SHOPPER: A PROBABILISTIC MODEL OF CONSUMER CHOICE WITH SUBSTITUTES AND COMPLEMENTS

Francisco J. R. Ruiz, David Blei and Susan Athey · 2020

By Soel Micheletti · Advanced Topics in Machine Learning and Data Science · May 2021

ETH zürich

.

SHOPPER: A PROBABILISTIC MODEL OF CONSUMER CHOICE WITH SUBSTITUTES AND COMPLEMENTS

Francisco J. R. Ruiz et al. · 2020

Source: https://produkte.migros.ch

By Soel Micheletti · May 2021 · ETH zürich

SHOPPER: A PROBABILISTIC MODEL OF CONSUMER CHOICE WITH SUBSTITUTES AND COMPLEMENTS

Francisco J. R. Ruiz et al. · 2020

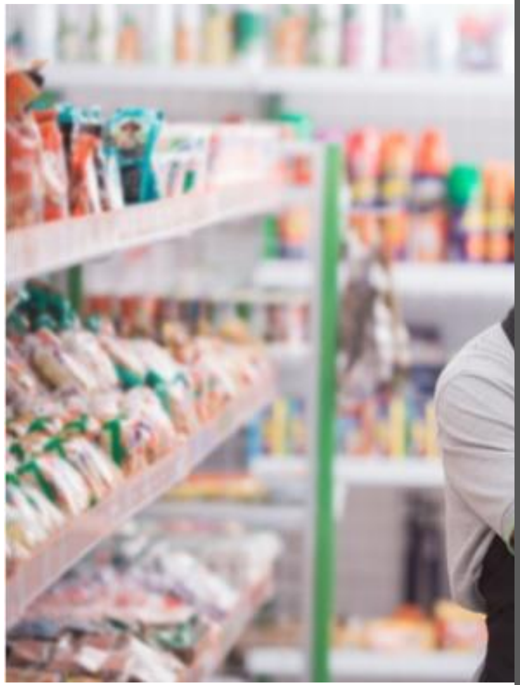
Source: https://produkte.migros.ch

By Soel Micheletti · May 2021 · ETH zürich

Source: https://cb4.com

Source: https://cb4.com

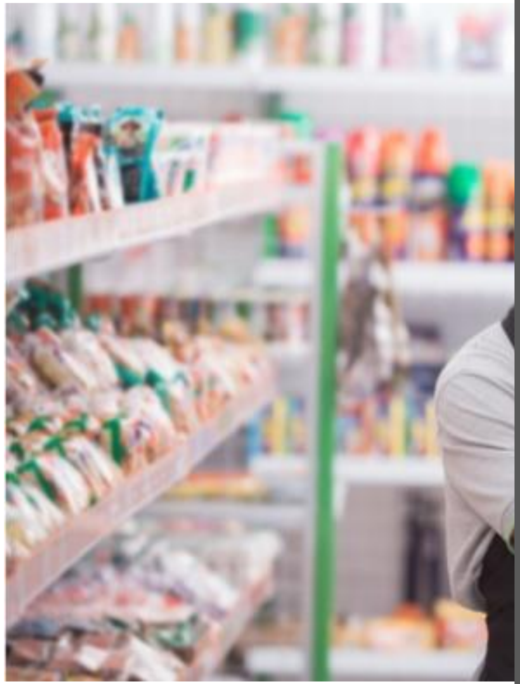
Source; https://atlas-network.com



Source: https://cb4.com

- What products should be promoted? By how much?
- Should the price of certain products be changed?
- Insights about both products and customers
- Form personalized recommendations

Source; https://atlas-network.com



Source: https://cb4.com

"MAKE ACCURATE PREDICTIONS UNDER PRICE INTERVENTIONS"

Francisco J. Ruiz

Source; https://atlas-network.com

Plan

THE MODEL

OPTIMIZATION/ INFERENCE

Plan

1. The Model

2. Optimization/ Inference

3. Results/ Discussion/ Openings

Plan

1. The Model

2. Optimization/ Inference

3. Results/ Discussion/ Openings

The Model

Intuition · Mathematical Formulation · Extensions

ITEM 1

ITEM 2

ITEM 3

ITEM 4

CHECK-OUT

The Model

Intuition · Mathematical Formulation · Extensions

ITEM 1	ITEM 2	ITEM 3	ITEM 4	CHECK-OUT
U_1	U_2	U_3	U_4	U_C

The Model

Intuition · Mathematical Formulation · Extensions

ITEM 1	ITEM 2	ITEM 3	ITEM 4	CHECK-OUT
U_1	U_2	U_3	U_4	U_C

- Price (sensitivity)
- Customer's preferences
- Item popularity
- Season of the year

...

