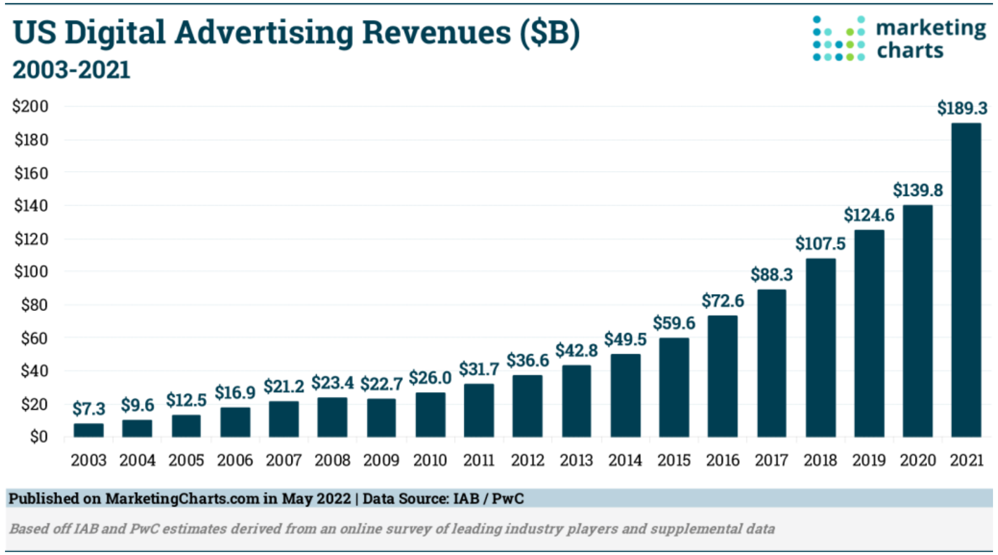
2022 IISA Annual Conference

Bengaluru

28th December 2022

Joint work with: Rashmi Ranjan Bhuyan (USC), Wreetabrata Kar (Purdue)

# Digital Marketing





- Digital promotions have revolutionized the way companies target customers with marketing offers.
- In 2021, an average firm was expected to allocate 45% of their total marketing budget to online media.
- Worldwide spending on digital advertising is projected to be 566 billion U.S. dollars in 2022.



# Digital Coupons



- Digital coupons offer discounts and rebates in electronic form through the use of coupon codes, vouchers, and other discounting methods.
- They are extensively used by marketers to attract and retain customers by promoting the desire to save money while making purchases.

- Digital coupons provide several benefits over traditional coupons.
  - Customers can easily store & search and so, they have much higher redemption rates than traditional coupons
  - Marketing managers benefit from very low designing and dissemination costs.

- Using digital coupons managers can very cheaply
  - assimilate different types of promotions (marketing interventions) in a single coupon
  - customize these coupons for different promotional needs.

# Digital Couponing: Promotion Mix Challenges

Digital coupons are easier to make and managers can put in a lot of choices simultaneously

**↔ Most coupons now-a-days are complex and multi-dimensional marketing inventions**


involves managerial decision on three attributes:
(a) kind of discount, which in this case is price,
(b) a product category, which is shoes, and
(c) the discount percentage level of 20%.

- It is extremely important for marketers to accurately understand how mixing these marketing components at different levels impact customers' responses to coupons.
- The roles of different promotion components in influencing consumer psychology have been separately studied in the existing literature.
- However, simultaneous estimation of the effects of these components in coupon marketing is challenging.
- Here, we provide a disciplined statistical methodology for conducting promo mix analysis by efficient estimating and attributing the impact of different marketing components used in coupon marketing.

