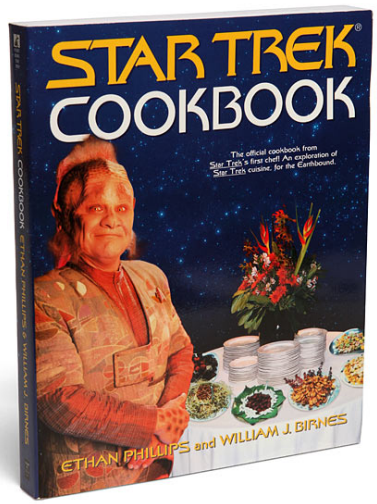
Economics and Probabilistic Machine Learning

David M. Blei
Departments of Computer Science and Statistics
Columbia University

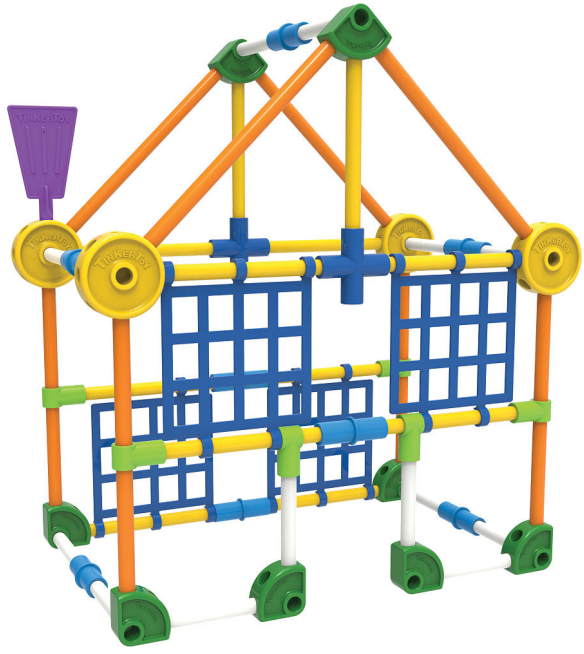
Modern probabilistic modeling

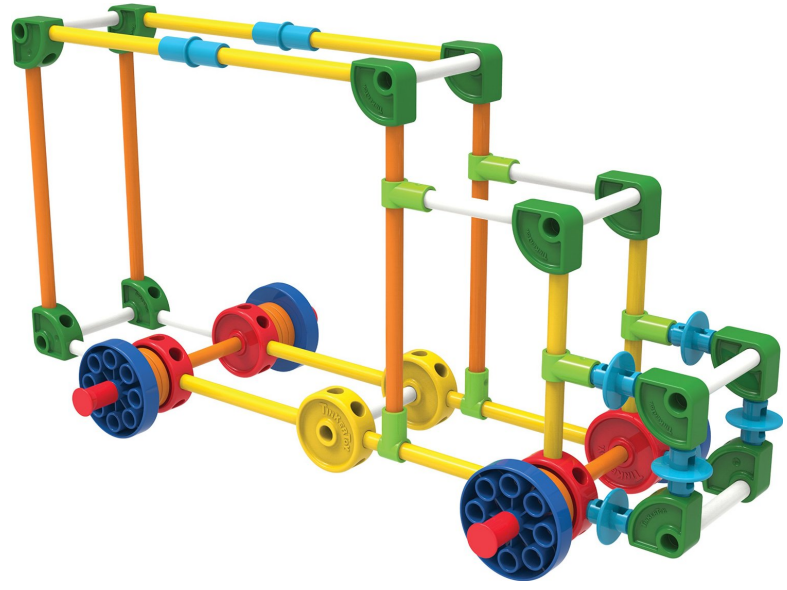
An efficient framework for discovering meaningful patterns in massive data.



How to use traditional machine learning and statistics to solve modern problems

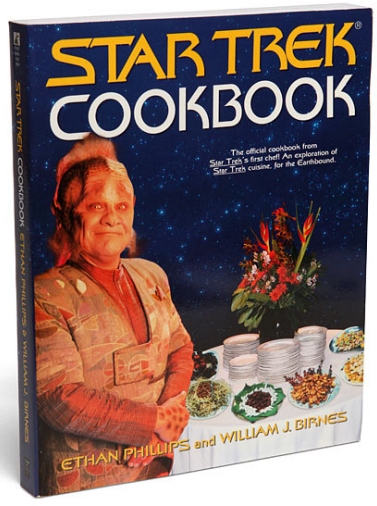
Probabilistic machine learning: tailored models for the problem at hand.



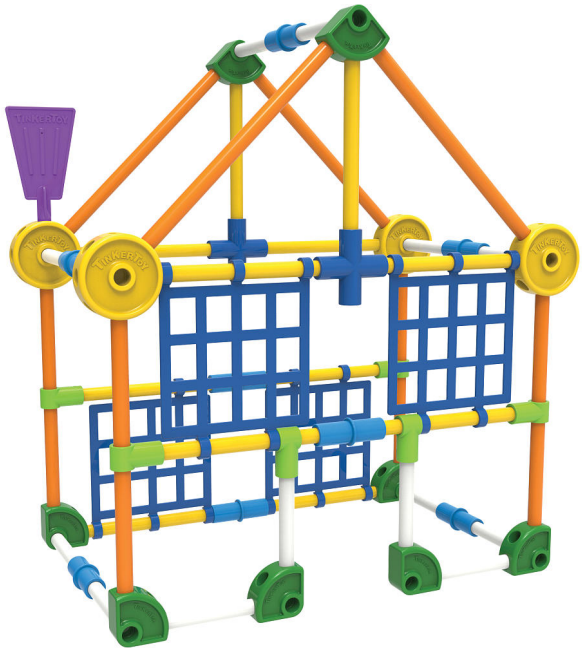


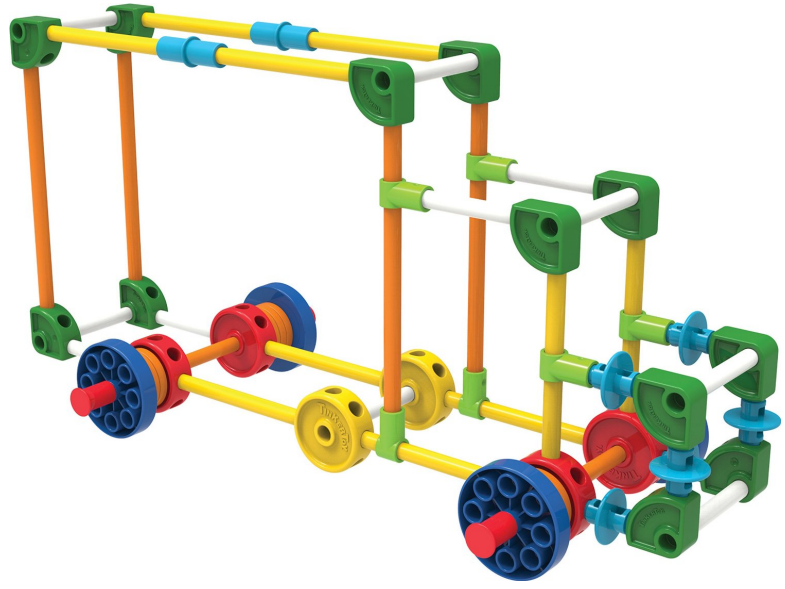
Probabilistic machine learning: tailored models for the problem at hand.

- Compose and connect reusable parts
- Driven by disciplinary knowledge and its questions
- Large-scale data, both in terms of data points and data dimension
- Focus on discovering and using structure in unstructured data
- Exploratory, observational, causal analyses

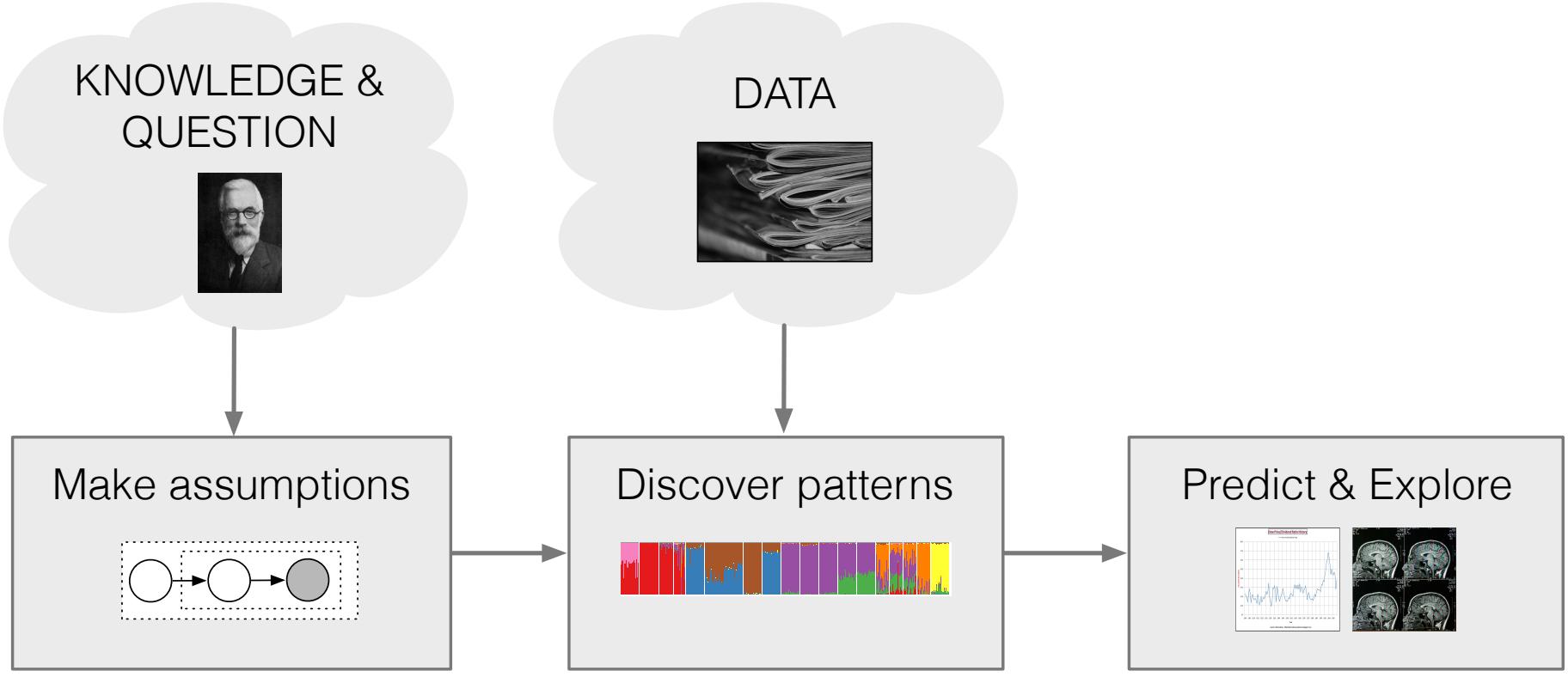


Many software packages available; typically fast and scalable



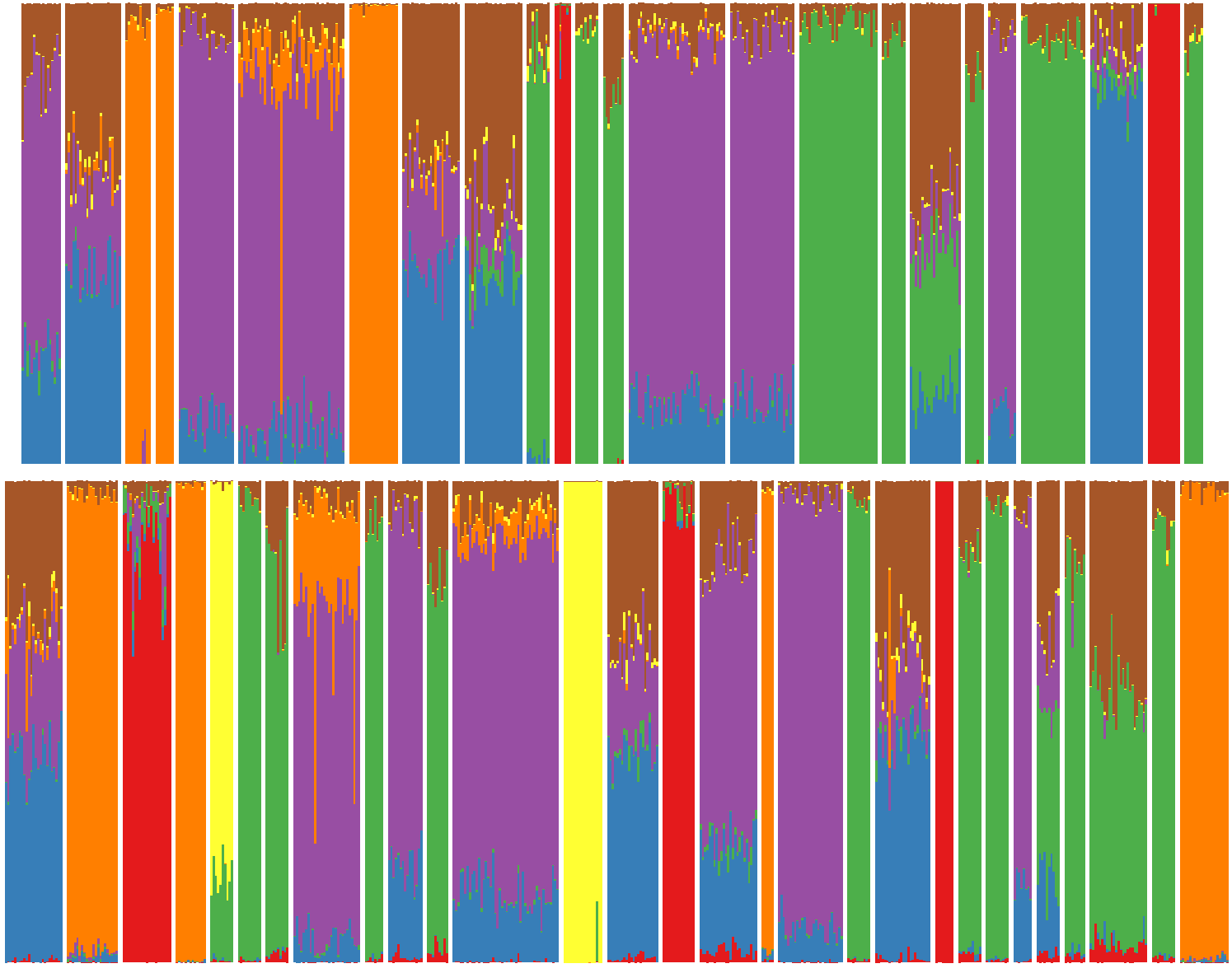


More challenging to implement; may not be fast or scalable

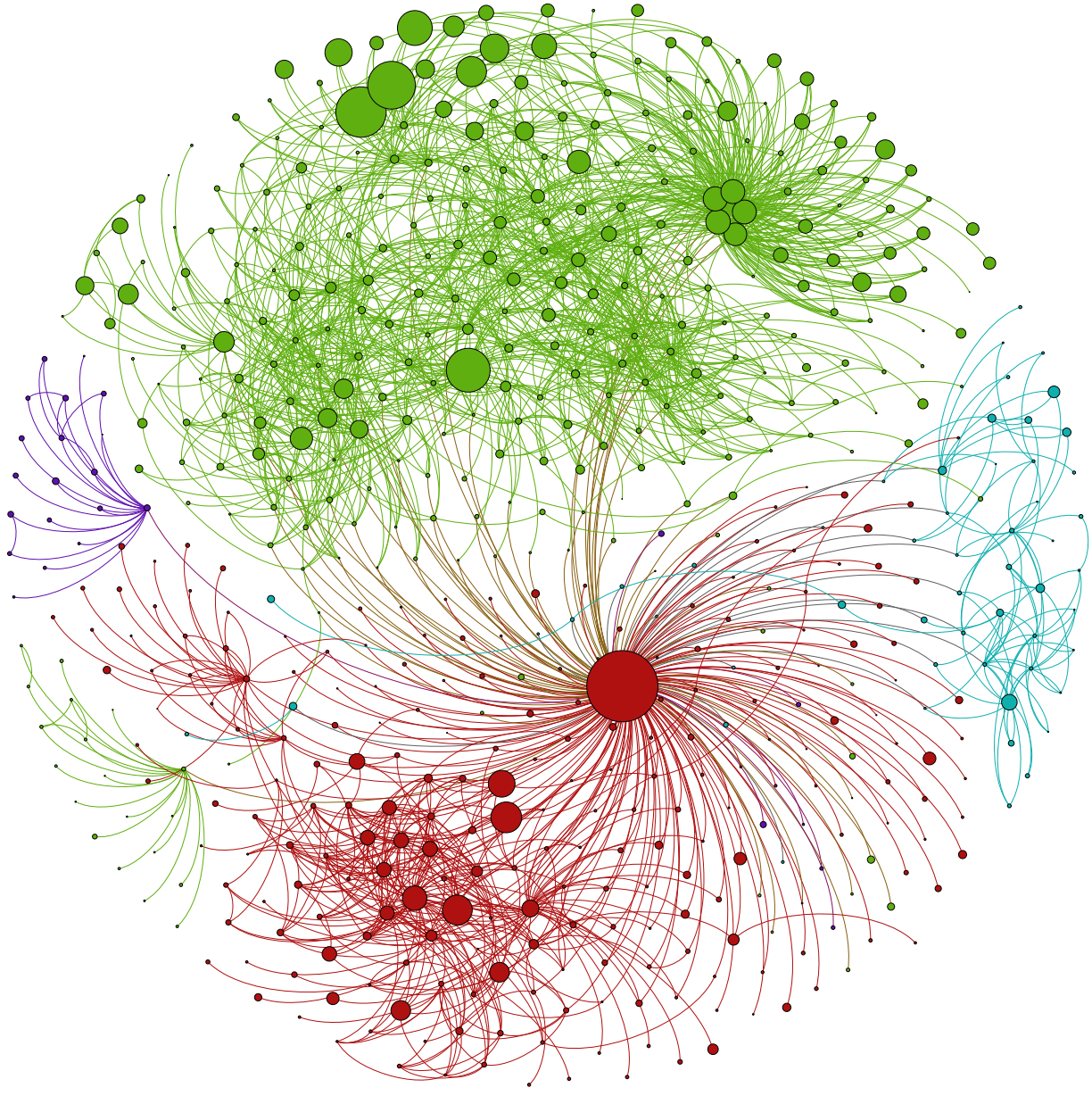


The probabilistic pipeline

- *Design* models that reflect our domain expertise and knowledge
- Given data, *compute* the approximate posterior of hidden variables
- Use the computation to *predict* the future or *explore* the patterns in your data.