The Model

Intuition · Mathematical Formulation · Extensions

ITEM 1	ITEM 2	ITEM 3	ITEM 4	CHECK-OUT
U_1	U_2	U_3	U_4	U_C

The Model

Intuition · Mathematical Formulation · Extensions

ITEM 1

$\tilde{U}_1$

ITEM 2

$\tilde{U}_2$

ITEM 3

$\tilde{U}_3$

ITEM 4

$\tilde{U}_4$

CHECK-OUT

$\tilde{U}_C$

The Model

Intuition · Mathematical Formulation · Extensions

ITEM 1	ITEM 2	ITEM 3	ITEM 4	CHECK-OUT
$\tilde{U}_1$	$\tilde{U}_2$	$\tilde{U}_3$	$\tilde{U}_4$	$\tilde{U}_C$

The Model

Intuition · Mathematical Formulation · Extensions

ITEM 1	ITEM 2	ITEM 3	ITEM 4	CHECK-OUT
U^*_1	U^*_2	U^*_3	U^*_4	U^*_C

The Model

Intuition · Mathematical Formulation · Extensions

ITEM 1	ITEM 2	ITEM 3	ITEM 4	CHECK-OUT
U^*_1	U^*_2	U^*_3	U^*_4	U^*_C

The Model

Intuition · **Mathematical Formulation** · Extensions

$\mathbf{y}_t = (y_{t1}, y_{t2}, \ldots, y_{tn})$

$\mathbf{y}_{t,\, i-1} = (y_{t1}, y_{t2}, \ldots, y_{t(i-1)})$

The Model

Intuition · **Mathematical Formulation** · Extensions

$\mathbf{y}_t = (y_{t1}, y_{t2}, \ldots, y_{tn})$

$\mathbf{y}_{t,\, i\text{-}1} = (y_{t1}, y_{t2}, \ldots, y_{t(i\text{-}1)})$

$$\Pr[y_{ti} = c \mid \mathbf{y}_{t,(i-1)}] = \frac{\exp\{\Psi(c, \boldsymbol{y}_{t,i-1})\}}{\sum_{c' \neq \boldsymbol{y}_{t,i-1}} \exp\{\Psi(c', \boldsymbol{y}_{t,i-1})\}}$$

The Model

Intuition · **Mathematical Formulation** · Extensions

$\mathbf{y}_t = (y_{t1}, y_{t2}, \ldots, y_{tn})$

$\mathbf{y}_{t,\, i\text{-}1} = (y_{t1}, y_{t2}, \ldots, y_{t(i\text{-}1)})$

$$\Pr[y_{ti} = c \mid \mathbf{y}_{t,(i-1)}] = \frac{\exp\{\Psi(c, \boldsymbol{y}_{t,i-1})\}}{\sum_{c' \neq \boldsymbol{y}_{t,i-1}} \exp\{\Psi(c', \boldsymbol{y}_{t,i-1})\}}$$

The Model

Intuition · **Mathematical Formulation** · Extensions

$$\Psi\left(c, \boldsymbol{y}_{t,i-1}\right) = \Psi_{tc} + \rho_c^{\mathrm{T}}\left(\frac{1}{\mathrm{i}-1}\sum_{j=1}^{i-1}\alpha_{y_{tj}}\right)$$

$\mathbf{y}_t = (y_{t1}, y_{t2}, \ldots, y_{tn})$

$\mathbf{y}_{t,\, i\text{-}1} = (y_{t1}, y_{t2}, \ldots, y_{t(i\text{-}1)})$

$$\Pr[y_{ti} = c \mid \mathbf{y}_{t,(i-1)}] = \frac{\exp\{\Psi\left(c, \boldsymbol{y}_{t,i-1}\right)\}}{\sum_{c' \neq \boldsymbol{y}_{t,i-1}} \exp\{\Psi\left(c', \boldsymbol{y}_{t,i-1}\right)\}}$$