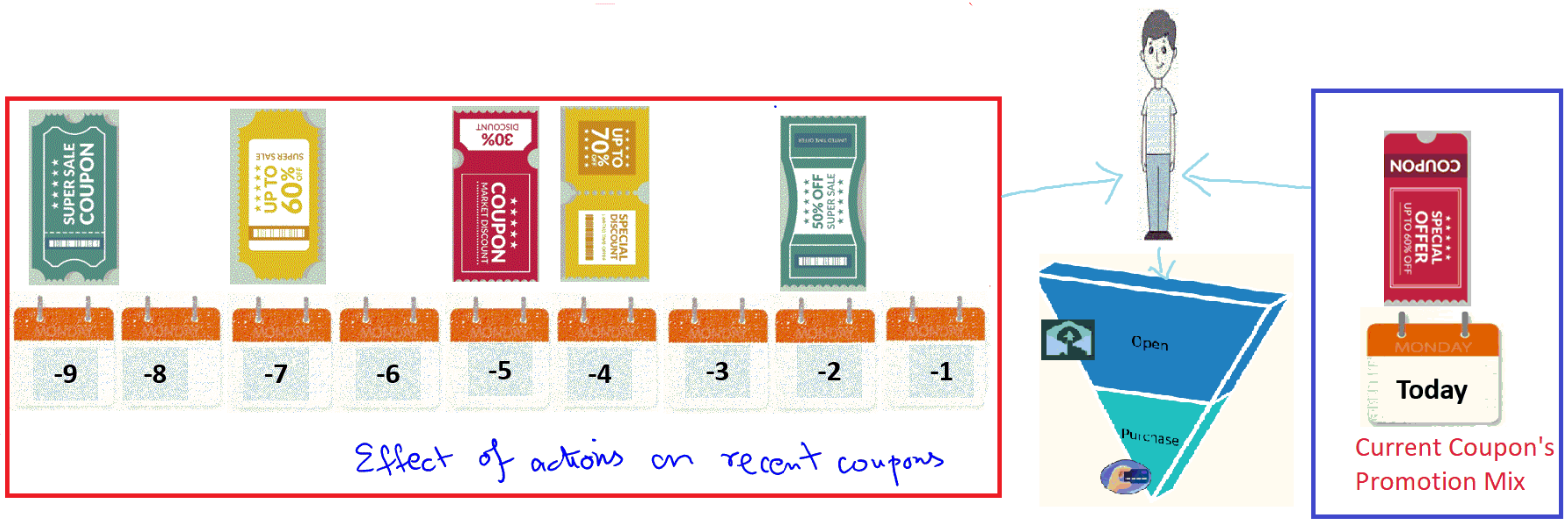
# Statistical Challenges in Promotion Mix analysis of Digital Coupons

Wide heterogeneity in the engagement levels and buying propensity among the customer base.

- It is important to group customers based on similarity of these attributes and then compare the effect of the promotions within segments.
- We use Bayesian Mixture model: GLMM model based clustering framework for grouping customers' based on their long-term behaviors.

Promotions received by customers in the recent past greatly affects their responses to the current promotion

- We use a stock variable that dynamically measures the carryover effects of coupons that a customer was exposed to in the recent past.
- We develop a 2-stage Dynamic Mixture Model (2DMM) and their coefficients are estimated by a block Markov Chain Monte Carlo (MCMC) algorithm on the combined two-stage likelihood.