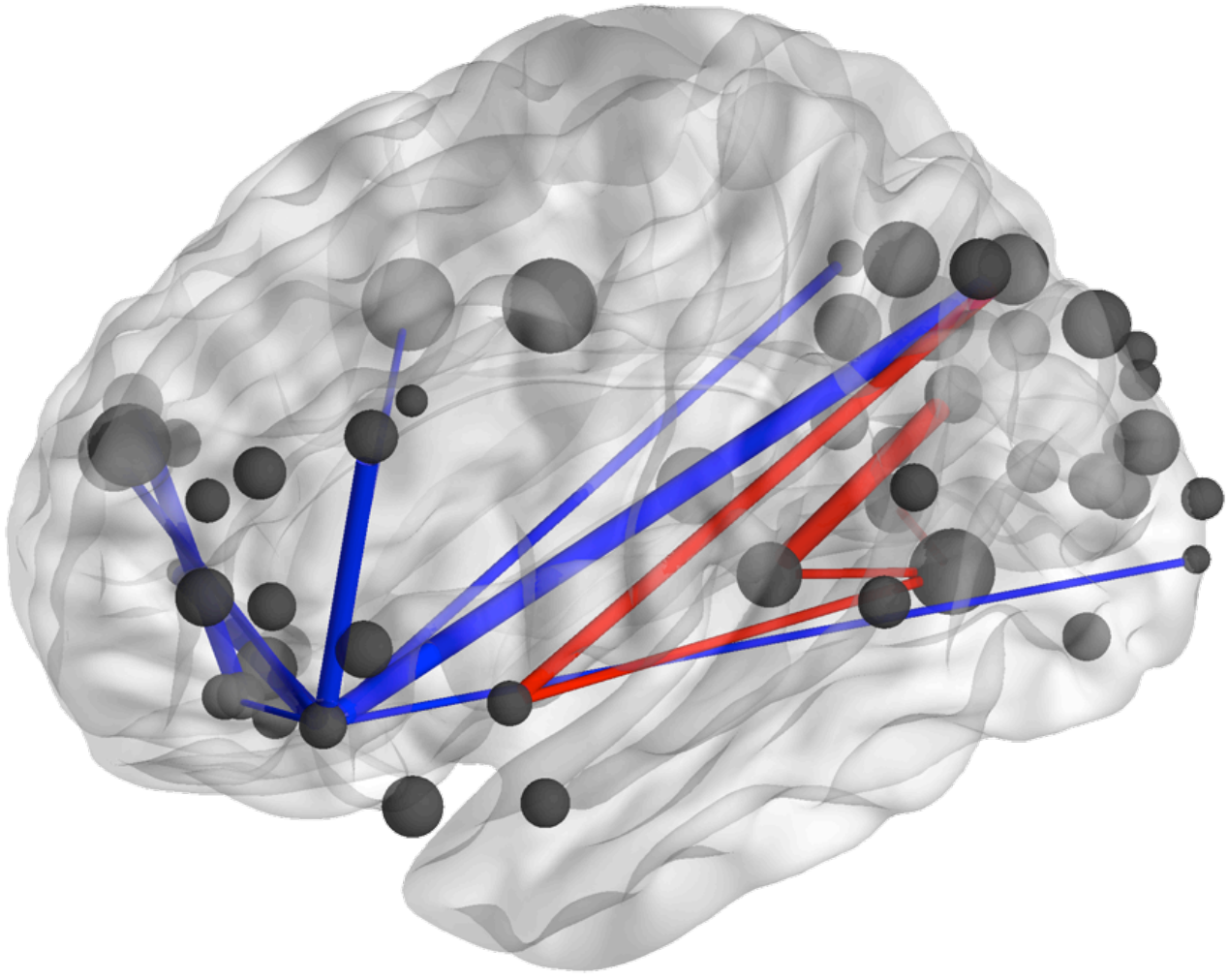


Population analysis of 2 billion genetic measurements

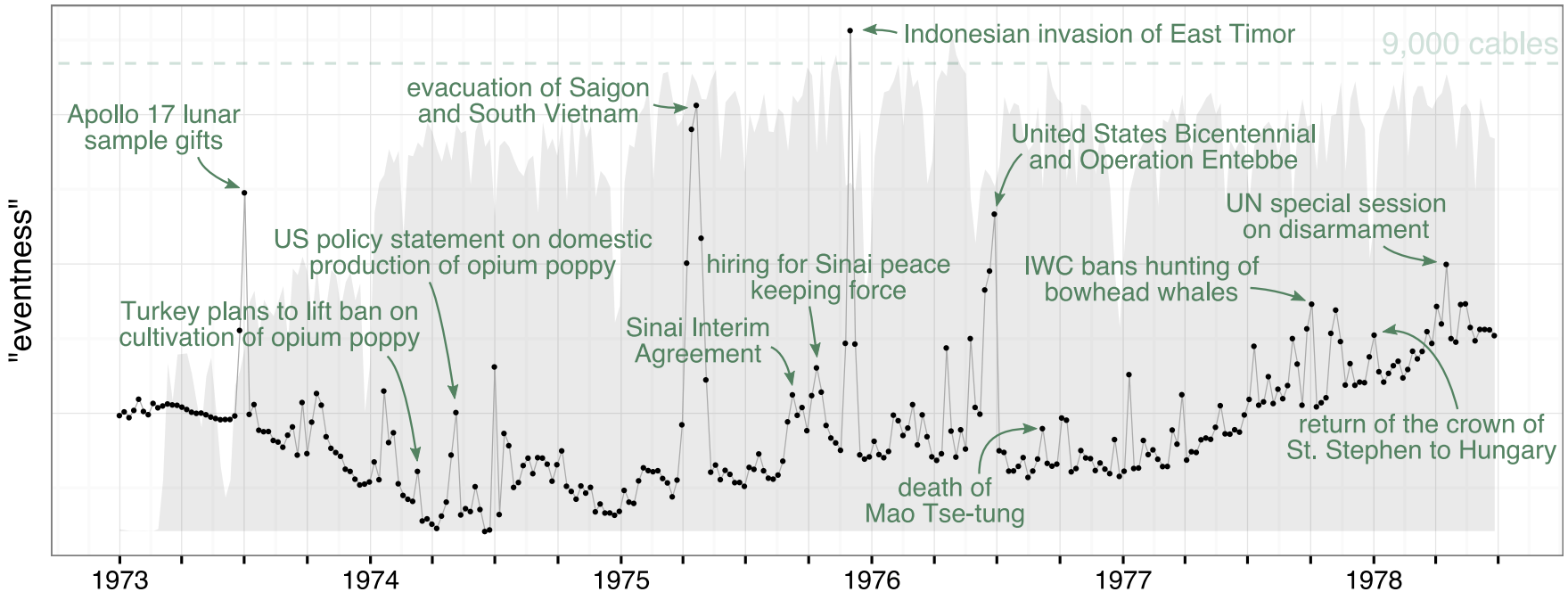


Communities discovered in a 3.7M node network of U.S. Patents

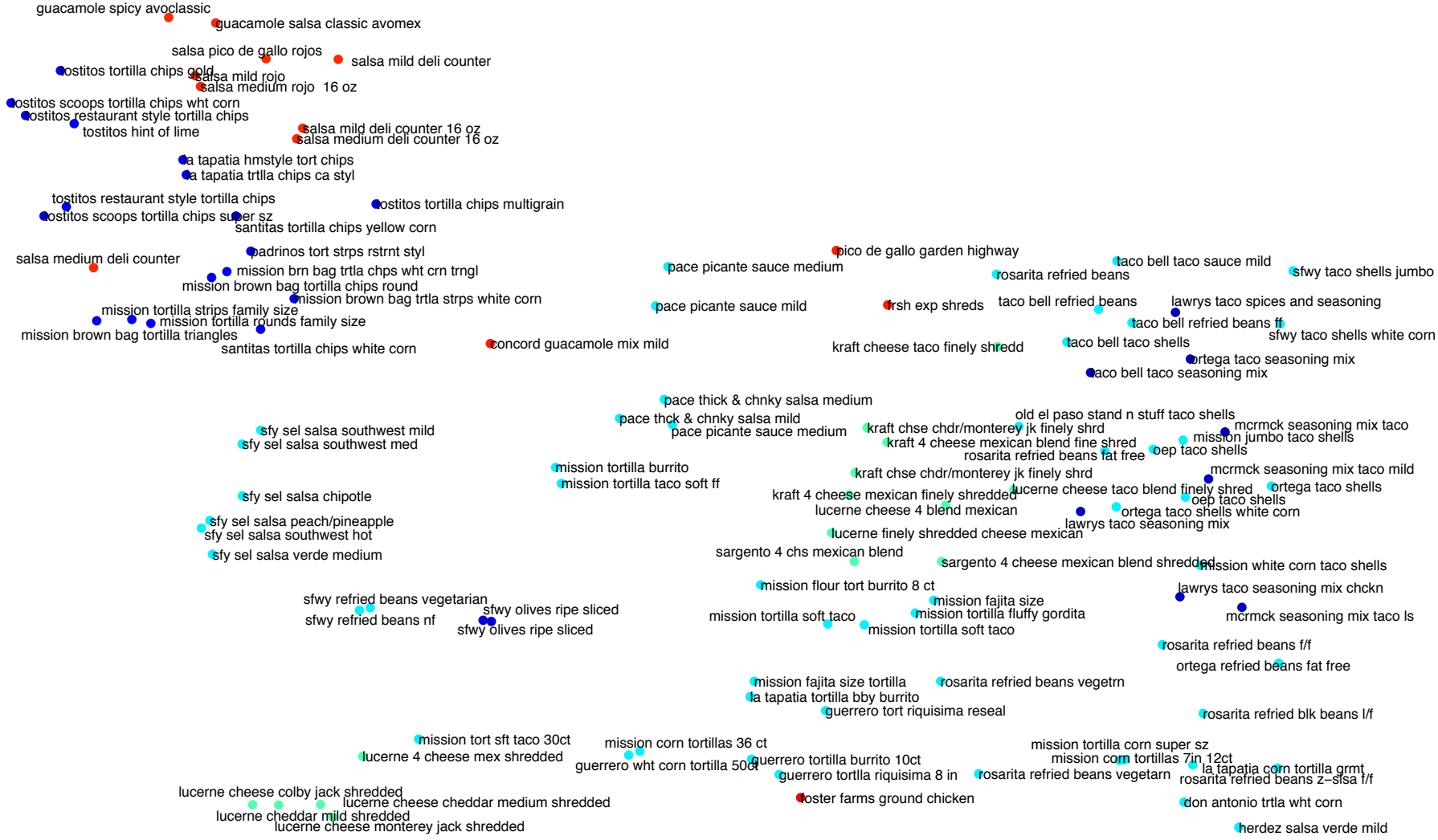
Topics found in 1.8M articles from the New York Times



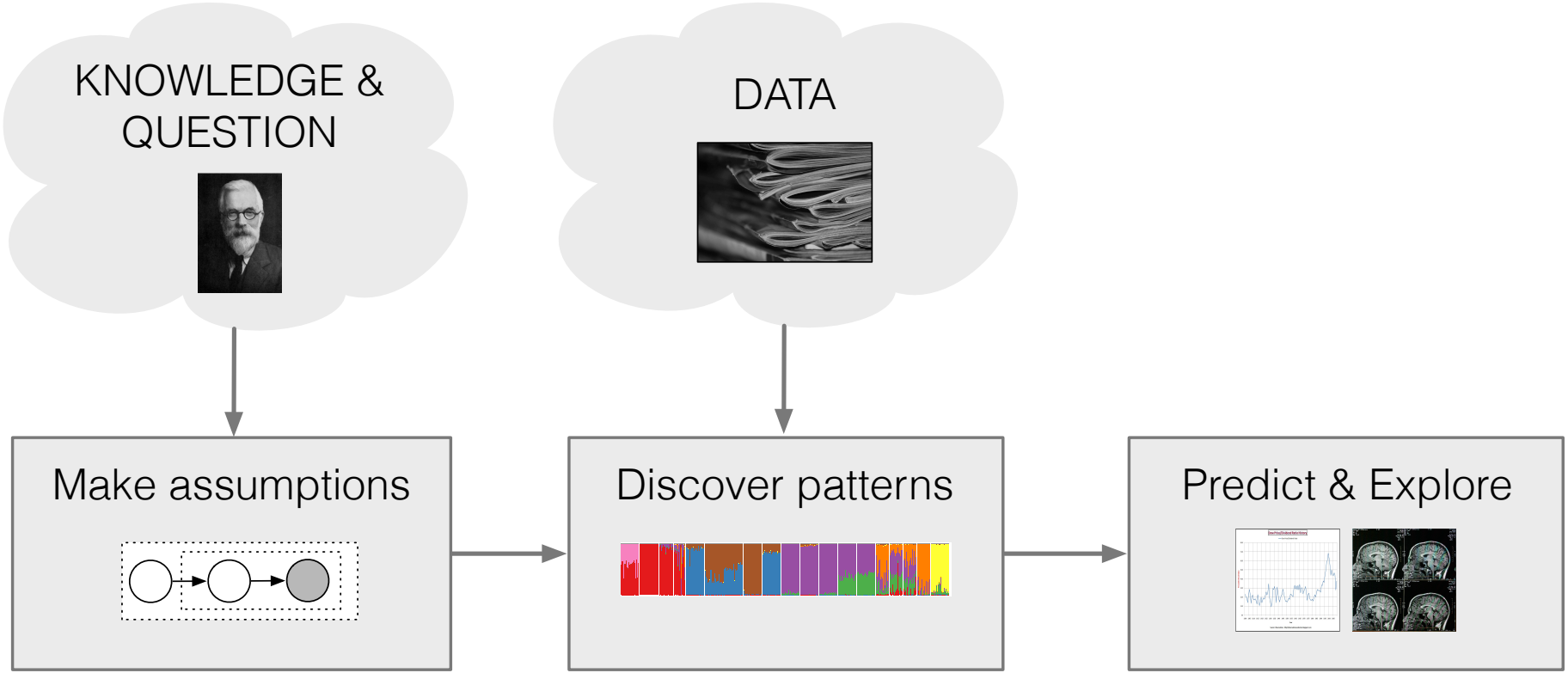
Neuroscience analysis of 220 million fMRI measurements



Events uncovered from 2M diplomatic cables

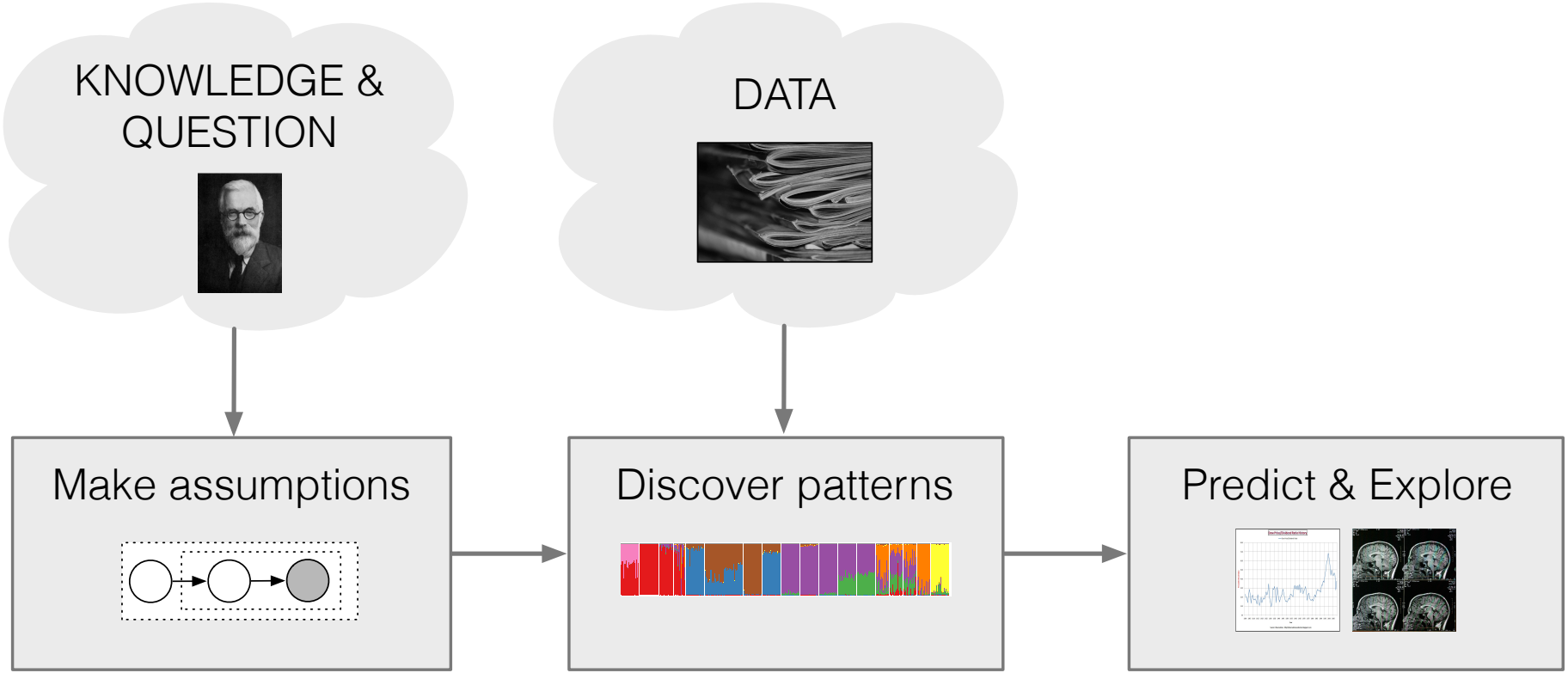


Item characteristics learned from 5.6M purchases



Our perspective:

- Customized data analysis is important to many fields.
- This pipeline separates assumptions, computation, application.
- It facilitates solving data science problems.