The Model

Intuition · **Mathematical Formulation** · Extensions

$\mathbf{y}_t = (y_{t1}, y_{t2}, \ldots, y_{tn})$

$\mathbf{y}_{t,\,i\text{-}1} = (y_{t1}, y_{t2}, \ldots, y_{t(i\text{-}1)})$

$$\Pr[y_{ti} = c \mid \mathbf{y}_{t,(i-1)}] = \frac{\exp\{\Psi(c, \boldsymbol{y}_{t,i-1})\}}{\sum_{c' \neq \boldsymbol{y}_{t,i-1}} \exp\{\Psi(c', \boldsymbol{y}_{t,i-1})\}}$$

$$\Psi(c, \boldsymbol{y}_{t,i-1}) = \Psi_{tc} + \rho_c^{\mathrm{T}} \left(\frac{1}{\mathrm{i}-1} \sum_{j=1}^{i-1} \alpha_{y_{tj}}\right)$$

Customer's preferences

The Model

Intuition · **Mathematical Formulation** · Extensions

$$\Psi\left(c, \boldsymbol{y}_{t,i-1}\right) = \Psi_{tc} + \rho_c^{\mathrm{T}}\left(\frac{1}{\mathrm{i}-1}\sum_{j=1}^{i-1}\alpha_{y_{tj}}\right)$$

Interaction's coefficient

$\mathbf{y}_t = (y_{t1}, y_{t2}, \ldots, y_{tn})$

$\mathbf{y}_{t,\, i\text{-}1} = (y_{t1}, y_{t2}, \ldots, y_{t(i\text{-}1)})$

$$\Pr[y_{ti} = c \mid \mathbf{y}_{t,(i-1)}] = \frac{\exp\{\Psi\left(c, \boldsymbol{y}_{t,i-1}\right)\}}{\sum_{c' \neq \boldsymbol{y}_{t,i-1}} \exp\{\Psi\left(c', \boldsymbol{y}_{t,i-1}\right)\}}$$

The Model

Intuition · **Mathematical Formulation** · Extensions

$\mathbf{y}_t = (y_{t1}, y_{t2}, \dots, y_{tn})$

$\mathbf{y}_{t,\, i\text{-}1} = (y_{t1}, y_{t2}, \dots, y_{t(i\text{-}1)})$

$$\Pr[y_{ti} = c \mid \mathbf{y}_{t,(i-1)}] = \frac{\exp\{\Psi(c, \boldsymbol{y}_{t,i-1})\}}{\sum_{c' \neq \boldsymbol{y}_{t,i-1}} \exp\{\Psi(c', \boldsymbol{y}_{t,i-1})\}}$$

$$\Psi(c, \boldsymbol{y}_{t,i-1}) = \Psi_{tc} + \rho_c^{\mathrm{T}} \Big(\frac{1}{\mathrm{i} - 1} \sum_{j=1}^{i-1} \alpha_{y_{tj}} \Big)$$

Latent representation of the item

The Model

Intuition · **Mathematical Formulation** · Extensions

$\mathbf{y}_t = (y_{t1}, y_{t2}, \ldots, y_{tn})$

$\mathbf{y}_{t,\, i\text{-}1} = (y_{t1}, y_{t2}, \ldots, y_{t(i\text{-}1)})$

$$\Pr[y_{ti} = c \mid \mathbf{y}_{t,(i-1)}] = \frac{\exp\{\Psi(c, \boldsymbol{y}_{t,i-1})\}}{\sum_{c' \neq \boldsymbol{y}_{t,i-1}} \exp\{\Psi(c', \boldsymbol{y}_{t,i-1})\}}$$

$$\Psi(c, \boldsymbol{y}_{t,i-1}) = \Psi_{tc} + \rho_c^{\mathrm{T}}\left(\frac{1}{\mathrm{i}-1}\sum_{j=1}^{i-1} \alpha_{y_{tj}}\right)$$

The Model

Intuition · **Mathematical Formulation** · Extensions

$$\Psi_{tc} = \lambda_c$$

Item popularity

$$\mathbf{y}_t = (y_{t1}, y_{t2}, \ldots, y_{tn})$$

$$\mathbf{y}_{t,\,i\text{-}1} = (y_{t1}, y_{t2}, \ldots, y_{t(i\text{-}1)})$$

$$\Pr[y_{ti} = c \mid \mathbf{y}_{t,(i-1)}] = \frac{\exp\{\Psi(c, \boldsymbol{y}_{t,i-1})\}}{\sum_{c' \neq \boldsymbol{y}_{t,i-1}} \exp\{\Psi(c', \boldsymbol{y}_{t,i-1})\}}$$

$$\Psi(c, \boldsymbol{y}_{t,i-1}) = \Psi_{tc} + \rho_c^{\mathrm{T}} \left(\frac{1}{\mathrm{i} - 1} \sum_{j=1}^{i-1} \alpha_{y_{tj}}\right)$$