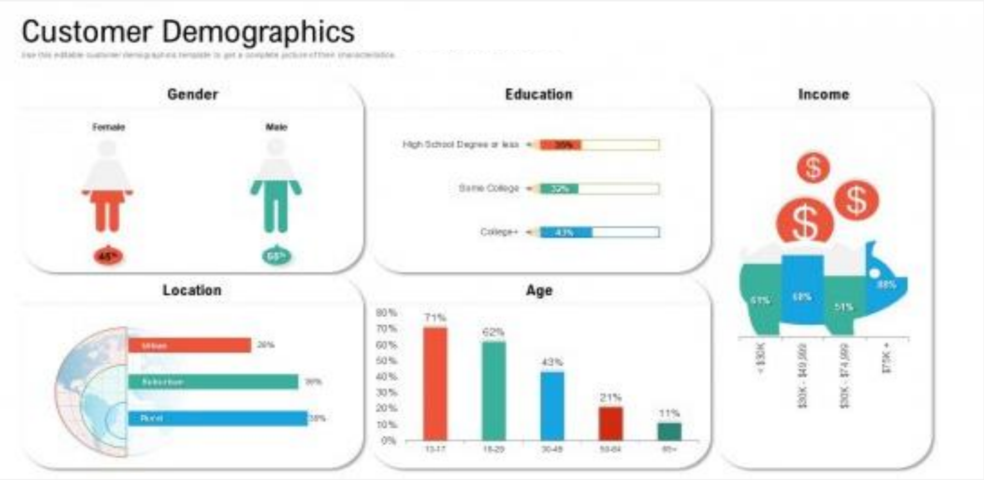


# Data

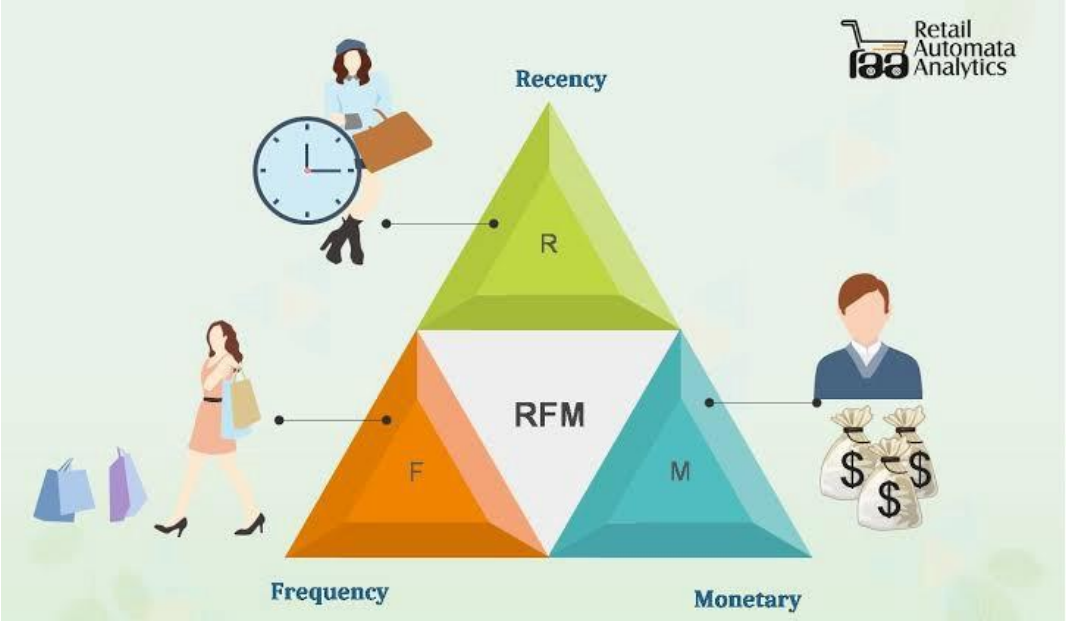
Long Term Association Data with the Firm: Summary of association over two years → Used in Stage I to segment customers

Current Marketing Campaign: 45 days window and detailed responses to promotions → Used in Stage II to estimate promotion efficiency on different customer segments

- Around 80k active customers
- Each customer’s demography data:
    - Gender
    - Income
    - Age



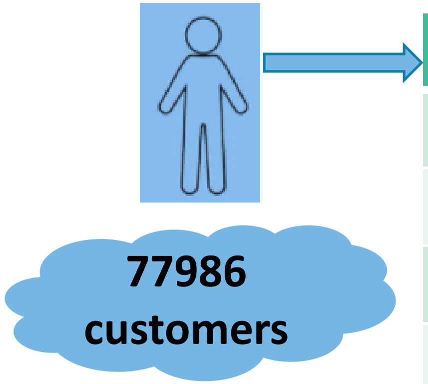
# Data: Stage I


- Long term association with the firm with historical data

- Recency: corresponds to days since last
  - Open
  - Click
  - Purchase
- Monetary: total purchase amount
- Frequency: number of purchases

# Data: Stage II




| Promotion | Open | Click | Purchase |
|---|---|---|---|
| P1 | Yes | Yes | No |
| P2 | No | - | - |
| . | . | . | . |
| P25 | Yes | Yes | Yes |

- Marketing window of 45 days
- Each customer receives around 25 promotional emails
- Each customer has a corresponding table


Each coupon has one or combination of marketing components (bag of features).

Examples of promotions:

Use promo code AB123. 50% off Swim + Free Ship & Returns!

Use promo code CD456. Mystery Savings - up to 30% off!

Use promo code EF789. Extra $10 off, your instant refund!

