Experience in Upper and Middle Sales Funnel Optimization Using Lift Modeling and Embeddings

Wai Gen Yee

May 1, 2021

International Conference on Data Mining

Data Science for Retail and E-Commerce Workshop, 2021

About...

- Links together restaurants, diners, and drivers
- Want to be the marketer go-to for restaurants
- 140,000+ restaurants in over 2,700 cities
- >20MM active diners
- >$5B annual food sales
- Other brands: Seamless, LevelUp, Tapingo,
- AllMenus and MenuPages

About...

ILLINOIS TECH

ORBITZ

GRUBHUB

Restaurants you love, delivered.

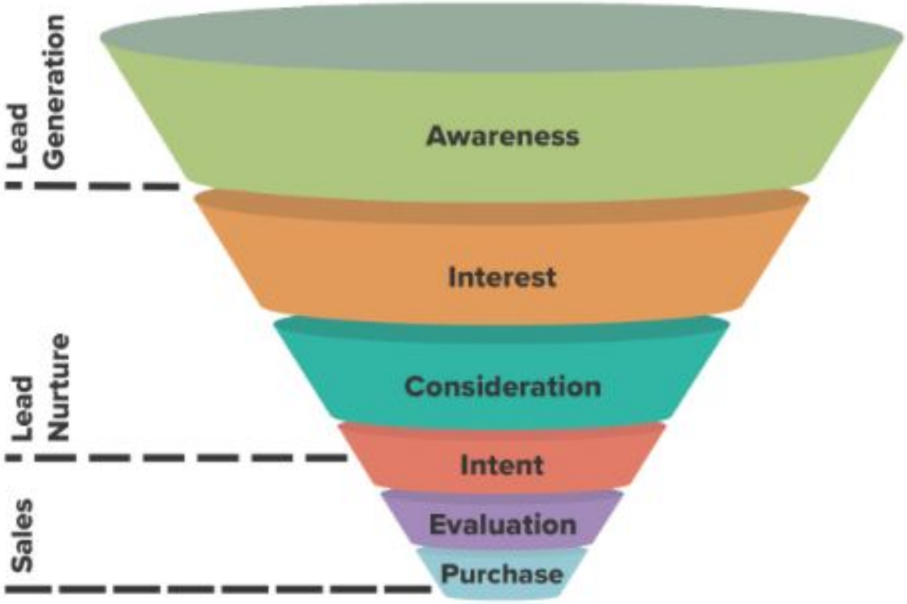
sears

Tripadvisor

Georgia Tech

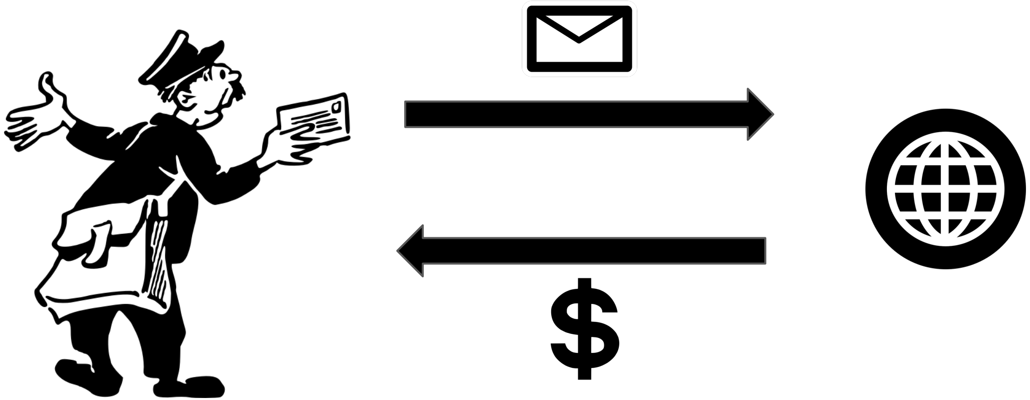
The Marketing Funnel

TrackMaven

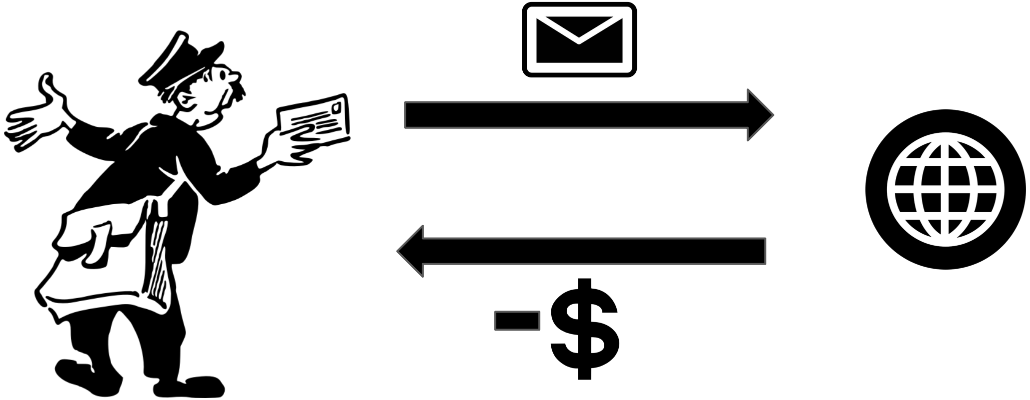


- Marketing campaigns and consumer research
- Events, advertising, tradeshows, blog, webinars, direct mail, viral campaigns, social media, search, media mentions, and more
- Engagement and introduction to positioning
- Emails, targeted content, classes, newsletters, and more
- Product information and specials
- Automated email campaigns, case studies, free trials, and more
- Product demos and shopping carts
- Sales adds to nurture stream
- Marketing and sales work to prove their product is best
- Sales transaction is completed