The Model

Intuition · **Mathematical Formulation** · Extensions

$\mathbf{y}_t = (y_{t1}, y_{t2}, \ldots, y_{tn})$

$\mathbf{y}_{t,i-1} = (y_{t1}, y_{t2}, \ldots, y_{t(i-1)})$

$$\Pr[y_{ti} = c \mid \mathbf{y}_{t,(i-1)}] = \frac{\exp\{\Psi(c, \boldsymbol{y}_{t,i-1})\}}{\sum_{c' \neq \boldsymbol{y}_{t,i-1}} \exp\{\Psi(c', \boldsymbol{y}_{t,i-1})\}}$$

$$\Psi_{tc} = \lambda_c + \theta_u^{\mathrm{T}} \alpha_c$$

$$\Psi(c, \boldsymbol{y}_{t,i-1}) = \Psi_{tc} + \rho_c^{\mathrm{T}}\left(\frac{1}{\mathrm{i}-1} \sum_{j=1}^{i-1} \alpha_{y_{tj}}\right)$$

Customer's preferences

The Model

Intuition · **Mathematical Formulation** · Extensions

$\mathbf{y}_t = (y_{t1}, y_{t2}, \dots, y_{tn})$

$\mathbf{y}_{t,\, i-1} = (y_{t1}, y_{t2}, \dots, y_{t(i-1)})$

$$\Pr[y_{ti} = c \mid \mathbf{y}_{t,(i-1)}] = \frac{\exp\{\Psi(c, \boldsymbol{y}_{t,i-1})\}}{\sum_{c' \neq \boldsymbol{y}_{t,i-1}} \exp\{\Psi(c', \boldsymbol{y}_{t,i-1})\}}$$

$$\Psi_{tc} = \lambda_c + \theta_u^{\mathrm{T}} \alpha_c - \gamma_{u_t}^{T} \beta_c \log(r_{tc})$$

Price effects

$$\Psi(c, \boldsymbol{y}_{t,i-1}) = \Psi_{tc} + \rho_c^{\mathrm{T}} \left(\frac{1}{\mathrm{i} - 1} \sum_{j=1}^{i-1} \alpha_{y_{tj}}\right)$$

The Model

Intuition · **Mathematical Formulation** · Extensions

$$\Psi_{tc} = \lambda_c + \theta_u^{\mathrm{T}}\alpha_c - \gamma_{u_t}^T\beta_c \log(r_{tc}) + \delta_{w_t}^T\mu_c$$

Seasonal effects

$\mathbf{y}_t = (y_{t1}, y_{t2}, \ldots, y_{tn})$

$\mathbf{y}_{t,\, i-1} = (y_{t1}, y_{t2}, \ldots, y_{t(i-1)})$

$$\Pr[y_{ti} = c \mid \mathbf{y}_{t,(i-1)}] = \frac{\exp\{\Psi(c, \boldsymbol{y}_{t,i-1})\}}{\sum_{c' \neq \boldsymbol{y}_{t,i-1}} \exp\{\Psi(c', \boldsymbol{y}_{t,i-1})\}}$$

$$\Psi(c, \boldsymbol{y}_{t,i-1}) = \Psi_{tc} + \rho_c^{\mathrm{T}}\left(\frac{1}{\mathrm{i}-1}\sum_{j=1}^{i-1} \alpha_{y_{tj}}\right)$$

The Model

Intuition · **Mathematical Formulation** · Extensions

We want to estimate the best parameters from the data. This allows us to:

- Answer conterfactual queries
- Decide whether two products are complements/ substitutes
- Gaining insights about customers and products

$\mathbf{y}_t = (y_{t1}, y_{t2}, \ldots, y_{tn})$

$\mathbf{y}_{t,i-1} = (y_{t1}, y_{t2}, \ldots, y_{t(i-1)})$

$$\Pr[y_{ti} = c \mid \mathbf{y}_{t,(i-1)}] = \frac{\exp\{\Psi(c, \boldsymbol{y}_{t,i-1})\}}{\sum_{c' \neq \boldsymbol{y}_{t,i-1}} \exp\{\Psi(c', \boldsymbol{y}_{t,i-1})\}}$$

$$\Psi(c, \boldsymbol{y}_{t,i-1}) = \Psi_{tc} + \rho_c^{\mathrm{T}} \left(\frac{1}{\mathrm{i}-1} \sum_{j=1}^{i-1} \alpha_{y_{tj}}\right)$$

$$\Psi_{tc} = \lambda_c + \theta_u^{\mathrm{T}} \alpha_c - \gamma_{u_t}^{T} \beta_c \log(r_{tc}) + \delta_{w_t}^{T} \mu_c$$