Table corresponding to the example

| Promotion | Percentage | Shipping | Returns | Mystery | Dollar |
|---|---|---|---|---|---|
| P1 | 1 | 1 | 1 | 0 | 0 |
| P2 | 1 | 0 | 0 | 1 | 0 |
| P3 | 0 | 0 | 0 | 0 | 1 |

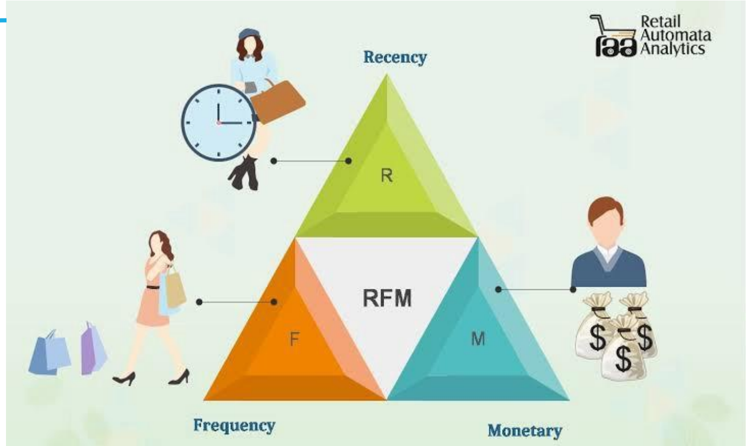
# Model

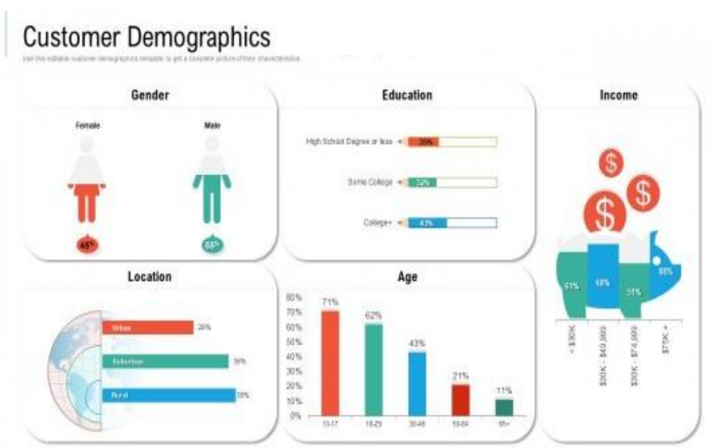
- Two problems at hand:
  - Create homogenous groups of customers (based on previous data)
  - Targeted coupons for each customer group using marketing components.

- 2- stage Dynamic Mixture Model (2DMM)
  - Stage 1: Customer Segmentation
  - Stage 2: Marketing components effects with dynamic short-term engagement

# Stage 1: Segmenting Customers

- Very high heterogeneity among the consumer base
- Estimating of coupon effects assuming invariance across customers will be misleading.
- Model based clustering with long-term association and demographic data




- Model as joint GLM's
- Two sets of demographic variables:
  - $w_i$ - global effects
  - $\widetilde{w_i}$ - cluster specific heterogenous effects

# Stage 1: Segmenting Customers

- Model as joint GLM's
- Two sets of demographic variables:
  - $w_i$ - global effects
  - $\widetilde{w_i}$ - cluster specific heterogenous effects
- h(.), clustering function
  - i.e h(32) = 5, means the $32^{nd}$ customer is in $5^{th}$ cluster

$$\mathrm{R_i^o} = \mathrm{w_i^T}\delta^o + \widetilde{w_i^T}\tau_{h(i)}^o + \epsilon_i^o$$

$$\mathrm{R_i^c} = \mathrm{w_i^T}\delta^c + \widetilde{w_i^T}\tau_{h(i)}^c + \epsilon_i^c$$

$$\mathrm{R_i^p} = \mathrm{w_i^T}\delta^p + \widetilde{w_i^T}\tau_{h(i)}^p + \epsilon_i^p$$

$$\log \mathrm{M_i} = \mathrm{w_i^T}\delta^M + \widetilde{w_i^T}\tau_{h(i)}^M + \epsilon_i^M$$

$$\mu(F_i) = \mathrm{w_i^T}\delta^F + \widetilde{w_i^T}\tau_{h(i)}^F$$

- Each recency variable (open/click/purchase) modelled linearly
- Monetary variable has a log-normal distribution
- Frequency variable is Poisson with mean as given

# Stage 1: Segmenting Customers

$$R_i^o = w_i^T\delta^o + \widetilde{w_i^T}\tau_{h(i)}^o + \epsilon_i^o$$

$$R_i^c = w_i^T\delta^c + \widetilde{w_i^T}\tau_{h(i)}^c + \epsilon_i^c$$

$$R_i^p = w_i^T\delta^p + \widetilde{w_i^T}\tau_{h(i)}^p + \epsilon_i^p$$

$$\log M_i = w_i^T\delta^M + \widetilde{w_i^T}\tau_{h(i)}^M + \epsilon_i^M$$

$$\mu(F_i) = w_i^T\delta^F + \widetilde{w_i^T}\tau_{h(i)}^F$$

- Main Goal: Estimate the clustering function
- We use a Bayesian approach, assume prior on the coefficients (gaussian priors)