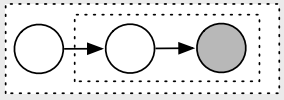
What we need:

- **Flexible** and **expressive** components for building models
- **Scalable** and **generic** inference algorithms
- **New applications** to stretch probabilistic modeling into new areas

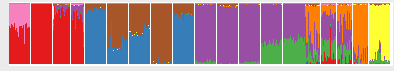
KNOWLEDGE & QUESTION

DATA

Make assumptions

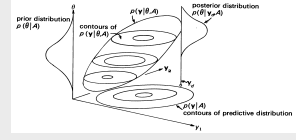


Discover patterns



Predict & Explore

Criticize model



Revise

- Here I discuss two threads of research with Susan Athey's group.
- Build probabilistic models to analyze large-scale consumer behavior; many consumers choosing among many items
- (Caveat: I'm not an economist.)
- also joint with Francisco Ruiz

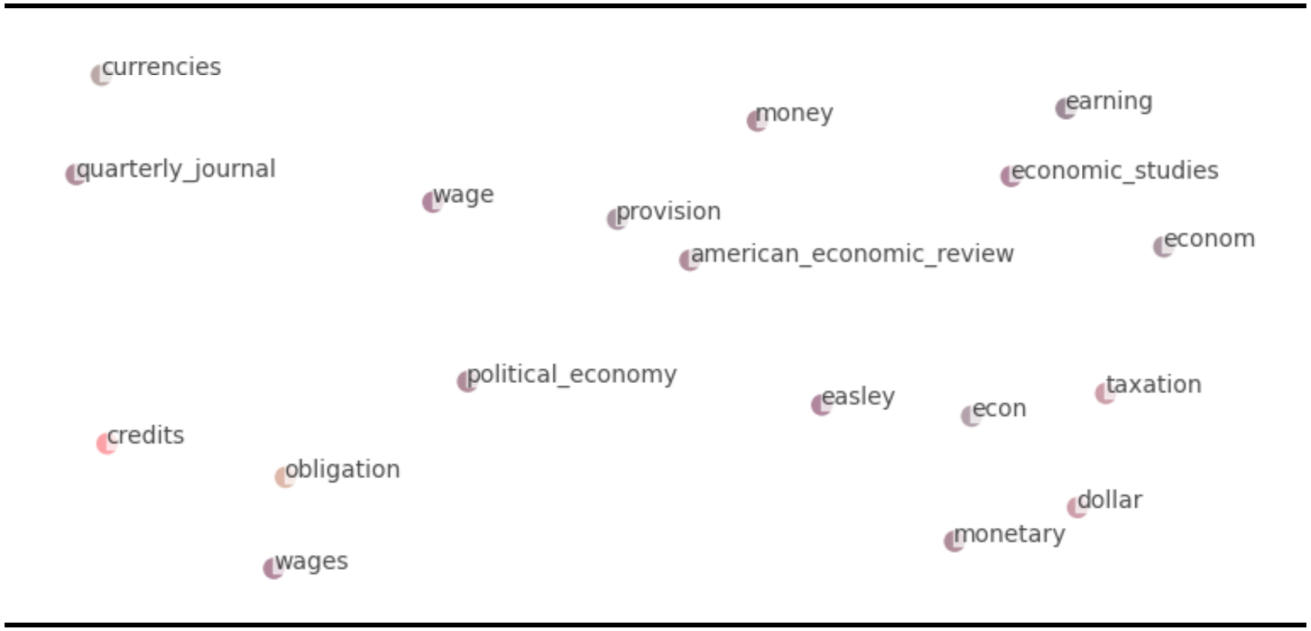
- Vision: a utility model for baskets of items:

$$U(\text{basket}) = [\text{subs/comps}] + [\text{shopper}] + [\text{prices}] + [\text{other}] + \epsilon$$

- Goals
  - design, fit, check, and revise this model
  - answer counterfactual questions about purchase behavior

Economic embeddings

Identifying substitutes and co-purchases in large-scale consumer data.



- **Word embeddings** are a powerful approach for analyzing language.
- Discovers a *distributed representation* of words
  - Distances appear to capture semantic similarity.
- Many variants, but each reflects the same main ideas:
  - Words are placed in a low-dimensional latent space
  - A word's probability depends on its distance to other words in its context

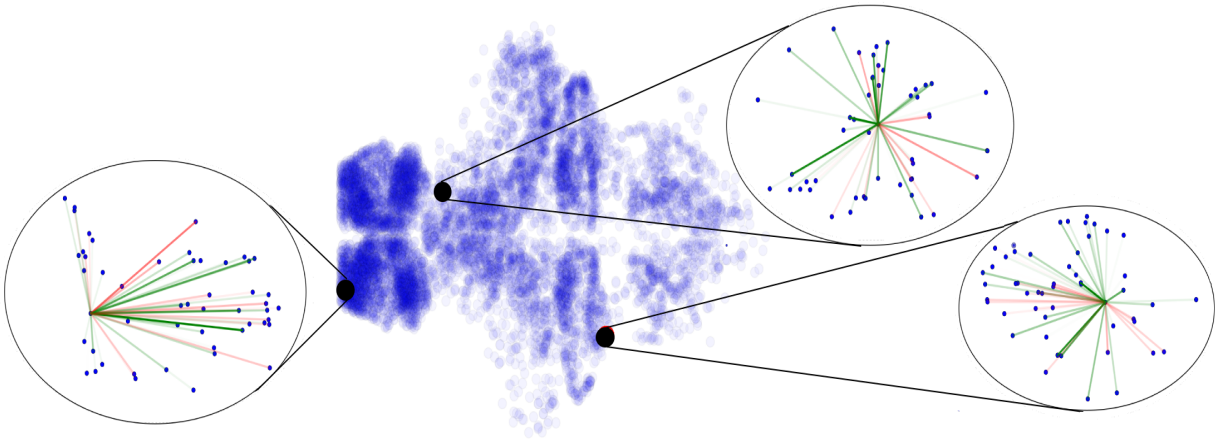
---

Bengio et al., A neural probabilistic language model. *Journal of Machine Learning Research*, 2003.
Mikolov et al., Efficient estimation of word representations in vector space. *Neural Information Processing Systems*, 2013.
Image: Paul Ginsparg

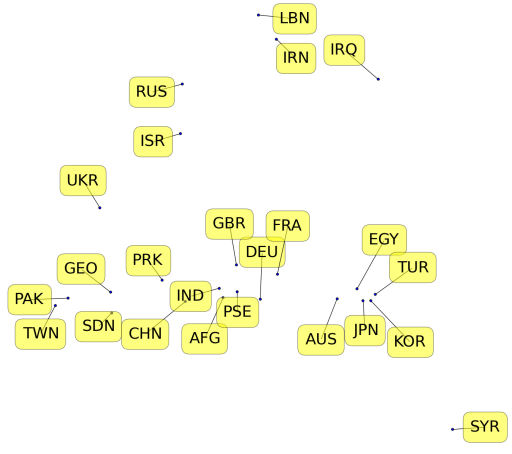
- **Exponential family embeddings** generalize this idea to other types of data.
- Use *generalized linear models* and *exponential families*
- Examples:

  neuroscience; recommender systems; networks; shopping baskets
- Bigger picture: A statistical perspective on ideas from neural networks.

Zebrafish brain activity



Interactions between countries



Vacation-town deli

Jam

PB #1

PB #2

Soda


Bread


Pizza

- Consider a vacation-town deli; it has six items.
- Customers either buy [pizza, soda] or [peanut butter, jam, bread]
  Customers only buy one type of peanut butter at a time.
- Items bought together (or not) are *co-purchased* (or not).
  The peanut butters are *substitutes*.
  (For now we ignore many issues, e.g., formal definitions, price, causality.)
- We would like to capture this purchase behavior.

Vacation-town deli

Jam

PB #1

PB #2

Soda


Bread

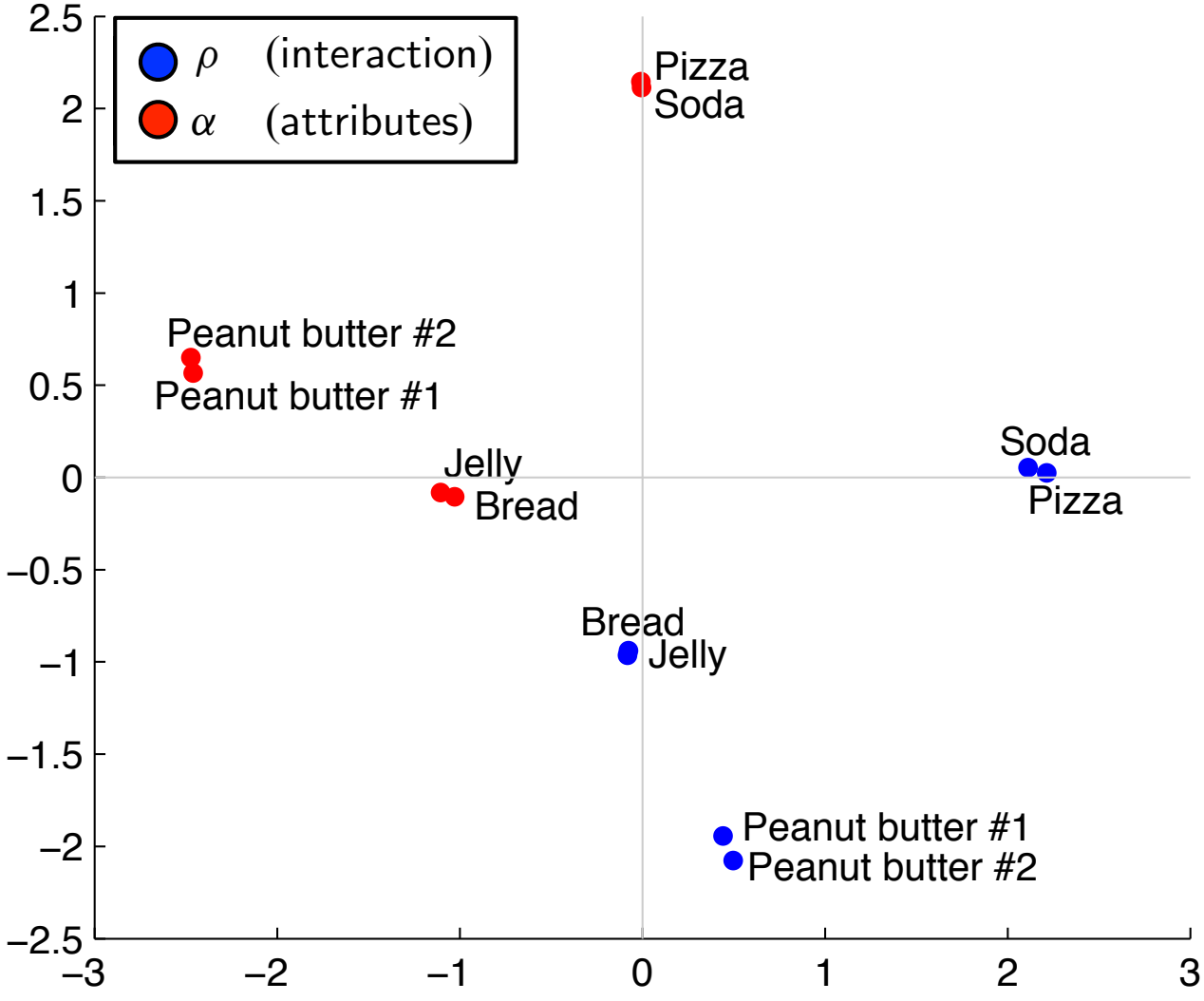

Pizza

- We endow each item with two (unknown) locations in a real space $\mathbb{R}^k$:
  an embedding ρ and context vector α.

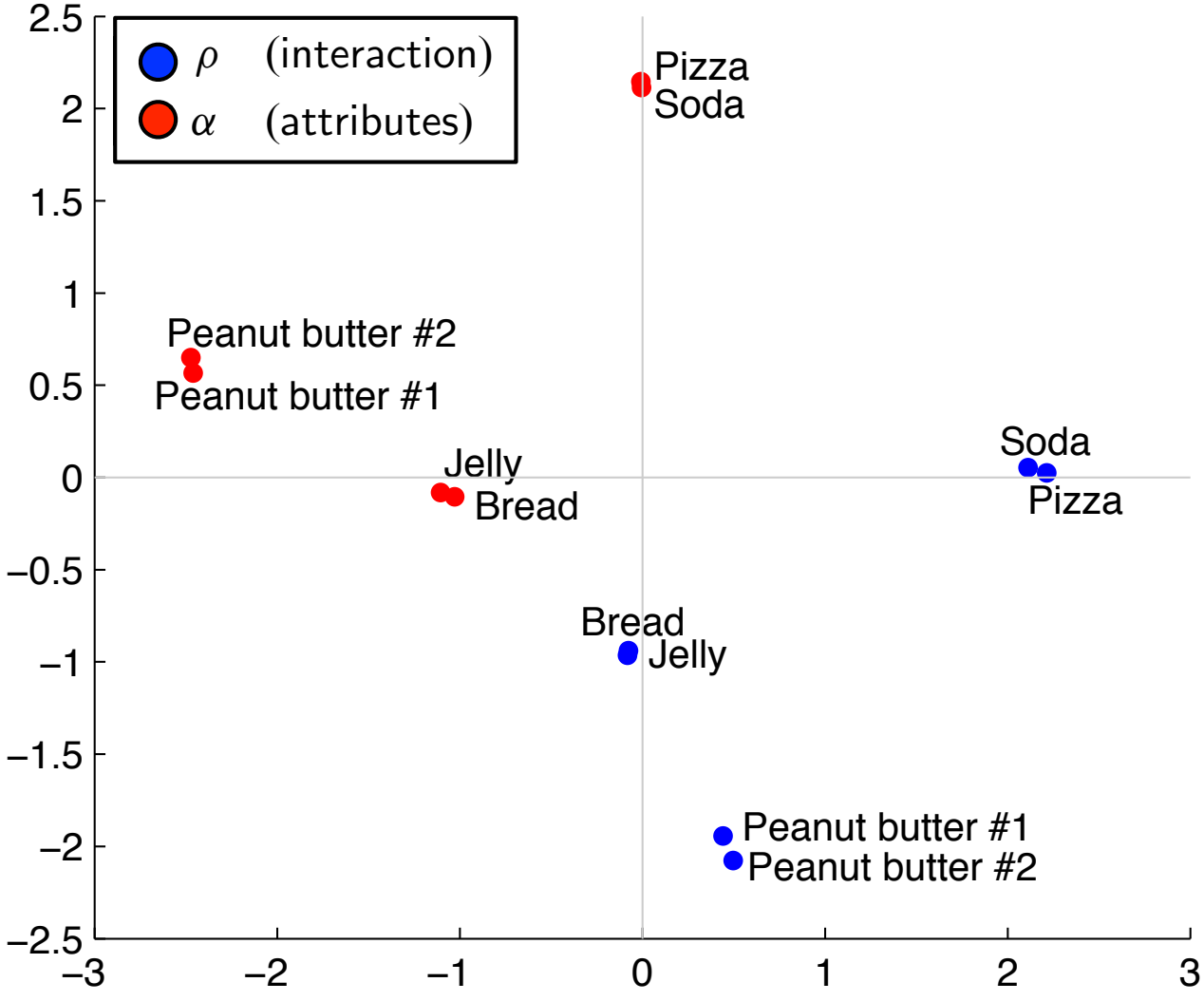
- The conditional probability of each item depends on its embedding and the context vectors of the other items in the basket,

$$x_{b,i} \mid \mathbf{x}_{b,-i} \sim \text{Poisson}\left(\exp\left\{\rho_i^\top \textstyle\sum_{j \neq i} \alpha_j x_{b,j}\right\}\right).$$