The Model

Thinking ahead

- More realistic consumer behaviour
- Consider how the current choice influences the next k choices

Plan

1. The Model

2. Optimization/ Inference

3. Results/ Discussion/ Openings

Optimization

Given a dataset with observed characteristics $\boldsymbol{x} = \{u_t, w_t, r_t\}_{t=1}^{N}$ and shopping baskets $\widetilde{\boldsymbol{y}} = \{\widetilde{\boldsymbol{y}}_{\boldsymbol{t}}\}_{t=1}^{N}$ we want to find the latent parameters $\ell = \{\alpha, \rho, \lambda, \theta, \gamma, \beta, \delta, \mu\}$ that maximize the MAP estimation.

$$\Pr[\ell \mid \widetilde{\boldsymbol{y}}, \boldsymbol{x}] = \frac{\Pr[\ell] \ \prod_t \Pr[\widetilde{\boldsymbol{y}}_{\boldsymbol{t}} | \ell, \boldsymbol{x}_{\boldsymbol{t}}]}{\Pr[\widetilde{\boldsymbol{y}} | \ \boldsymbol{x}]}$$

Optimization: the problem is intractable

Approximate $\Pr[\ell \mid \tilde{\boldsymbol{y}}, \boldsymbol{x}]$ with a tractable approximating distribution $q(\ell, \nu)$ that minimizes the KL divergence. Or, equivalently, maxizing the ELBO

$$argmax_{\nu} \mathbb{E}_{q(\ell,\nu)}[\log \Pr[\ell, \tilde{\boldsymbol{y}} | \boldsymbol{x}] - \log q(\ell, \nu)]$$

Optimization: the problem is intractable

Approximate $\Pr[\ell \mid \tilde{\boldsymbol{y}}, \boldsymbol{x}]$ with a tractable approximating distribution $q(\ell, \nu)$ that minimizes the KL divergence. Or, equivalently, maxizing the ELBO

$$argmax_{\nu} \mathbb{E}_{q(\ell,\nu)}[\log \Pr[\ell, \tilde{\boldsymbol{y}} | \boldsymbol{x}] - \log q(\ell, \nu)]$$

This gives a tractable approximated distribution $q(\ell, \nu^*)$ that allows to estimate the final parameters ℓ.

Main Ideas to tackle intractable problems

Stochastic optimization addresses

- Large datasets (expensive gradients)
- Intractable expectations

Variational lower bounds on ELBO addresses

- Unordered baskets
- Large number of items

Plan

1. The Model

2. Optimization/ Inference

3. Results/ Discussion/ Openings

Plan

1.The Mod

2. Optimiz

3. Results/

"LOOKING AT THE GOF IS NOT ENOUGH TO DO COUNTERFACTUALS"

Susan Athey

Dataset

- 6 mio. items
- 570 000 baskets
- 3200 customers
- 5600 unique items
- 97 weeks of data

Source: https://en.wikipedia.org

Quantitative Results

Model	Log-likelihood All (540K)	Price ± 2.5% (231K)	Price ± 5% (139K)	Price ± 15% (25K)
B-Emb (Rudolph et al. (2016))	−5.119 (0.001)	−5.119 (0.002)	−5.148 (0.002)	−5.250 (0.006)
P-Emb (Rudolph et al. (2016))	−5.160 (0.001)	−5.138 (0.002)	−5.204 (0.002)	−5.311 (0.005)
HPF (Gopalan, Hofman and Blei (2015))	−4.914 (0.002)	−4.931 (0.002)	−4.994 (0.003)	−5.061 (0.009)
SHOPPER (I+U)	−4.744 (0.002)	−4.743 (0.003)	−4.776 (0.003)	−4.82 (0.01)
SHOPPER (I+U+S)	−4.730 (0.002)	−4.778 (0.003)	−4.801 (0.004)	−4.83 (0.01)
SHOPPER (I+U+P)	−4.728 (0.002)	−4.753 (0.003)	−**4.747** (0.004)	−4.69 (0.01)
SHOPPER (I+U+P+S)	−**4.724** (0.002)	−**4.741** (0.003)	−4.774 (0.004)	−**4.64** (0.01)

Source: SHOPPER paper

Qualitative Results: Inference of Latent Factors

Mollusks	Organic vegetables	Granulated sugar	Cat food dry/moist
finfish all other—frozen	organic fruits	flour	cat food wet
crustacean nonshrimp	citrus	baking ingredients	cat litter & deodorant
shrimp family	cooking vegetables	brown sugar	pet supplies