2 step MCMC :

1. Fix h(.) update the other parameters ($\delta^o$, $\delta^c$, $\delta^p$, …) with Metropolis-Hastings
2. For each customer 'i' update h(i) to the cluster with max log-likelihood

# Stage 2: Marketing Component effects

- Promotion effect for each segment of customer, sum of
  - marketing components effect and
  - residual effect (sparse)

$$\mu_{k,h} = \sum_{1}^{J} \theta_{k,j} \nu_{j,h} + \gamma_{k,h}$$

- With the marketing components effect, managers can create new optimal coupons for each segment

Example of promotions:

Use promo code AB123. 50% off Swim + Free Ship & Returns!

Use promo code CD456. Mystery Savings - up to 30% off!

Use promo code EF789. Extra $10 off, your instant refund!

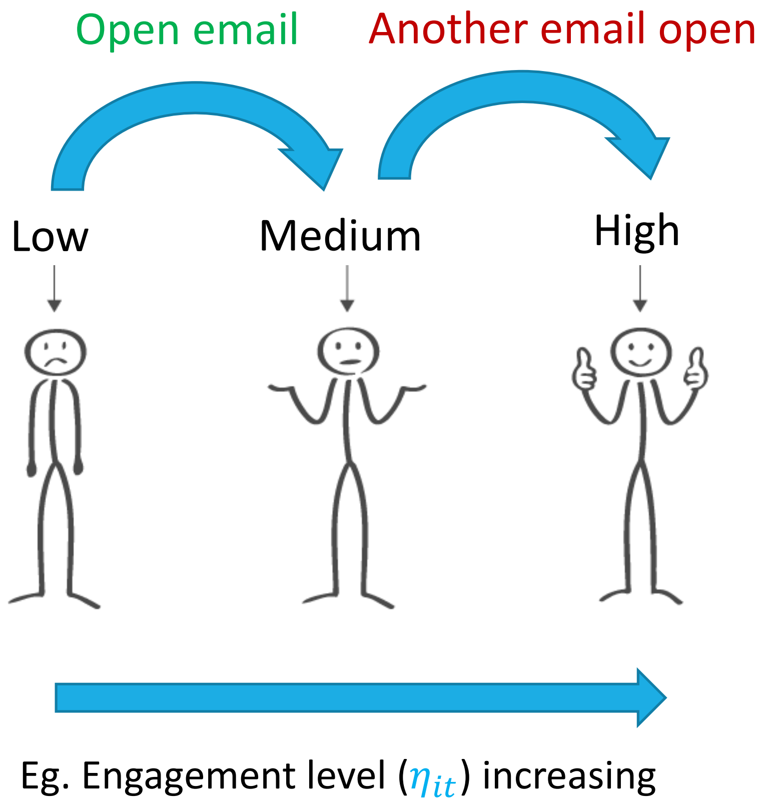
| Promotion | Percentage | Shipping | Returns | Mystery | Dollar |
|---|---|---|---|---|---|
| P1 | 1 | 1 | 1 | 0 | 0 |
| P2 | 1 | 0 | 0 | 1 | 0 |
| P3 | 0 | 0 | 0 | 0 | 1 |

# Stage 2: Short Term Dynamic Engagement

- The promotional coupons are sequential, introduce a dynamic engagement level '$\eta_{it}$'

$$\eta_{i,t} = \mathbf{1}\{Y_{i,t-1} = 1\} + \delta\eta_{i,t-1}$$

- The stock variable level considers the recent engagement as well as the past engagement.