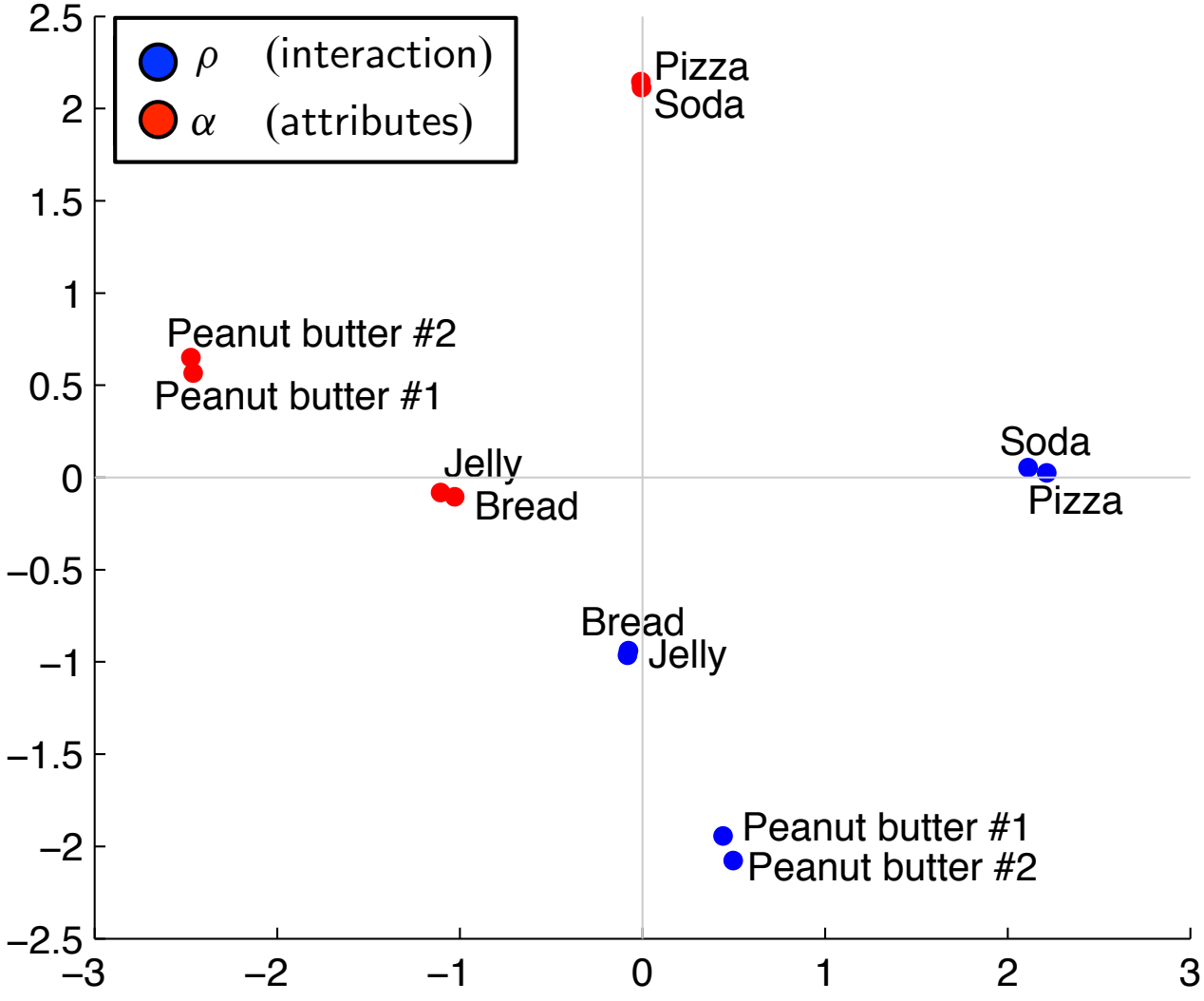
- α_i are latent product attributes
  ρ_i indicate how product i interacts with other products' attributes



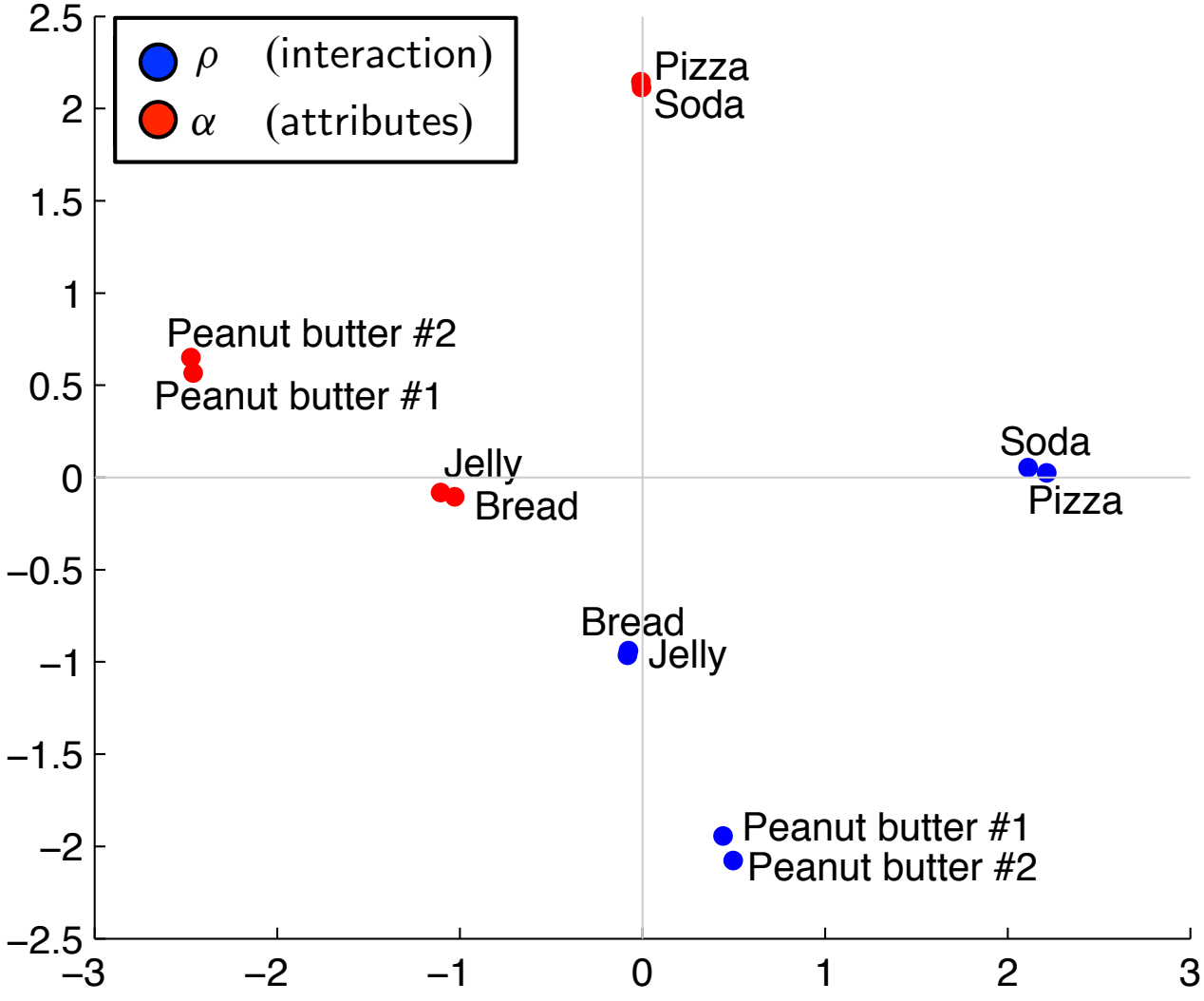
Pizza and soda are never bought with bread, jam, and PB; and vice versa



Bread, jam, and PB are bought together



PB #1 is never bought with PB #2



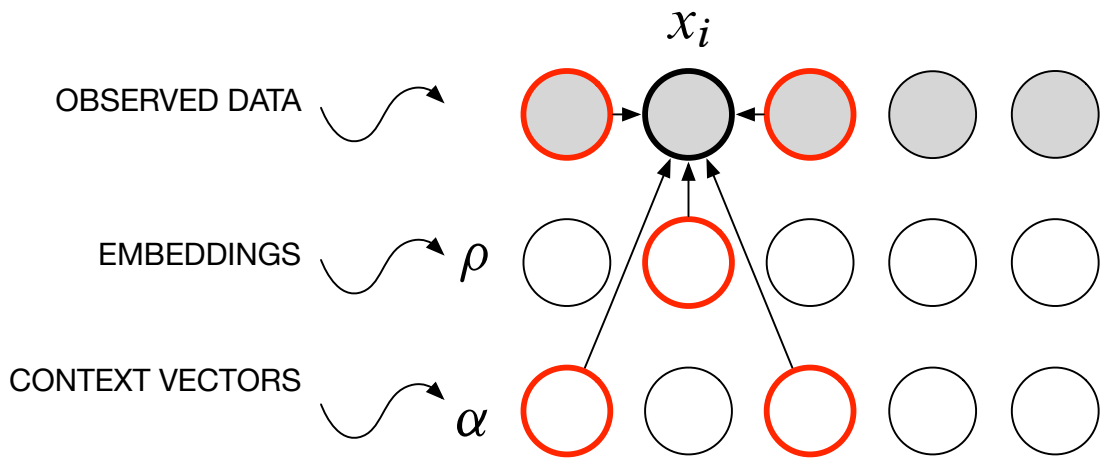
PB #1 is bought with similar items as PB #2

Exponential Family Embedding

- The goal of an EF-EMB is to discover a useful representation of data
- Observations $\mathbf{x} = x_{1:n}$, where x_i is a D-vector
- Examples:

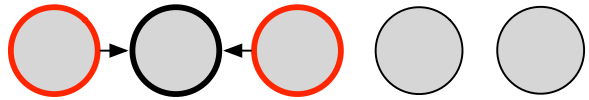
DOMAIN	INDEX	VALUE
Language	position in text i	word indicator
Neuroscience	neuron and time (n, t)	activity level
Network	pair of nodes (s, d)	edge indicator
Shopping	item and basket (d, b)	number purchased

Exponential Family Embedding



- Three ingredients:
  context, *conditional exponential family*, *embedding structure*
- Two latent variables per data index, an embedding and a context vector
- Model each data point conditioned on its context and latent variables.
- The latent variables interact in the conditional.
  How depends on which indices are in context and which one is modeled

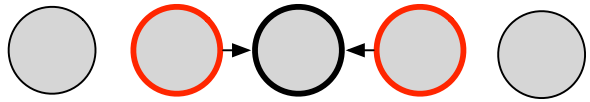
Context



- Each data point i has a *context* c_i, a set of indices of other data points.
- We model the conditional of the data point given its context, $p(x_i \mid \mathbf{x}_{c_i})$.
- Examples

DOMAIN	DATA POINT	CONTEXT
Language	word	surrounding words
Neuroscience	neuron activity	activity of surrounding neurons
Network	edge	other edges on the two nodes
Shopping	purchased item	other item counts on the same trip

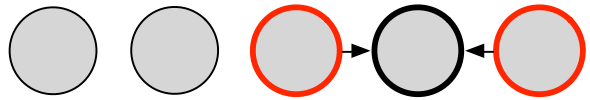
Context



- Each data point i has a *context* c_i, a set of indices of other data points.
- We model the conditional of the data point given its context, $p(x_i \mid \mathbf{x}_{c_i})$.
- Examples

DOMAIN	DATA POINT	CONTEXT
Language	word	surrounding words
Neuroscience	neuron activity	activity of surrounding neurons
Network	edge	other edges on the two nodes
Shopping	purchased item	other item counts on the same trip

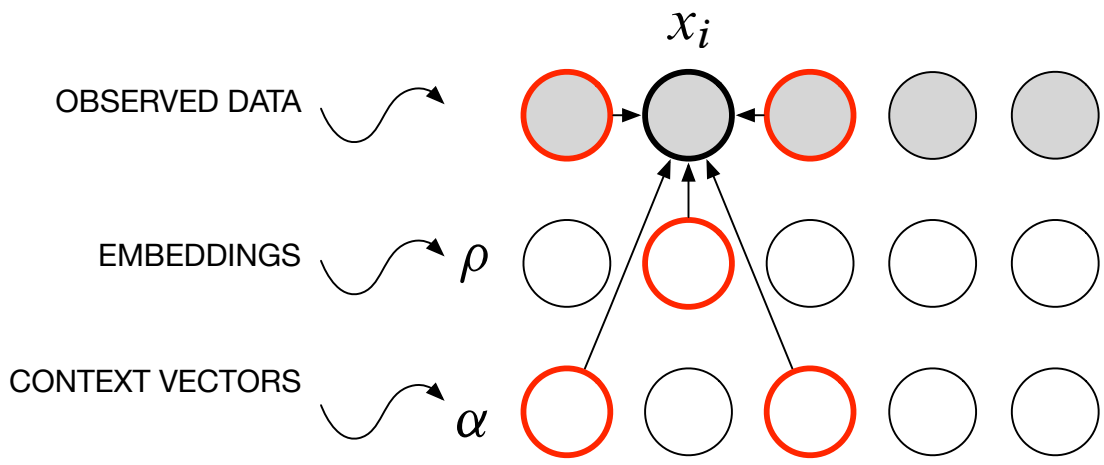
Context



- Each data point i has a *context* c_i, a set of indices of other data points.
- We model the conditional of the data point given its context, $p(x_i \mid \mathbf{x}_{c_i})$.
- Examples

DOMAIN	DATA POINT	CONTEXT
Language	word	surrounding words
Neuroscience	neuron activity	activity of surrounding neurons
Network	edge	other edges on the two nodes
Shopping	purchased item	other item counts on the same trip

Conditional exponential family



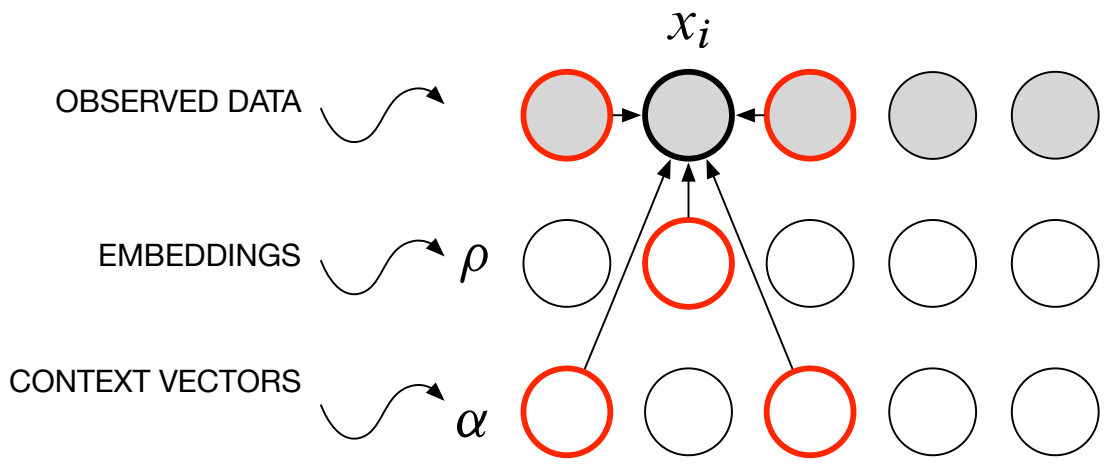
- The EF-EMB has latent variables for each data point's index:
  an *embedding* $\rho[i]$ and a *context vector* $\alpha[i]$

- These are used in the conditional of each data point,

$$x_i \mid \mathbf{x}_{c_i} \sim \text{exp-fam}(\eta(\mathbf{x}_{c_i}; \rho[i], \alpha[c_i]), t(x_i)).$$

(Poisson for counts, Gaussian for reals, Bernoulli for binary, etc.)

Conditional exponential family

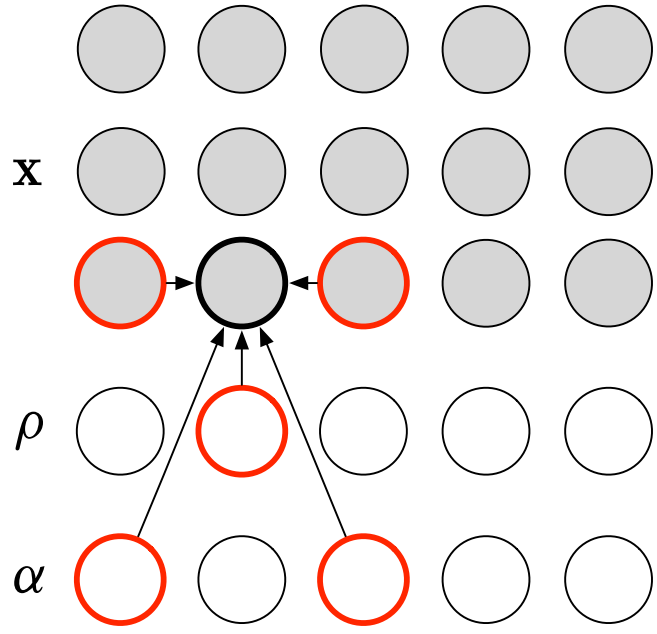


- The natural parameter combines the embedding and context vectors,

$$\eta_i(\mathbf{x}_{c_i}) = f\left(\rho[i]^\top \sum_{j \in c_i} \alpha[j] x_j\right),$$

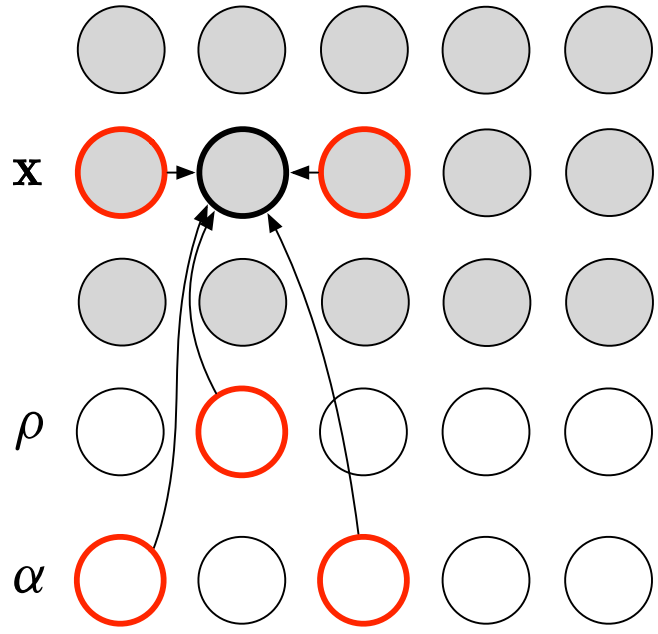
- E.g., an item's embedding (interaction) helps determine its count; its context vector (attributes) helps determine other item's counts

Embedding Structure



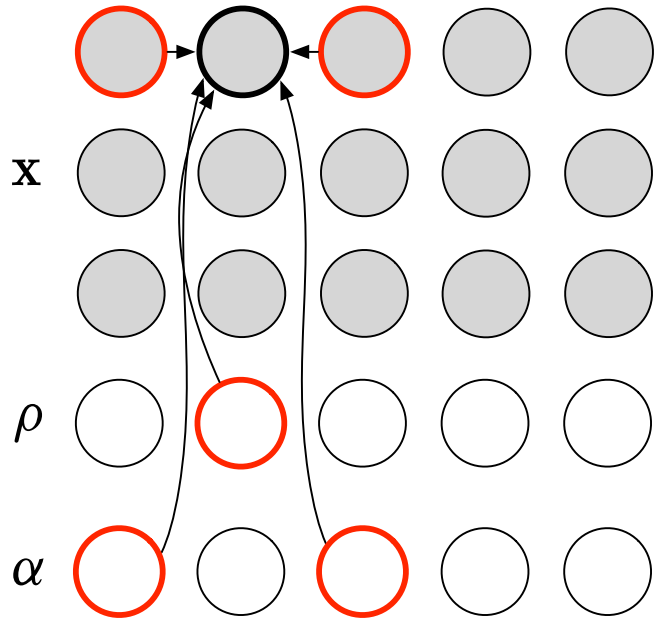
- The embedding structure determines how parameters are shared.
- E.g., $\rho[i] = \rho[j]$ for $i = (\mathrm{Oreos}, t)$ and $j = (\mathrm{Oreos}, u)$.
- Sharing enables learning about an object, such as a neuron, node, or item.