Source: SHOPPER paper

Qualitative Results: Seasonal Effects

Halloween candy		Cherries		Turkey—frozen	
3.46	2006/10/25	3.07	2006/06/28	3.56	2005/11/16
3.34	2005/10/26	3.01	2006/07/12	3.30	2006/11/15
2.81	2005/10/19	2.85	2006/06/21	2.64	2005/11/23
		⋮			
−1.28	2005/11/23	−3.59	2006/10/11	−1.25	2006/06/21
−1.31	2007/01/03	−3.89	2006/10/18	−1.29	2006/07/05
−1.33	2005/11/16	−4.54	2006/10/25	−1.30	2006/07/19

Source: SHOPPER paper

Qualitative Results: Complements and Substitutes

Complementarity metric:

$$C_{cc'} = \frac{1}{2}(\rho_c^T \alpha_c + \rho_{c'}^T \alpha_c)$$

Exchangeability metric:

$$E_{cc'} = \frac{1}{2}\left(D_{KL}(p_{\cdot|c} || p_{\cdot|c'}) + D_{KL}(p_{\cdot|c'} || p_{\cdot|c})\right)$$

query	complementarity score		exchangeability score	
mission tortilla taco 1	2.40	taco bell seasoning mix	0.05	mission fajita
	2.26	mcrmck seasoning mix	0.07	mission tortilla taco 2
	2.24	lawrys seasoning mix	0.13	mission tortilla fluffy gordita
(private) hot dog buns	2.99	bp franks meat	0.11	ball park hot dog buns
	2.63	bp franks bun size	0.13	(private) hot dog potato buns 1
	2.37	bp franks beed bun length	0.15	(private) hot dog potato buns 2

Source: SHOPPER paper

Openings on this Line of Research

- Econometrics + Machine Learning: https://www.youtube.com/watch?v=mjGPF5R5JOQ&t=347s

Openings on this Line of Research

- Econometrics + Machine Learning: https://www.youtube.com/watch?v=mjGPF5R5JOQ&t=347s

- Counterfactual inference is used also on other scenarios: news consumption; what happens to the market if a shop closes...

Openings on this Line of Research

- Econometrics + Machine Learning: https://www.youtube.com/watch?v=mjGPF5R5JOQ&t=347s
- Counterfactual inference is used also on other scenarios: news consumption; what happens to the market if a shop closes…
- The authors used the same ideas in a very interesting paper to answer counterfactual queries about restaurant locations: https://arxiv.org/pdf/1801.07826.pdf

Openings on this Line of Research

I. Empirical Model and Estimation

We model the consumer's choice of restaurant conditional on deciding to go out to lunch. We assume that the consumer selects the restaurant that maximizes utility, where the utility of user u for restaurant i on her t-th visit is

$$U_{uit} = \lambda_i + \theta_u^\top \alpha_i + \mu_i^\top \delta_{w_{ut}} - \gamma_u^\top \beta_i \cdot \log(d_{ui}) + \epsilon_{uit},$$

Openings on this Line of Research

- Econometrics + Machine Learning: https://www.youtube.com/watch?v=mjGPF5R5JOQ&t=347s
- Counterfactual inference is used also on other scenarios: news consumption; what happens to the market if a shop closes…
- The authors used the same ideas in a very interesting paper to answer counterfactual queries about restaurant locations: https://arxiv.org/pdf/1801.07826.pdf
- Matrix factorization, embedding methods and Bayesian inference.

My Take

\+ SHOPPER is good at counterfactuals (if training data is good enough)

\+ Interpretability: extraction of valuable latent features

+/- Practical paper, but there is still a gap to use the model on business decisions

\- More details on the related work in the community would have been appreciated

Thank you for your attention.

References

- Counterfactual Inference for Consumer Choice Across Many Product Categories: https://arxiv.org/pdf/1906.02635.pdf
- Code Implementation SHOPPER: https://github.com/franrruiz/shopper-src

Backup Slides

How to use the model for Business Decisions

- Not clear how answering counterfactual queries about prices on a specific basket can help taking business decisions.
- Detect complements that are located physically far away in the store and put them together.
- We can detect the favorite products of a certain (class of) customer(s) and decide to promote them.
- Interpretation of the latent parameters not too useful here. Does this model generalize in more "interesting" scenarios?

How to use the model for Business Decisions

- Not clear how a specific basket can help
- Detect complem tore and put them toget
- We can detect t er(s) and decide to prom
- Interpretation this model generalize in more “interesting” scenarios?