# Stage 2: Logistic Model for Promotion Efficacy

$$\text{logit}(\boldsymbol{P}(Y_{it} = 1)) = x_{ik}^{T}\alpha + \widetilde{x_{ik}^{T}}\beta_h + \eta_{it} + \underbrace{\sum_1^J \theta_{k,j}\nu_{j,h} + \gamma_{k,h}}_{} + a_h$$

**Indices**

- Promotion k
- Customer i
- Cluster h
- Time t

*With cluster indices h fixed

$\alpha$ , $\beta_h$ captures the customer invariant, varying across segments covariates

$$\eta_{i,t} = \mathbf{1}\{Y_{i,t} = 1\} + \delta\eta_{i,t-1}$$

$$\mu_{k,h} = \sum_1^J \theta_{k,j}\nu_{j,h} + \gamma_{k,h}$$

$a_h$ random effect capturing covariance across clusters

**Block MCMC:**

1. Update all parameters except ‘$\delta$’, use slightly modified Polya-Gamma augmentation for logistic regression
2. Update ‘$\delta$’ by perturbing, update $\eta_{it}$ correspondingly, compute likelihood and accept/reject (MH)

# Path algorithm for combining the two stages

- We develop a path algorithm that provides path of estimates of the parameters in 2DMM as the number of clusters (m) varies over 1,...,M for some large M.

- Our proposed path algorithm merges clusters in stage I of 2DMM in a bottom up fashion based on minimizing the squared error loss. We provide fast asymptotically efficient updates of the parameter estimates in stage II of 2DMM as clusters are merged.

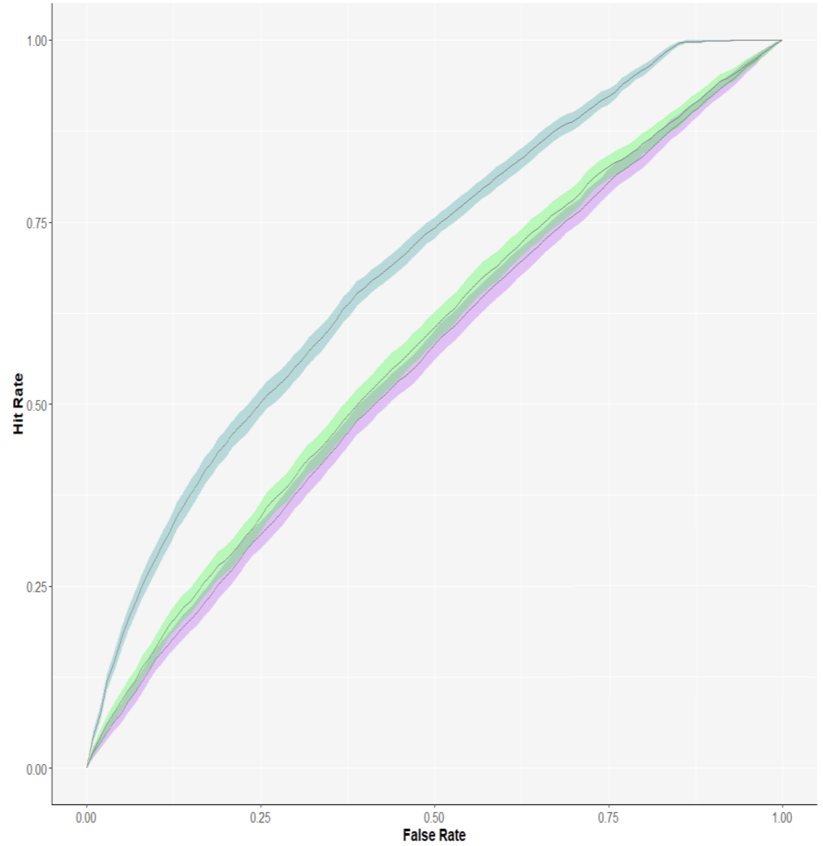
- Our proposed novel path algorithm for 2DMM provides integrative analysis of the promotion mix effects at different desired heterogeneity levels.

# Results

- Held out promotional period data of 19496 customers
- Ran three models
  - 2DMM (with 5 clusters)
  - Naive logistic model: $logit(\boldsymbol{P}(Y_{it} = 1)) = x_{ik}^{T}\theta + \gamma_k$
  - Dynamic logistic model: $logit(\boldsymbol{P}(Y_{it} = 1)) = x_{ik}^{T}\theta + \eta_{i,t} + \gamma_k$

- 2DMM provided a better AUC (0.71) compared to naive logistic model (0.59) and dynamic logistic regression's AUC (0.62). **20.3% improvement ☺**

- **Summary:** Combined with the dynamic stock model, segmenting customers based on past behavior can significantly improve coupon effect estimation.





Reference: Mukhopadhay S, Kar W, Mukherjee G. *Estimating Promotion Effectiveness in Email Marketing: A high-dimensional Bayesian Joint Model for Nested Imbalanced Data*, Annals of Applied Statistics, 2022.


Thank You ☺