Embedding Structure



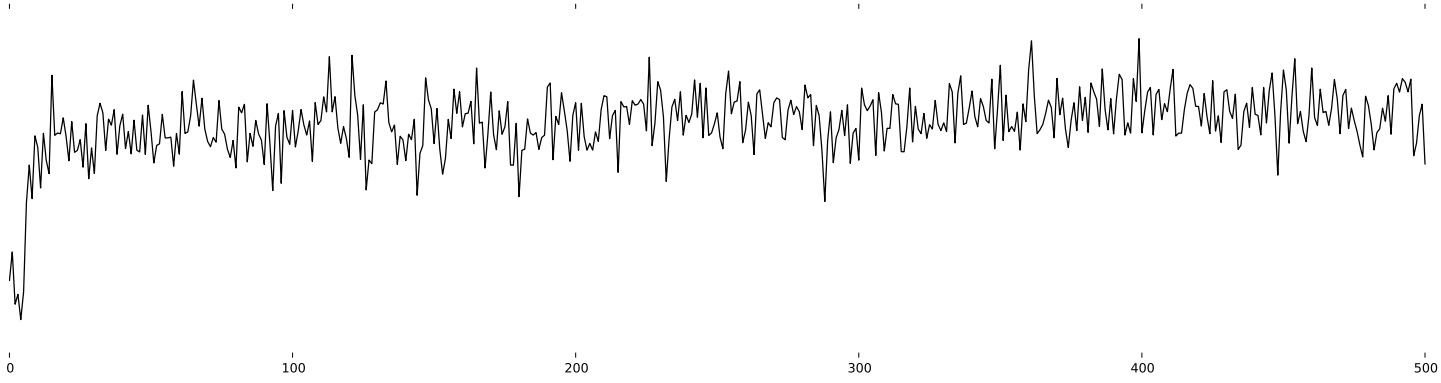
- The embedding structure determines how parameters are shared.
- E.g., $\rho[i] = \rho[j]$ for $i = (\mathrm{Oreos}, t)$ and $j = (\mathrm{Oreos}, u)$.
- Sharing enables learning about an object, such as a neuron, node, or item.

Embedding Structure



- The embedding structure determines how parameters are shared.
- E.g., $\rho[i] = \rho[j]$ for $i = (\mathrm{Oreos}, t)$ and $j = (\mathrm{Oreos}, u)$.
- Sharing enables learning about an object, such as a neuron, node, or item.

Pseudolikelihood

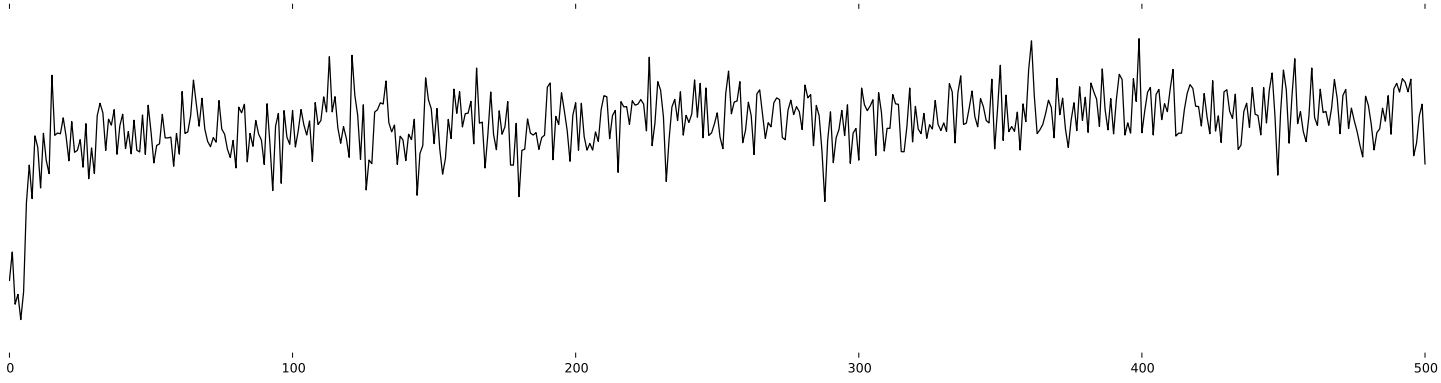


- We model each data point, conditional on the others.
- Combine these ingredients in a “pseudo-likelihood” (i.e., a utility)

$$\mathcal{L}(\boldsymbol{\rho}, \boldsymbol{\alpha}) = \sum_{i=1}^{n} \left(\eta_i^\top t(x_i) - a(\eta_i) \right) + \log f(\boldsymbol{\rho}) + \log g(\boldsymbol{\alpha}).$$

- Fit with stochastic optimization; exponential families simplify the gradients.

Pseudolikelihood



- The objective resembles a collection of GLM likelihoods.
- The gradient is

$$\nabla_{\rho[j]}\mathcal{L} = \sum_{i=1}^{I} \big(t(x_i) - \mathbb{E}[t(x_i)]\big)\nabla_{\rho[j]}\eta_i + \nabla_{\rho[j]} \log f(\rho[j]).$$

- (Stochastic gradients give justification to NN ideas like “negative sampling.”)

Market basket analysis

- Data | Purchase counts of items in shopping trips at a large grocery store
  - Category-level | 478 categories; 635,000 trips; 6.8M purchases
  - Item-level | 5,675 items; 620,000 trips; 5.6M purchases
- Context | Other items purchased at the same trip
- Structure | Embeddings for each item are shared across trips
- Family | Poisson (and we downweight the zeros)

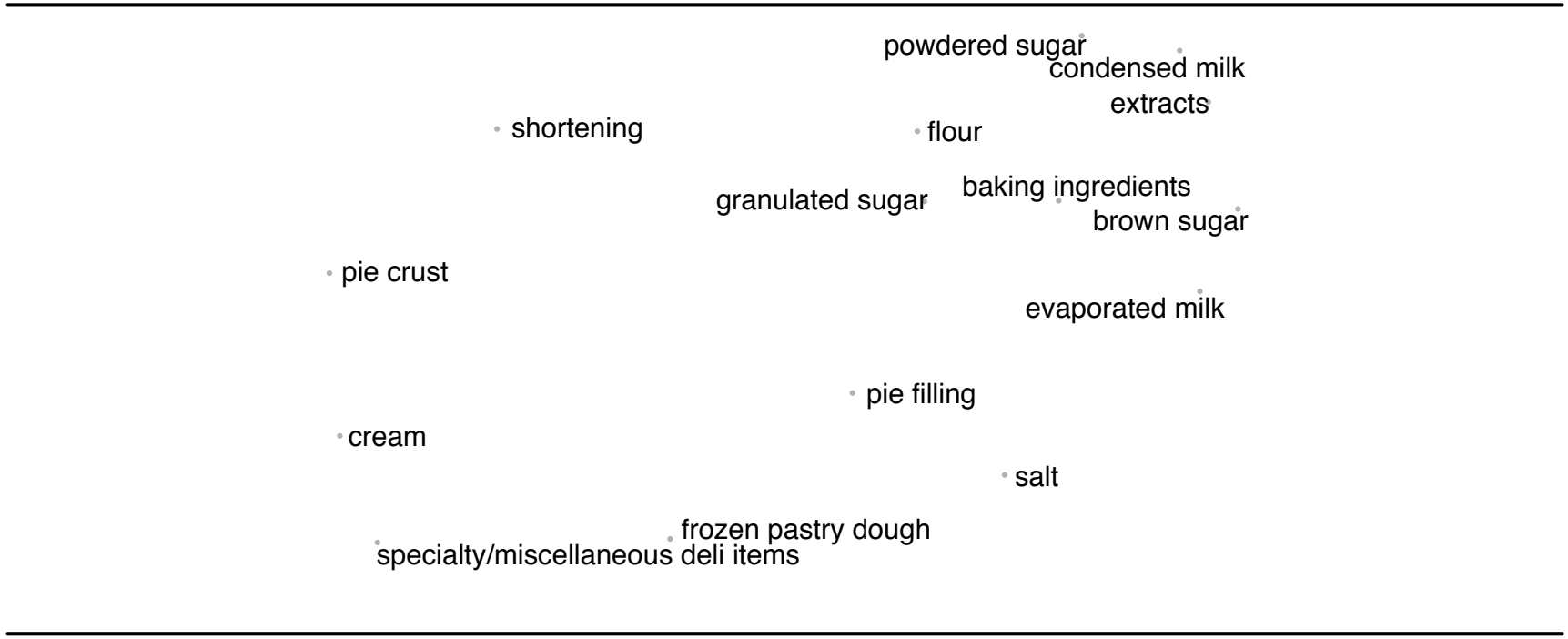
Market basket analysis

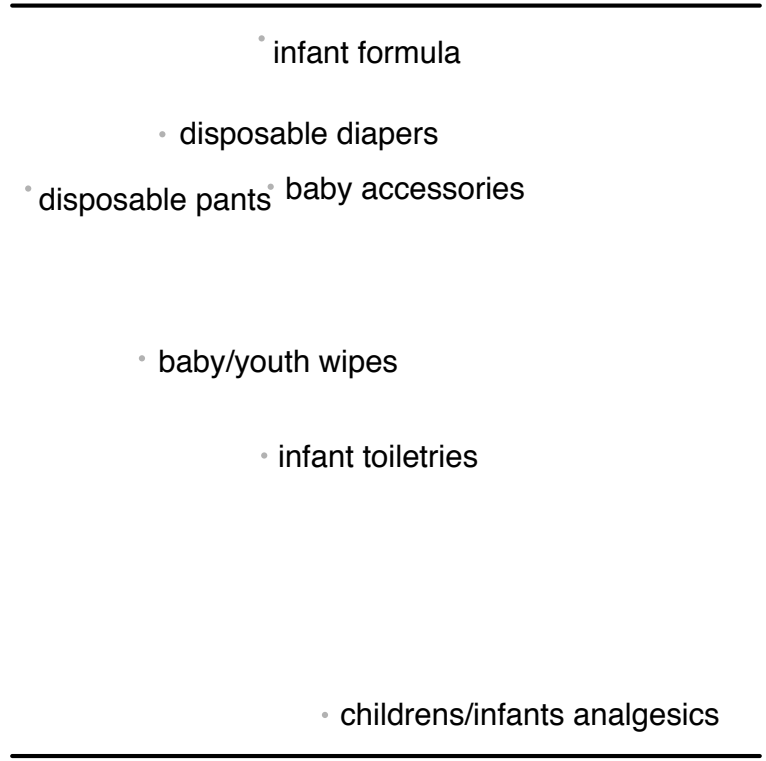
- Recall the conditional probability

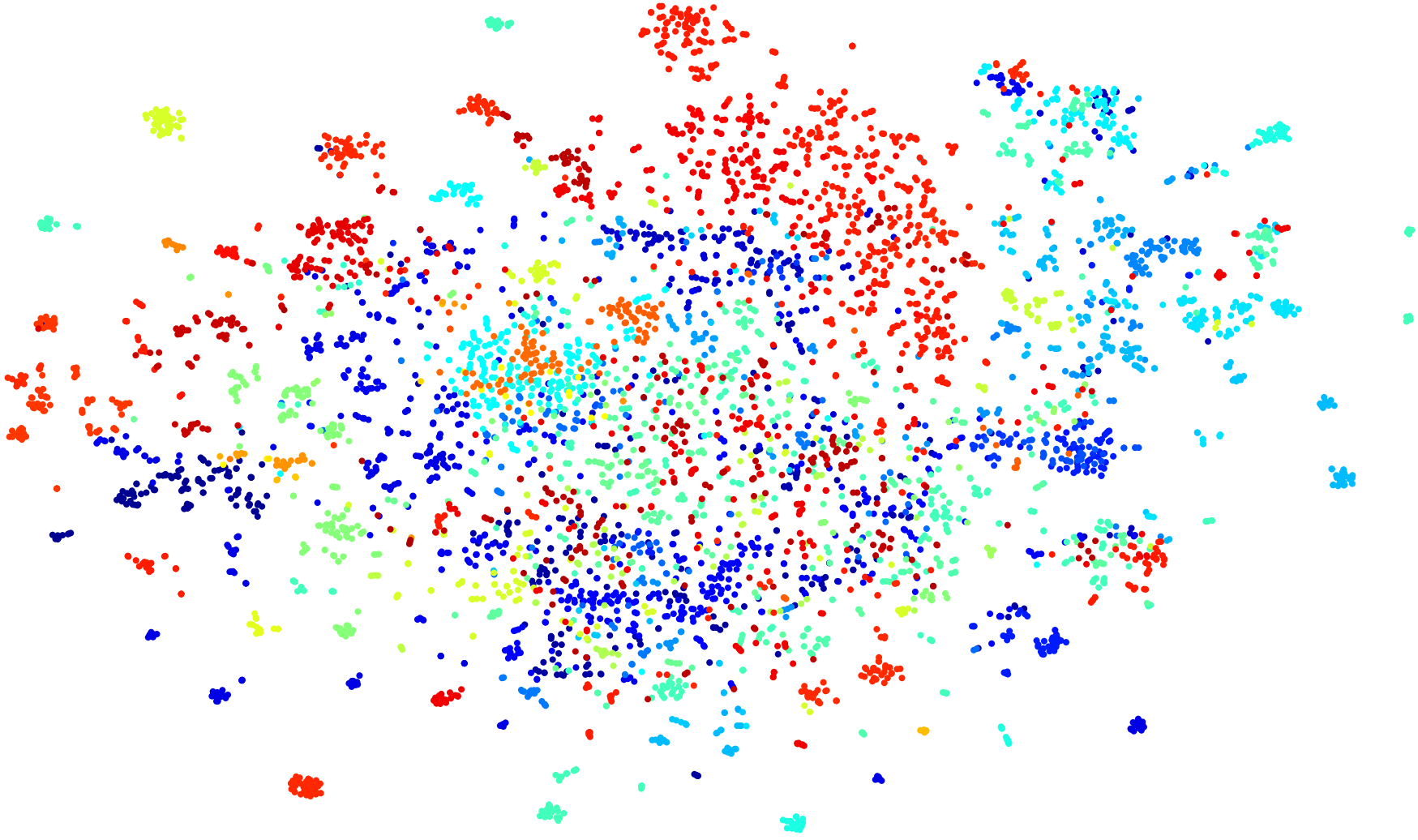
$$x_i \mid \mathbf{x}_{-i} \sim \text{Poisson}\left(\exp\left\{\rho_i^\top \textstyle\sum_{j \neq i} \alpha_j x_j\right\}\right).$$

- α_i reflects attributes of item i
- ρ_i reflects the interaction of item i with attributes of other items.

A 2D representation of category attributes α_i

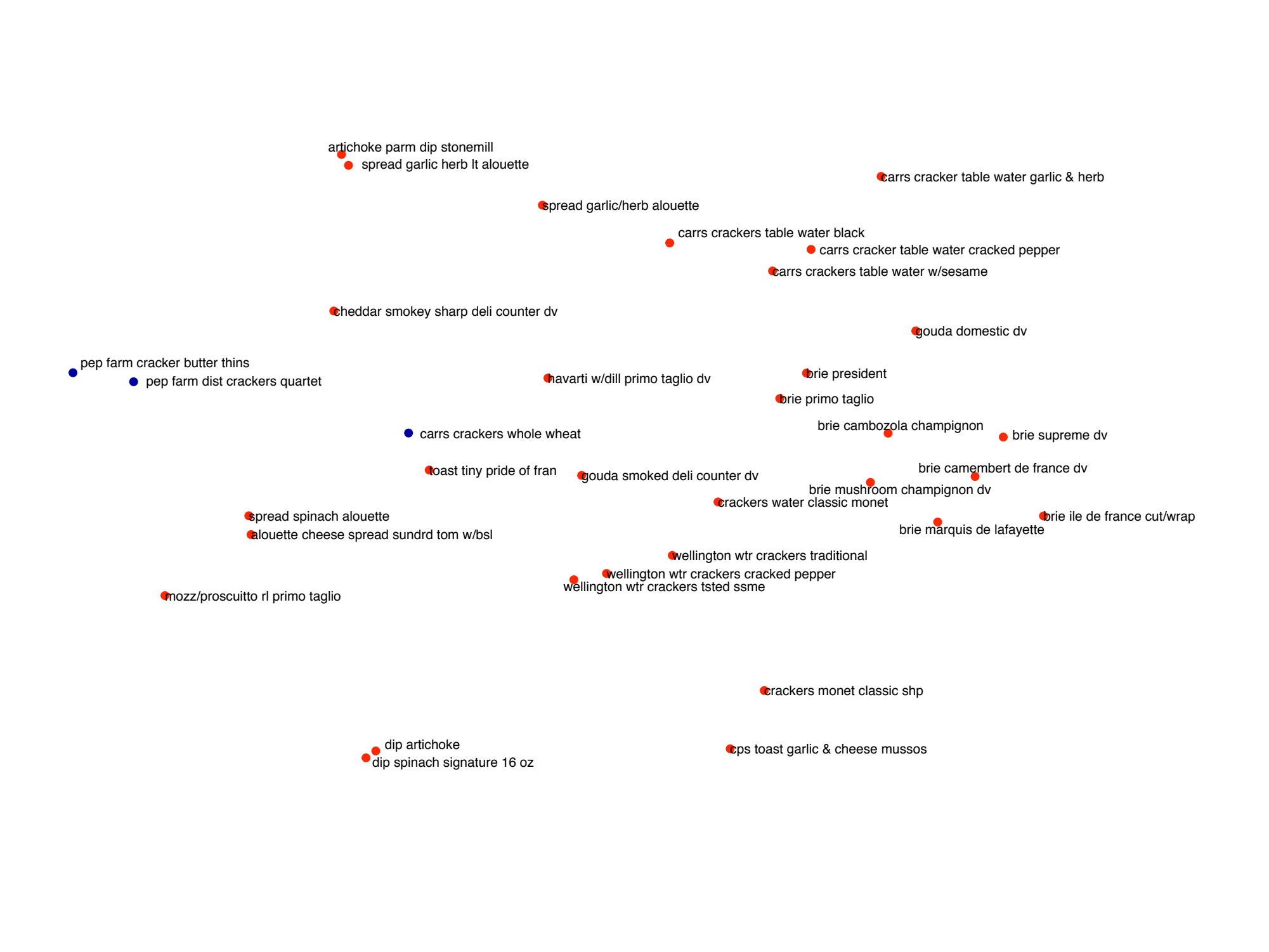

powdered sugar
condensed milk
extracts
shortening
flour
baking ingredients
granulated sugar
brown sugar
pie crust
evaporated milk
pie filling
cream
salt
frozen pastry dough
specialty/miscellaneous deli items