REINFORCEMENT
LEARNING
APPROACH

Why Thinking Ahead?

Illustration · Results

- Better description of consumer’s behavior
- Choice: pick A with utility 1 or B with utility 2. If you picked 1 you can take an item with utility 10, if you picked B you can take an item with utility 5 ➔ Greedily taking B in step 1 is suboptimal.

Toy Simulation: Performance Boost with Thinking Ahead

Illustration · Results

- Items: Coffee, Diapers, Ramen, Candy, Hot dogs, Hot Dog Buns, Taco Shells, Taco Seasoning
- Customers: 50% are new parents and never buy candy or ramen; 50% are students and nevery buy cofee of diapers
- Either you by Taco Shells AND Taco Seasoning or none. Same for Hot Dogs and Hot Dog Buns.
- About preferred items: if price is low buy with probability 0.95, otherwise 0.1. High price with probability 0.4.

Complementary 1	Complementary 2	Probability
LOW	LOW	0.5 for both
LOW	HIGH	0.85 for low, 0.15 for high
HIGH	LOW	0.85 for low, 0.15 for high
HIGH	HIGH	Does not happen

Toy Simulation: Performance Boost with Thinking Ahead

Illustration · **Results**

		stage 1: *diapers*	stage 2: *hot dogs*	stage 3: *hot dog buns*	stage 4: *checkout*
non think-ahead	diapers	**0.31**	0.00	0.00	0.00
	coffee (↑)	0.03	0.02	0.05	0.21
	ramen	0.00	0.00	0.00	0.00
	candy	0.00	0.00	0.00	0.00
	hot dogs	0.18	**0.25**	0.00	0.00
	hot dog buns	0.17	0.25	**0.79**	0.00
	taco shells (↑)	0.14	0.19	0.00	0.00
	taco seasoning	0.17	0.24	0.00	0.00
	checkout	0.00	0.05	0.16	**0.79**
think-ahead	diapers	**0.37**	0.00	0.00	0.00
	coffee (↑)	0.02	0.02	0.07	0.10
	ramen	0.00	0.00	0.00	0.00
	candy	0.00	0.00	0.00	0.00
	hot dogs	0.24	**0.34**	0.00	0.00
	hot dog buns	0.24	0.42	**0.85**	0.00
	taco shells (↑)	0.06	0.10	0.00	0.00
	taco seasoning	0.06	0.10	0.00	0.00
	checkout	0.00	0.02	0.08	**0.90**

Source: SHOPPER paper

Thinking Ahead on a Larger Dataset

Illustration · **Results**

	Three items	Entire baskets
Non think-ahead	-4.795 (0.005)	-4.96 (0.02)
Think-ahead	$\mathbf{-4.747}$ (0.004)	$\mathbf{-4.91}$ (0.02)

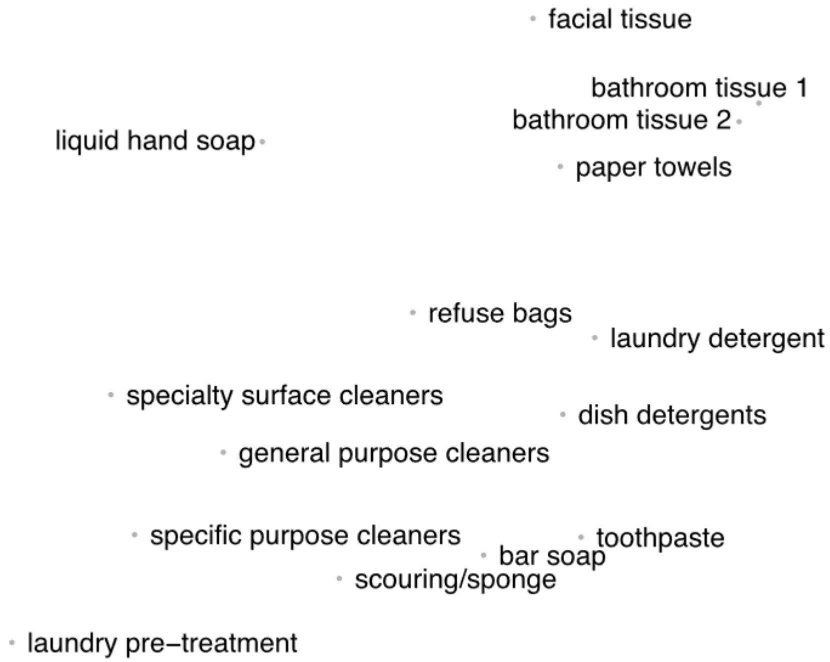
Source: SHOPPER paper

The Dataset

- 570 878 baskets rom a single grocery store
- 6 mio. purchases from 5590 unique items; 3206 different customers
- Span of 97 weeks
- Heterogeneous data (e.g. Prices change regularly)
- Training + validaton + test split

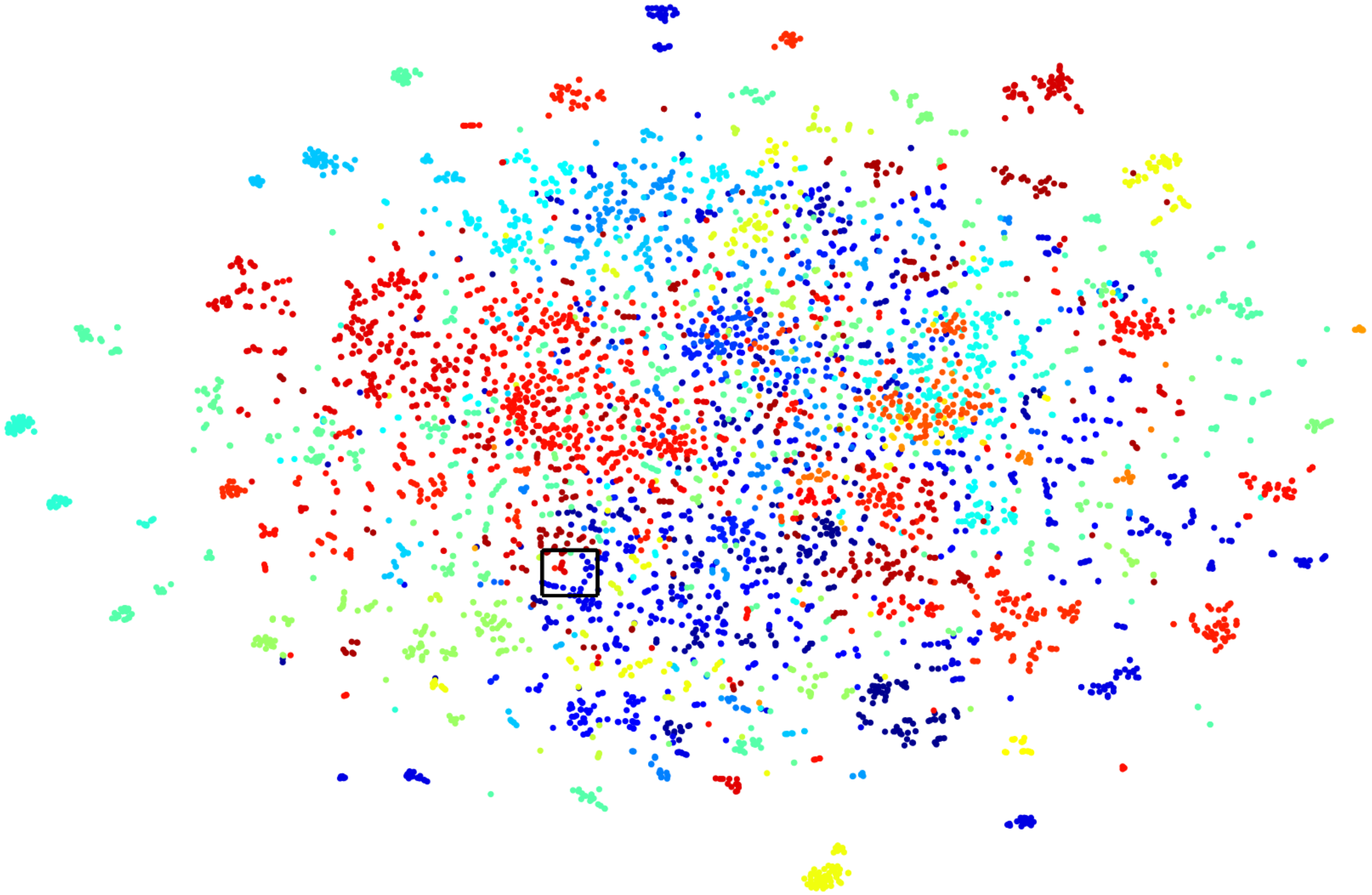
Inference of Latent Features

(a) Pet food and supplies.



(b) Cleaning and hygiene.

Qualitative Results: Inference of Latent Factors



Source: SHOPPER paper