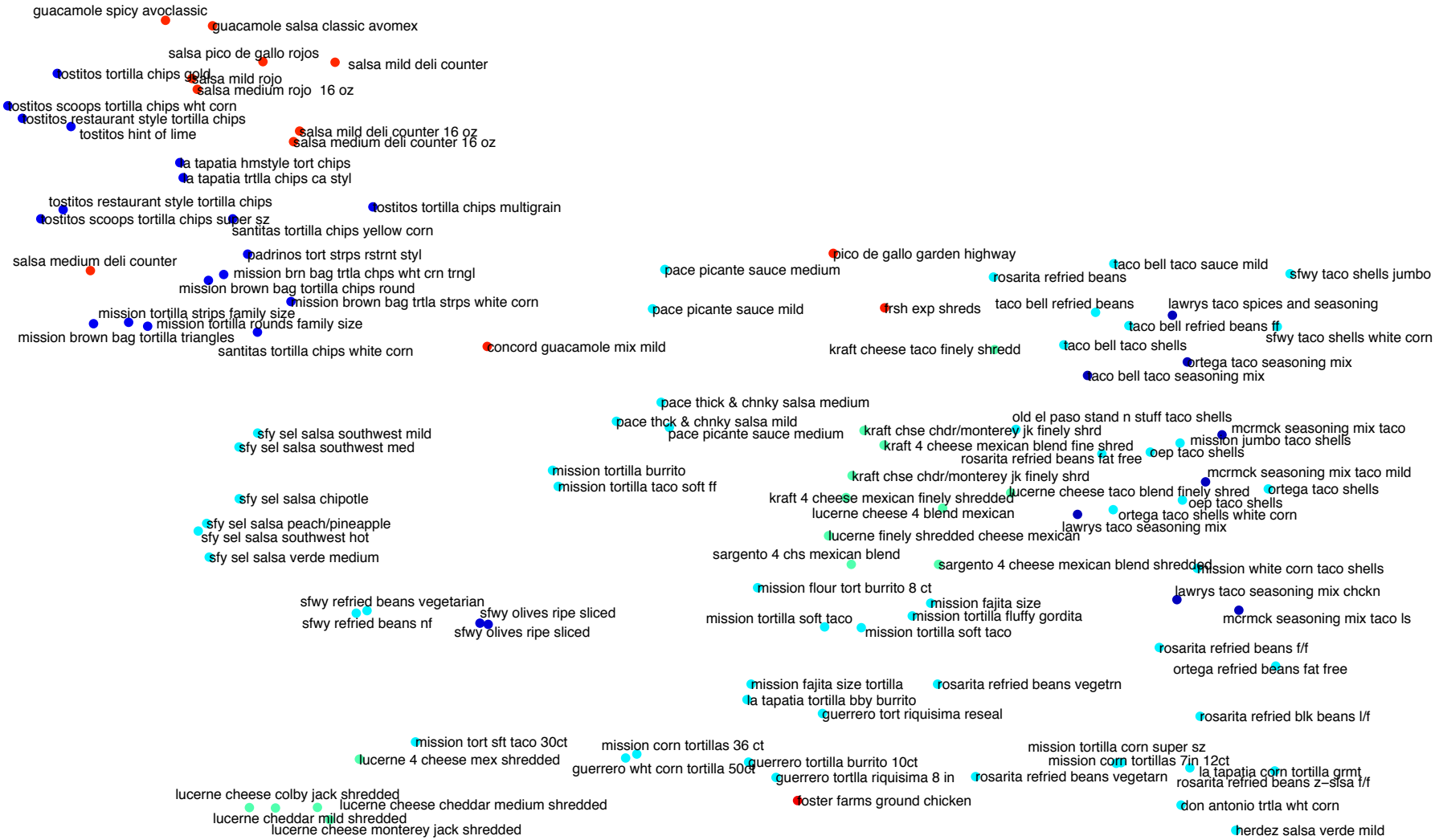

infant formula
disposable diapers
disposable pants
baby accessories
baby/youth wipes
infant toiletries
childrens/infants analgesics



A 2D representation of item attributes α_i


pasta chkn herb parigianna buitoni
ravioli 4 cheese lt buitoni
angel hair buitoni
tortellini ital sausage buitoni
sauce alfredo light buitoni
tortellini chkn/prsct buitoni
sauce pesto basil r/f buitoni
tortellini 3 cheese buitoni
tortellini mixed cheese buitoni
tortellini chse fmly sz buitoni
tortellini herb chck fam sz buitoni
ravioli cheese buitoni
ravioli 4 cheese buitoni
fettucine fresh egg buitoni
linguine fresh egg buitoni
sauce alfredo buitoni
sauce alfredo fmly sz buitoni
sauce marinara fresh buitoni
genova meat sauce
genova ricotta ravioli
genova meat ravioli


artichoke parm dip stonemill
spread garlic herb lt alouette
carrs cracker table water garlic & herb
spread garlic/herb alouette
carrs crackers table water black
carrs cracker table water cracked pepper
carrs crackers table water w/sesame
cheddar smokey sharp deli counter dv
gouda domestic dv
pep farm cracker butter thins
pep farm dist crackers quartet
havarti w/dill primo taglio dv
brie president
brie primo taglio
carrs crackers whole wheat
brie cambozola champignon
brie supreme dv
toast tiny pride of fran
gouda smoked deli counter dv
brie camembert de france dv
brie mushroom champignon dv
crackers water classic monet
spread spinach alouette
brie ile de france cut/wrap
alouette cheese spread sundrd tom w/bsl
brie marquis de lafayette
wellington wtr crackers traditional
wellington wtr crackers cracked pepper
wellington wtr crackers tsted ssme
mozz/proscuitto rl primo taglio
crackers monet classic shp
dip artichoke
dip spinach signature 16 oz
cps toast garlic & cheese mussos


guacamole spicy avoclassic
guacamole salsa classic avomex
salsa pico de gallo rojos
salsa mild deli counter
tostitos tortilla chips gold
salsa mild rojo
salsa medium rojo 16 oz
tostitos scoops tortilla chips wht corn
tostitos restaurant style tortilla chips
tostitos hint of lime
salsa mild deli counter 16 oz
salsa medium deli counter 16 oz
la tapatia hmstyle tort chips
la tapatia trtlla chips ca styl
tostitos restaurant style tortilla chips
tostitos tortilla chips multigrain
tostitos scoops tortilla chips super sz
santitas tortilla chips yellow corn
salsa medium deli counter
padrinos tort strps rstrnt styl
mission brn bag trtla chps wht crn trngl
mission brown bag tortilla chips round
mission brown bag trtla strps white corn
mission tortilla strips family size
mission tortilla rounds family size
mission brown bag tortilla triangles
santitas tortilla chips white corn
concord guacamole mix mild
pico de gallo garden highway
pace picante sauce medium
taco bell taco sauce mild
rosarita refried beans
sfwy taco shells jumbo
pace picante sauce mild
rsh exp shreds
taco bell refried beans
lawrys taco spices and seasoning
taco bell refried beans ff
sfwy taco shells white corn
kraft cheese taco finely shredd
taco bell taco shells
ortega taco seasoning mix
taco bell taco seasoning mix
pace thick & chnky salsa medium
pace thck & chnky salsa mild
pace picante sauce medium
old el paso stand n stuff taco shells
kraft chse chdr/monterey jk finely shrd
mcrmck seasoning mix taco
kraft 4 cheese mexican blend fine shred
mission jumbo taco shells
rosarita refried beans fat free
oep taco shells
sfy sel salsa southwest mild
sfy sel salsa southwest med
mission tortilla burrito
mission tortilla taco soft ff
kraft chse chdr/monterey jk finely shrd
mcrmck seasoning mix taco mild
lucerne cheese taco blend finely shred
ortega taco shells
sfy sel salsa chipotle
kraft 4 cheese mexican finely shredded
oep taco shells
lucerne cheese 4 blend mexican
ortega taco shells white corn
lawrys taco seasoning mix
sfy sel salsa peach/pineapple
sfy sel salsa southwest hot
lucerne finely shredded cheese mexican
sfy sel salsa verde medium
sargento 4 chs mexican blend
sargento 4 cheese mexican blend shredded
mission white corn taco shells
mission flour tort burrito 8 ct
lawrys taco seasoning mix chckn
sfwy refried beans vegetarian
mission fajita size
sfwy olives ripe sliced
mission tortilla soft taco
mission tortilla fluffy gordita
mcrmck seasoning mix taco ls
sfwy refried beans nf
sfwy olives ripe sliced
mission tortilla soft taco
rosarita refried beans f/f
ortega refried beans fat free
mission fajita size tortilla
rosarita refried beans vegetrn
la tapatia tortilla bby burrito
guerrero tort riquisima reseal
rosarita refried blk beans l/f
mission tort sft taco 30ct
mission corn tortillas 36 ct
mission tortilla corn super sz
lucerne 4 cheese mex shredded
guerrero wht corn tortilla 50ct
guerrero tortilla burrito 10ct
mission corn tortillas 7in 12ct
la tapatia corn tortilla grmt
guerrero tortlla riquisima 8 in
rosarita refried beans vegetarn
rosarita refried beans z-slsa f/f
lucerne cheese colby jack shredded
lucerne cheese cheddar medium shredded
foster farms ground chicken
don antonio trtla wht corn
lucerne cheddar mild shredded
lucerne cheese monterey jack shredded
herdez salsa verde mild

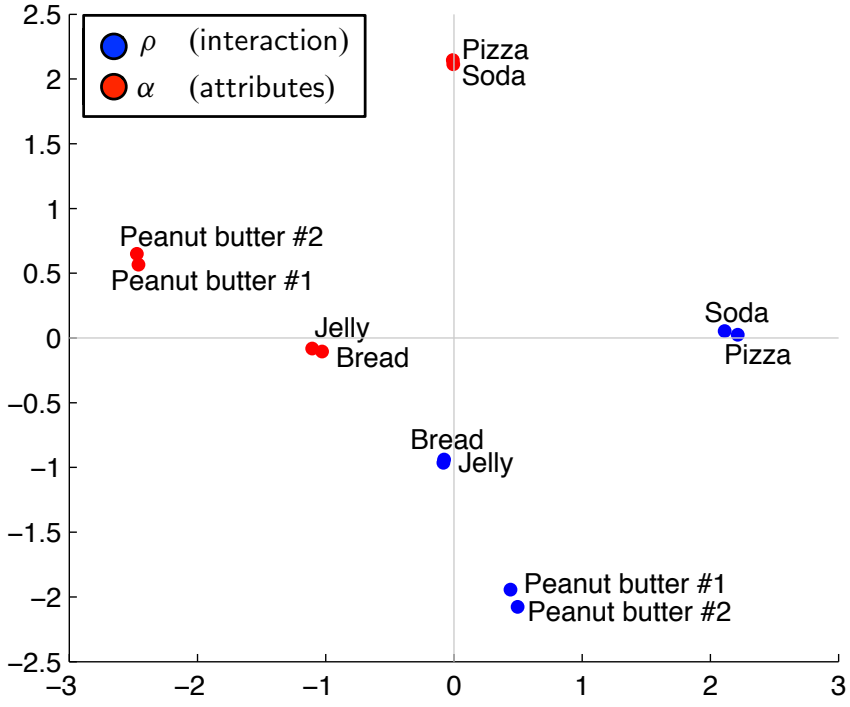
Category level

MODEL	$K = 20$	$K = 50$	$K = 100$
Poisson embedding	-7.497	-7.284	-7.199
Poisson embedding (downweighting zeros)	-7.110	-6.994	$\mathbf{-6.950}$
Additive Poisson embedding	-7.868	-8.191	-8.414
Hierarchical Poisson factorization	-7.740	-7.626	-7.626
Poisson PCA	-8.314	-9.51	-11.01

Item level

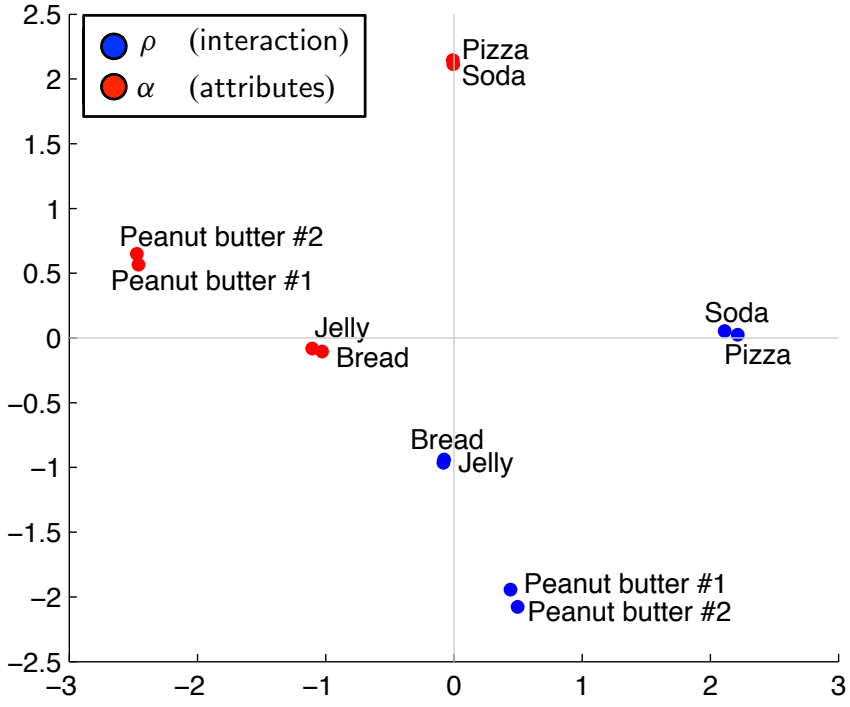
MODEL	K=50	K=100
Poisson embedding	-7.72	$\mathbf{-7.64}$
Hierarchical Poisson factorization	-7.86	-7.87

Gopalan et al., Scalable recommendation with hierarchical Poisson factorization. *Uncertainty in Artificial Intelligence*, 2015.
Collins et al., A generalization of principal component analysis to the exponential family. *Neural Information Processing Systems*, 2002.



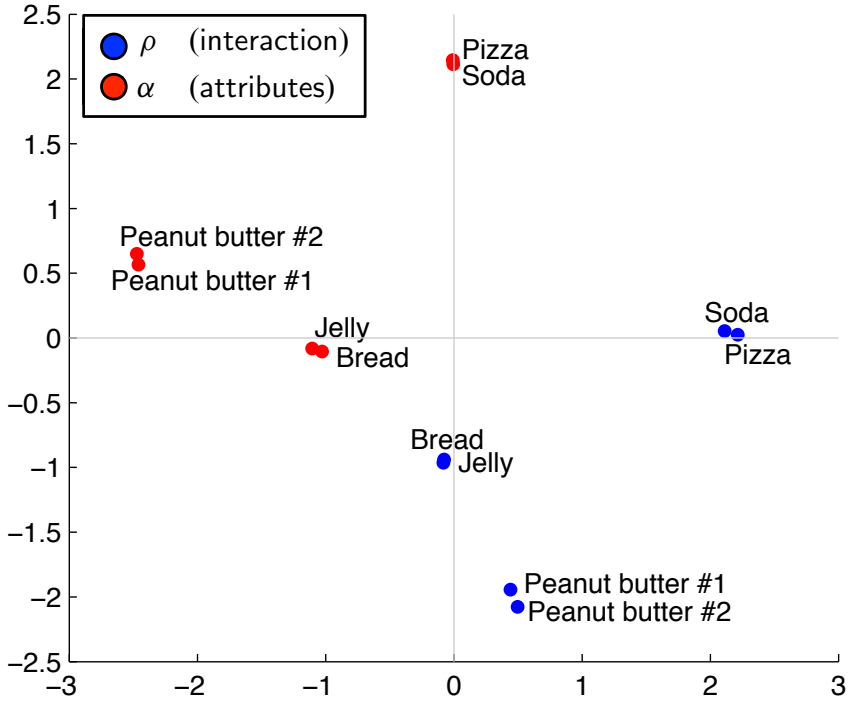
We want to use this fit to understand purchase patterns.

- *Exchangeables* have a similar effect on the purchase of other items.
- *Same-category items* tend to be exchangeable and rarely purchased together (e.g., two types of peanut butter).
- *Complements* are purchased (or not purchased) together (e.g., hot dogs and buns).



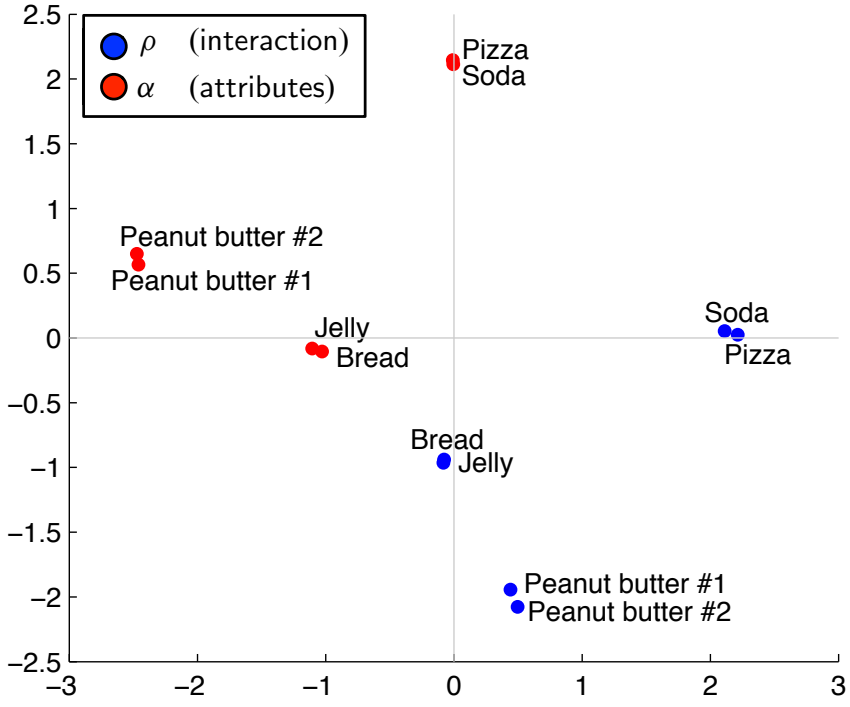
- PB #1 and PB #2 induce similar distributions of other items
- But they are rarely purchased together.
- Define the sigmoid function between two items,

$$\sigma_{ki} \triangleq \frac{1}{\left(1+\exp\{-\rho_k^\top \alpha_i\}\right)} \quad ; \quad \bar{\sigma}_{ki} \triangleq 1-\sigma_{ki}$$



- The "substitute predictor" is

$$-\left(\sum_{k/\{i,j\}} \sigma_{ki} \log\left(\frac{\sigma_{ki}}{\sigma_{kj}}\right) + \bar{\sigma}_{ki} \log\left(\frac{\bar{\sigma}_{ki}}{\bar{\sigma}_{kj}}\right)\right) - \sigma_{ji} \log\left(\frac{\sigma_{ji}}{1-\sigma_{ji}}\right).$$



- The "complement predictor" is the negative of the last term

$$\sigma_{ji} \log \left(\frac{\sigma_{ji}}{1 - \sigma_{ji}} \right).$$

- Notes:
  - We use the symmetrized version of both quantities
  - These quantities generalize to other exponential families

ITEM 1	ITEM 2	SCORE (RANK)
organic vegetables	organic fruits	6.18 (01)
vegetables (<10 oz)	beets (>=10 oz)	5.64 (02)
baby food	disposable diapers	3.43 (32)
stuffing	cranberries	3.30 (36)
gravy	stuffing	3.23 (37)
pie filling	evaporated milk	3.09 (42)
deli cheese	deli crackers	2.87 (55)
dry pasta/noodles	tomato paste/sauce/puree	2.73 (63)
mayonnaise	mustard	2.61 (69)
cake mixes	frosting	2.49 (78)

Example co-purchases at the category level

ITEM 1	ITEM 2	SCORE (RANK)
bouquets	roses	0.20 (01)
frozen pizza 1	frozen pizza 2	0.18 (02)
bottled water 1	bottled water 2	-0.07 (03)
carbonated soft drinks 1	carbonated soft drinks 2	-0.12 (04)
orange juice 1	orange juice 2	-0.37 (05)
bathroom tissue 1	bathroom tissue 2	-0.58 (06)
bananas 1	bananas 2	-0.61 (07)
salads-convenience 1	salads-convenience 2	-0.63 (08)
potatoes 1	potatoes 2	-0.66 (09)
bouquets	blooming	-1.18 (10)

Top ten potential substitutes at the category level

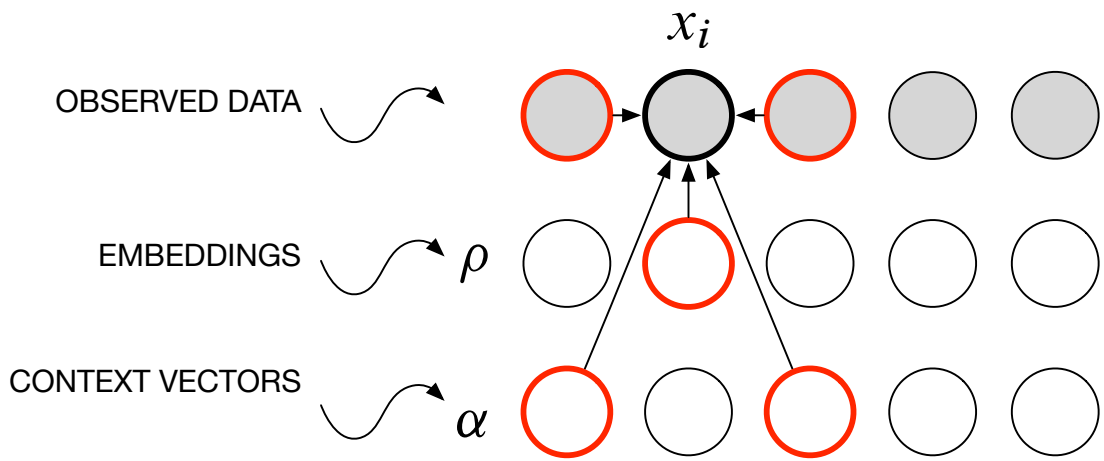
ITEM 1	ITEM 2	SCORE (RANK)
ygrt peach ff	ygrt mxd berry ff	19.83 (0001)
s&w beans garbanzo	s&w beans red kidney	14.42 (0002)
whiskas cat fd beef	whiskas cat food tuna/chicken	8.45 (0149)
parsnips loose	rutabagas	8.32 (0157)
celery hearts organic	apples fuji organic	4.36 (0995)
85p ln gr beef patties 15p fat	sesame buns	4.35 (1005)
kiwi imported	mangos small	3.22 (1959)
colby jack shredded	taco bell taco seasoning mix	2.89 (2472)
star magazine	in touch magazine	2.87 (2497)
seasoning mix fajita	mission tortilla corn super sz	2.87 (2500)

Example co-purchases at the UPC level

ITEM 1	ITEM 2	SCORE (RANK)
coffee drip grande	coffee drip venti	-0 .33 (001)
sandwich signature reg	sandwich signature lrg	-1.17 (020)
market bouquet	alstromeria/rose bouquet	-2.89 (186)
sushi shoreline combo	sushi full moon combo	-3.76 (282)
semifreddis bread baguette	crusty sweet baguette	-7.65 (566)
orbit gum peppermint	orbit gum spearmint	-7.96 (595)
snickers candy bar	3 musketeers candy bar	-7.97 (598)
cheer lndry det color guard	all lndry det liquid fresh rain	-7.99 (602)
coors light beer btl	coors light beer can	-8.12 (621)
greek salad signature	neptune salad signature	-8.15 (630)

Example potential substitutes at the UPC level

Summary and Questions

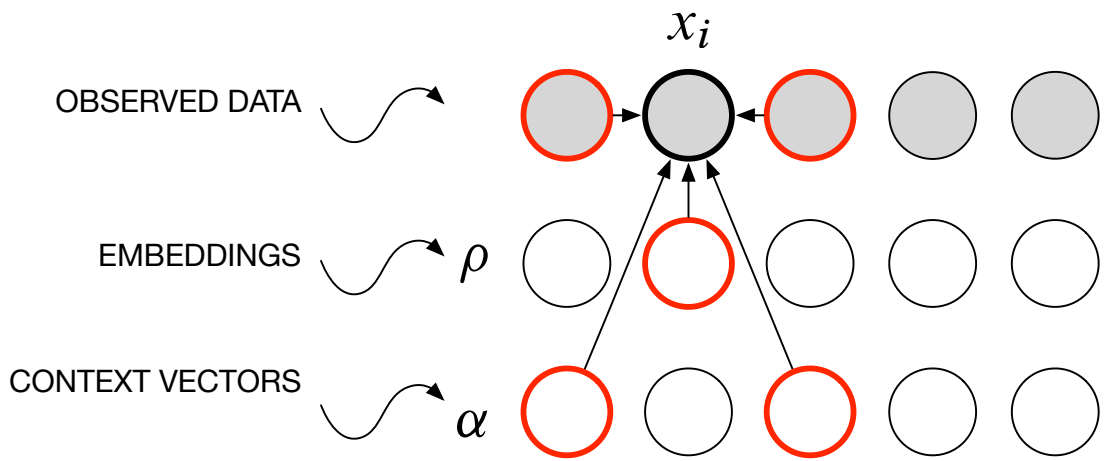


- Word embeddings have become a staple in natural language processing
  We distilled its essential elements, generalized to consumer data
- Compared to classical factorization, good performance in many data
  - movie ratings, neural activity, scientific reading, shopping baskets

---

Rudolph et al. Exponential Family Embeddings. *Neural Information Processing Systems*, 2016.
Liang et al. Modeling User Exposure in Recommendation. *Recommendation Systems*, 2016.
Ranganath et al. Deep Exponential Families. *Artificial Intelligence and Statistics*, 2015.

Summary and Questions



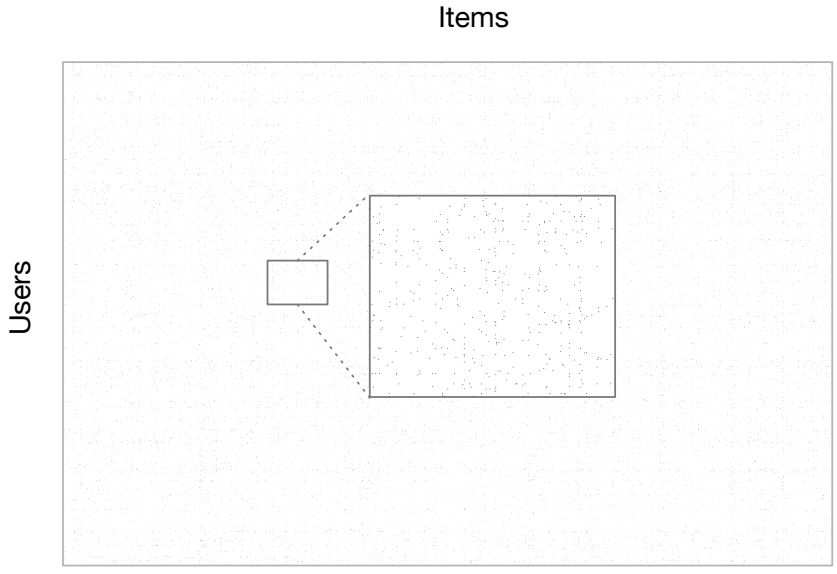
- How can we capture higher-order structure in the embeddings?
- Why downweight the zeros?
- How can we include price and other complexities?

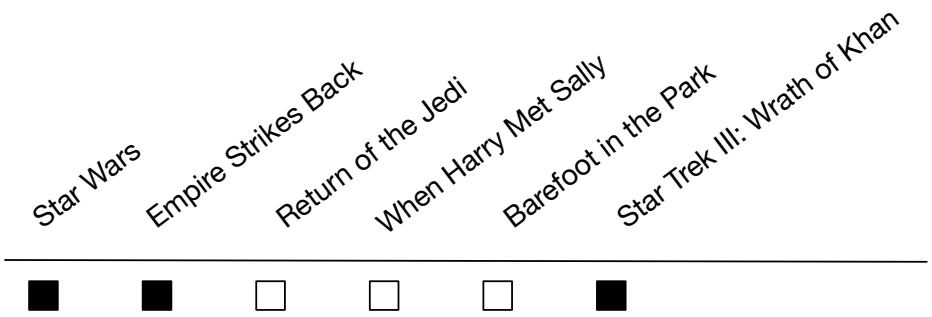
Rudolph et al. Exponential Family Embeddings. *Neural Information Processing Systems*, 2016.
Liang et al. Modeling User Exposure in Recommendation. *Recommendation Systems*, 2016.
Ranganath et al. Deep Exponential Families. *Artificial Intelligence and Statistics*, 2015.

Poisson factorization

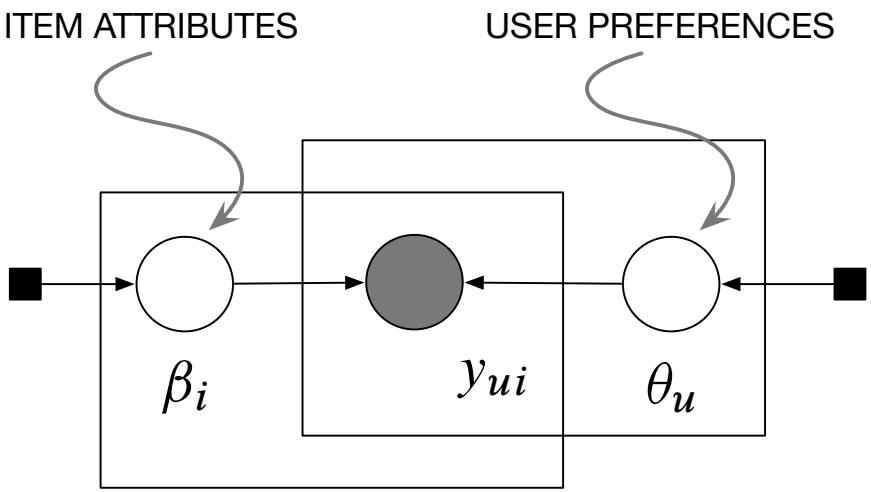
A computationally efficient method for discovering correlated preferences

- *Economics*
  - Look at items within one category (e.g. yoghurt)
  - Try to estimate the effects of interventions (e.g., coupons, price, layout)
- *Machine learning*
  - Look at all items
  - Estimate user preferences and make predictions (recommendations)
  - Ignore causal effects of interventions





- *Implicit data* is about users interacting with items
  - clicks
  - “likes”
  - purchases
- Less information than explicit data (e.g. ratings), but more prevalent

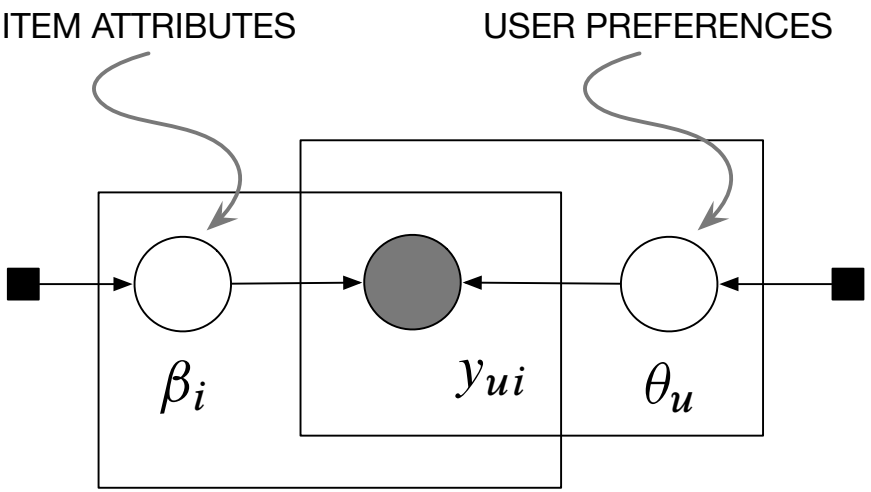


$$\theta_{uk} \sim \text{Gam}(\cdot, \cdot)$$
$$\beta_{ik} \sim \text{Gam}(\cdot, \cdot)$$
$$y_{ui} \sim \text{Poisson}\left(\theta_u^\top \beta_i\right)$$

Poisson factorization

- Assumptions
  - Users (consumers) have *latent preferences* θ_u.
  - Items have *latent attributes* β_i.
  - How many items a shopper purchased comes from a Poisson.
- The posterior $p(\boldsymbol{\theta}, \boldsymbol{\beta} \mid \mathbf{y})$ reveals purchase patterns.

Gopalan et al., Scalable recommendation with hierarchical Poisson factorization. *Uncertainty in Artificial Intelligence*, 2015.



$$\theta_{uk} \sim \text{Gam}(\cdot, \cdot)$$
$$\beta_{ik} \sim \text{Gam}(\cdot, \cdot)$$
$$y_{ui} \sim \text{Poisson}\left(\theta_u^\top \beta_i\right)$$

Advantages

- captures heterogeneity of users
- implies a distribution of total consumption
- efficient approximation, only requires non-zero data

Gopalan et al., Scalable recommendation with hierarchical Poisson factorization. *Uncertainty in Artificial Intelligence*, 2015.

News articles from the *New York Times*

“Business Self-Help”

Stay Focused And Your Career Will Manage Itself
To Tear Down Walls You Have to Move Out of Your Office
Self-Reliance Learned Early
Maybe Management Isn't Your Style
My Copyright Career

“Personal Finance”

In Hard Economy for All Ages Older Isn't Better It's Brutal
Younger Generations Lag Parents in Wealth-Building
Fast-Growing Brokerage Firm Often Tangles With Regulators
The Five Stages of Retirement Planning Angst
Signs That It's Time for a New Broker

“All Things Airplane”

Flying Solo
Crew-Only 787 Flight Is Approved By FAA
All Aboard Rescued After Plane Skids Into Water at Bali Airport
Investigators Begin to Test Other Parts On the 787
American and US Airways May Announce a Merger This Week

Scientific articles from Mendeley

“Astronomy”

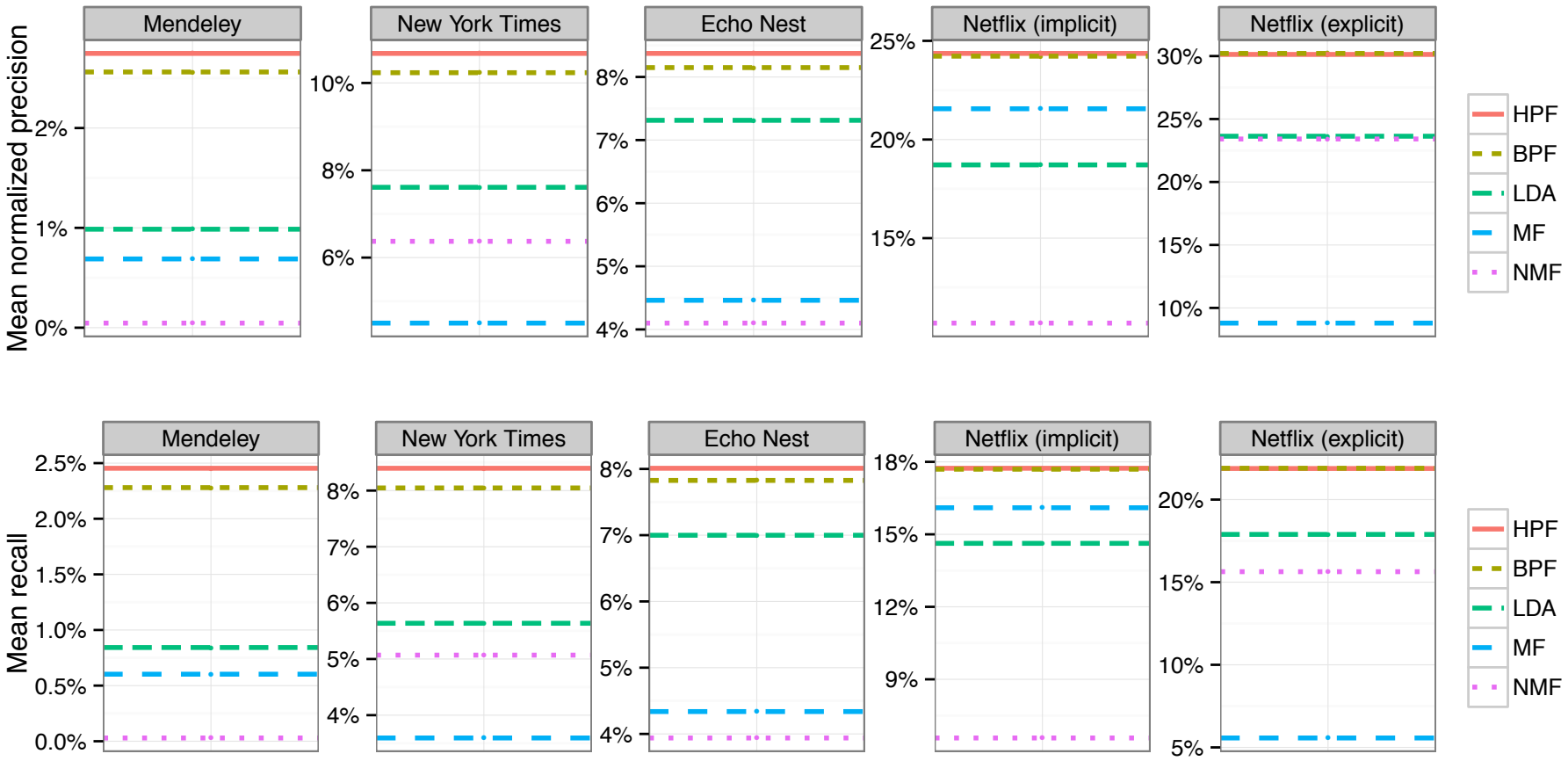
Theory of Star Formation
Error estimation in astronomy: A guide
Astronomy & Astrophysics
Measurements of Omega from 42 High-Redshift Supernovae
Stellar population synthesis at the resolution of 2003

“Biodiesel”

Biodiesel from microalgae.
Biodiesel from microalgae beats bioethanol
Commercial applications of microalgae
Second Generation Biofuels
Hydrolysis of lignocellulosic materials for ethanol production

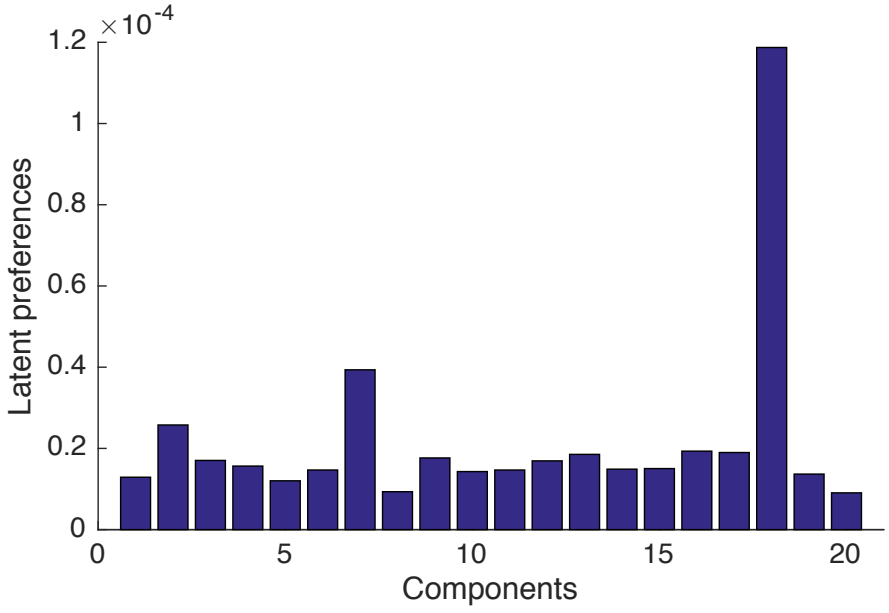
“Political Science”

Social Capital: Origins and Applications in Modern Sociology
Increasing Returns, Path Dependence, and Politics
Institutions, Institutional Change and Economic Performance
Diplomacy and Domestic Politics
Comparative Politics and the Comparative Method

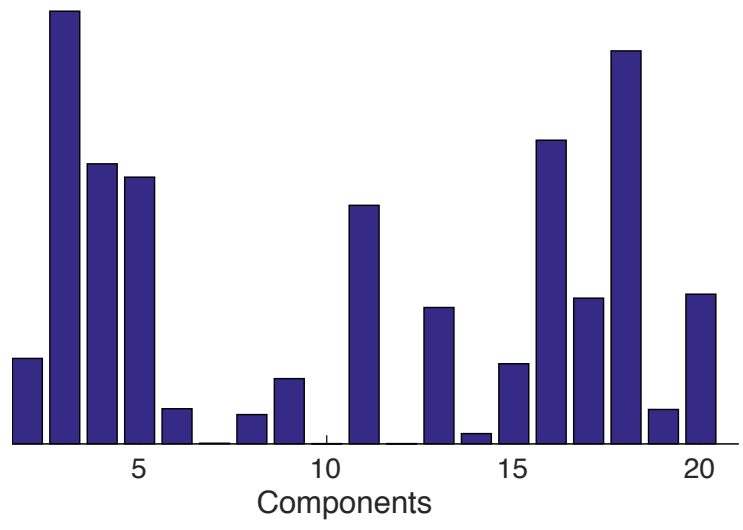

Mean normalized precision
Mendeley
New York Times
Echo Nest
Netflix (implicit)
Netflix (explicit)
0%
1%
2%
6%
8%
10%
4%
5%
6%
7%
8%
15%
20%
25%
10%
15%
20%
25%
30%
HPF
BPF
LDA
MF
NMF
Mean recall
Mendeley
New York Times
Echo Nest
Netflix (implicit)
Netflix (explicit)
0.0%
0.5%
1.0%
1.5%
2.0%
2.5%
4%
5%
6%
7%
8%
4%
5%
6%
7%
8%
9%
12%
15%
18%
5%
10%
15%
20%
HPF
BPF
LDA
MF
NMF

“FRUIT”	**“CAT CARE”**	**“BABY ESSENTIALS”**	**“HEALTHY”**
stone fruit	cat food wet	baby food	health and milk substitutes
pears	cat food dry	starbucks coffee	organic vegetables
tropical fruit	cat litter & deodorant	disposable diapers	organic fruits
apples	canned fish	infant formula	cold cereal
grapes	paper towels	baby/youth wipes	vegetarian / organic frozen

Consumer #1 : “Cats and Babies”



Consumer #2 : “Healthy and Cats”



Poisson factorization and economics

- Consider a utility model of a single purchase with Gumbel error,

$$U(y_{ui}) = \log(\theta_u^\top \beta_i) + \epsilon.$$

- Suppose a shopper u buys N items. Then

$$y_u \mid N \sim \text{Multi}(N, \pi_u)$$
$$\pi_{ui} \propto \exp\{\theta_u^\top \beta_i\}.$$

- Thus, the unconditional distribution of counts is Poisson factorization,

$$y_{uj} \sim \text{Poisson}(\theta_u^\top \beta_j).$$

Poisson factorization and economics

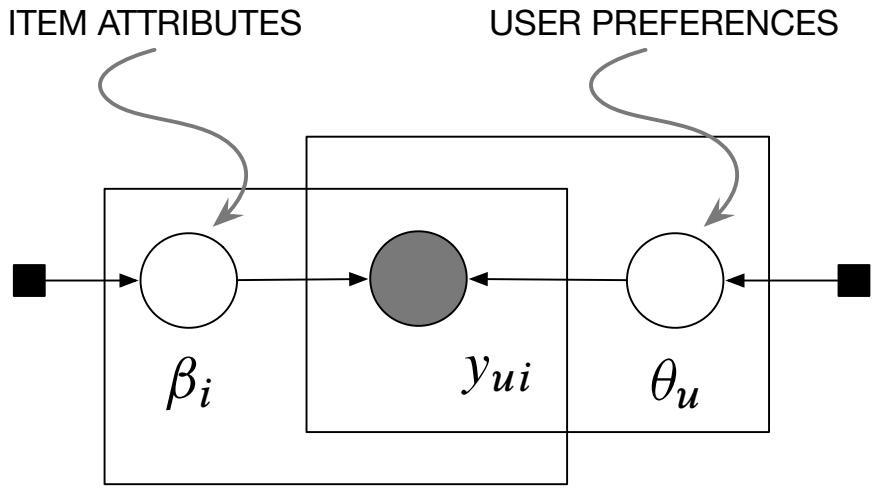
- With this connection, we can devise new utility models, e.g.,

$$U(y_{ui}) = \log(\theta_u^\top \beta_i + \alpha_u \exp\{-c \cdot \text{price}_i) + \epsilon.$$

- ...and other factors
  - time of day
  - in stock
  - date
  - observed item characteristics & category
  - demographic information about the shopper

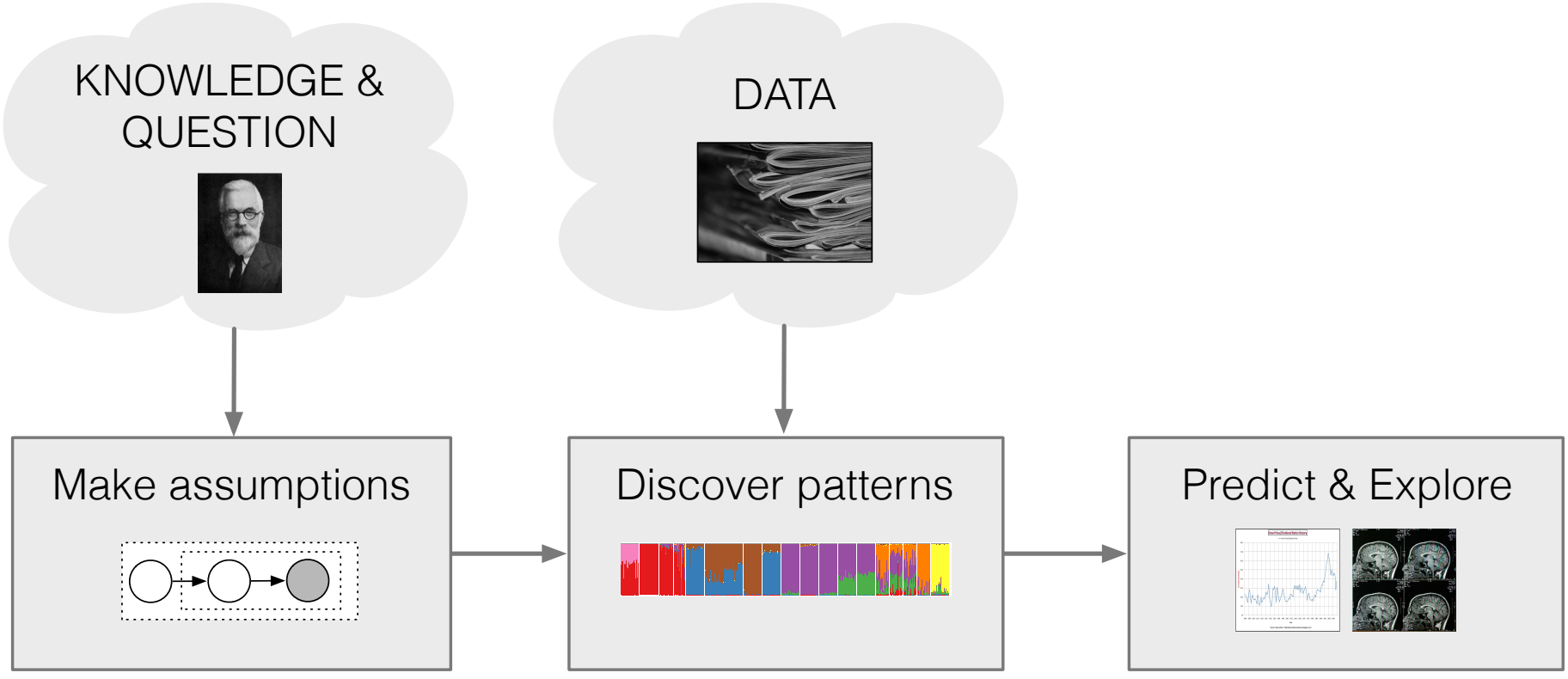
- Inference is still efficient.

  With assumptions, we can answer counterfactual questions.



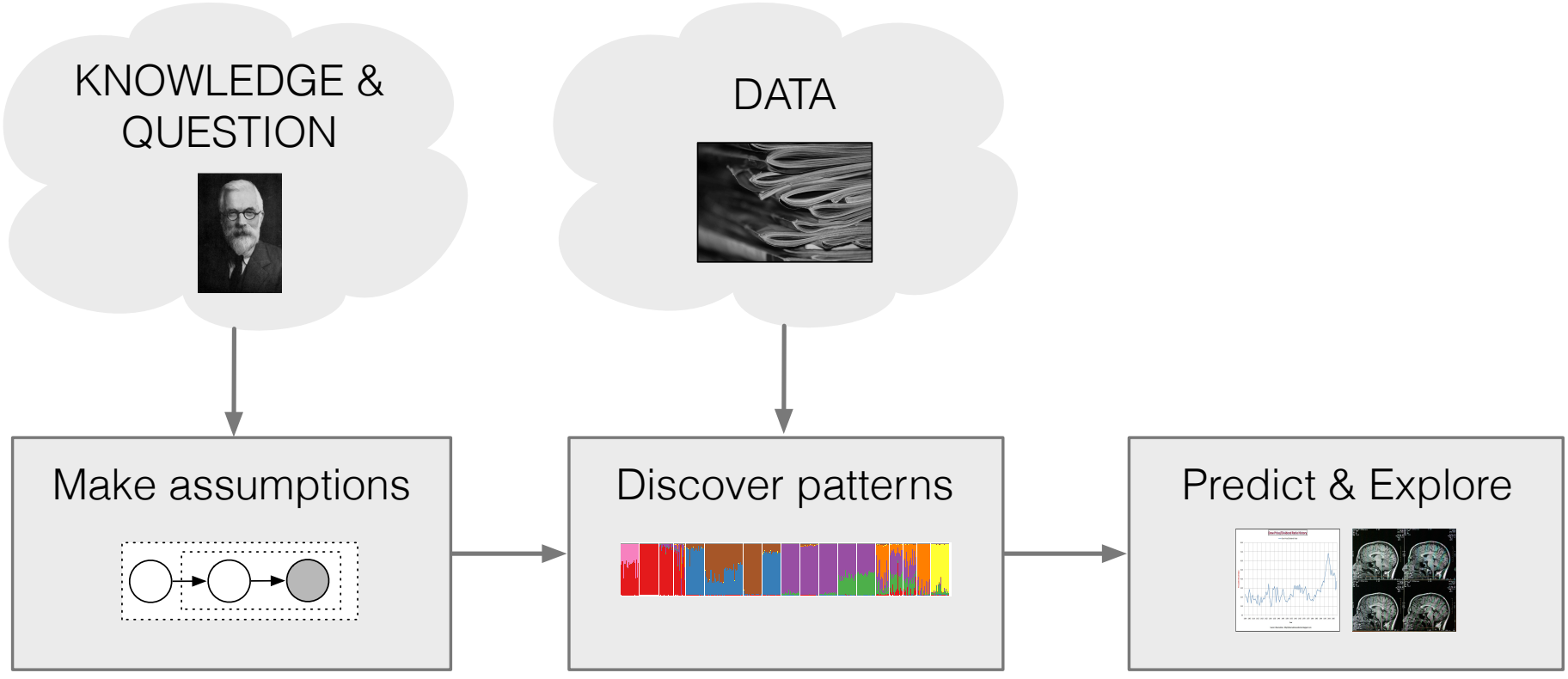
$$\theta_{uk} \sim \text{Gam}(\cdot, \cdot)$$
$$\beta_{ik} \sim \text{Gam}(\cdot, \cdot)$$
$$y_{ui} \sim \text{Poisson}\left(\theta_u^\top \beta_i\right)$$

- Poisson factorization efficiently analyzes large-scale purchase behavior.
- Next steps
  - include notions of co-purchases and substitutes
  - include time of day at a level that is unconfounded
  - include price and stock out; answer counterfactual questions
- Research in recommendation systems can help economic analyses.



Probabilistic machine learning: design expressive models to analyze data.

- Tailor your method to your question and knowledge
- Use generic and scalable inference to analyze large data sets
- Form predictions, hypotheses, inferences, and revise the model



Opportunities for economics and machine learning

- Push economics to high-dimensional data and scalable computation
- Push ML to explainable models, applied causal inference, new problems
- Develop new modeling methods together