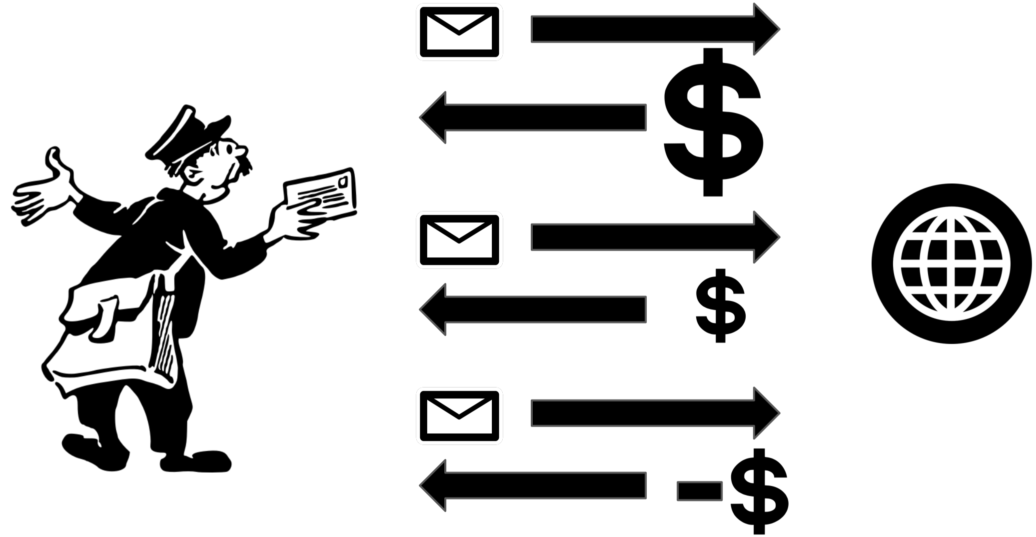
Assumption: Make Money, Scale



Assumption: Lose Money, Don’t Scale



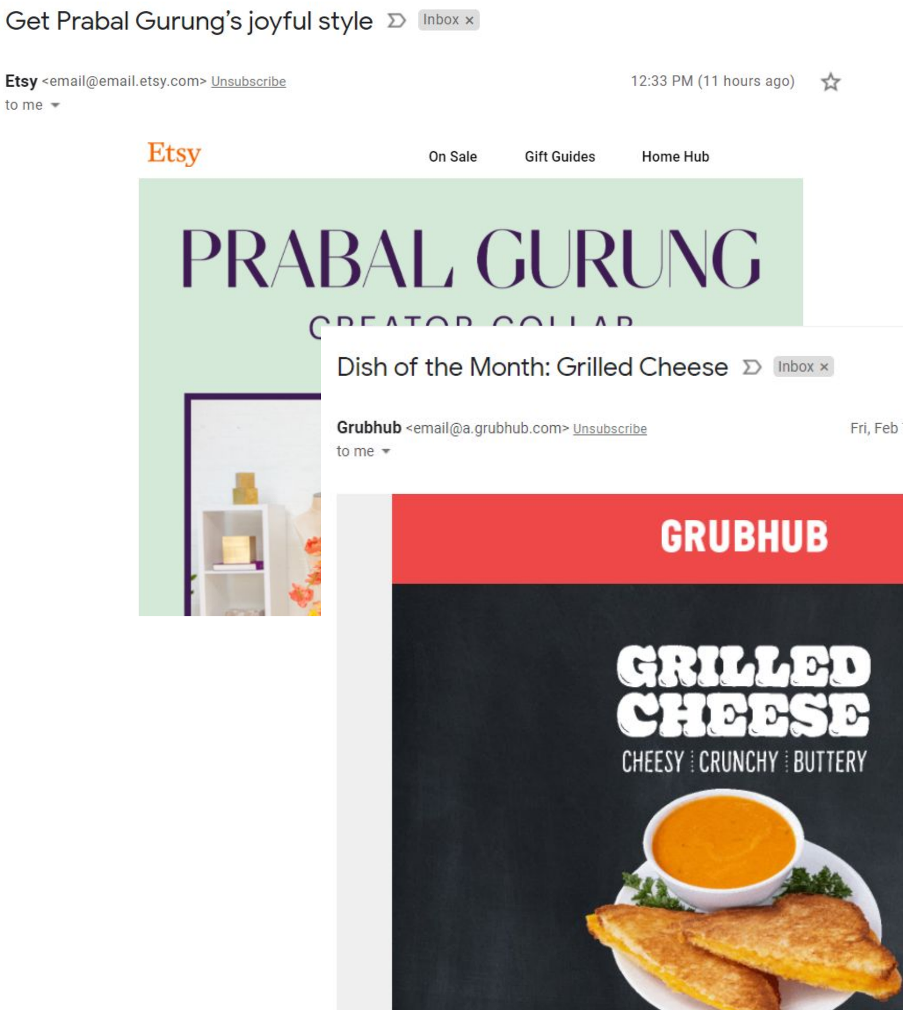
Assumption: Make Money, Scale



What Really Happens… And What We Should Do

Approach

- Test on a simple brand email campaign
- Use training data from a prior email
- Predict revenue for each email / customer segment pair



Approach

Training data:

- Training data is from a prior randomized experiment
  - 10% random holdout
- Predictors: features about the customer
  - Purchase behavior, email interaction, time on platform...
- Target: revenue generated 7 days after the email was sent

- One predictive function per email variant, including holdout
- Send variant v to customer c where f(v, c) = revenue is maximized

Results

- FAIL!
- Offline analysis

		recommendation	
		don't send	send
	don't send	DD	DS
actual	send	SD	SS

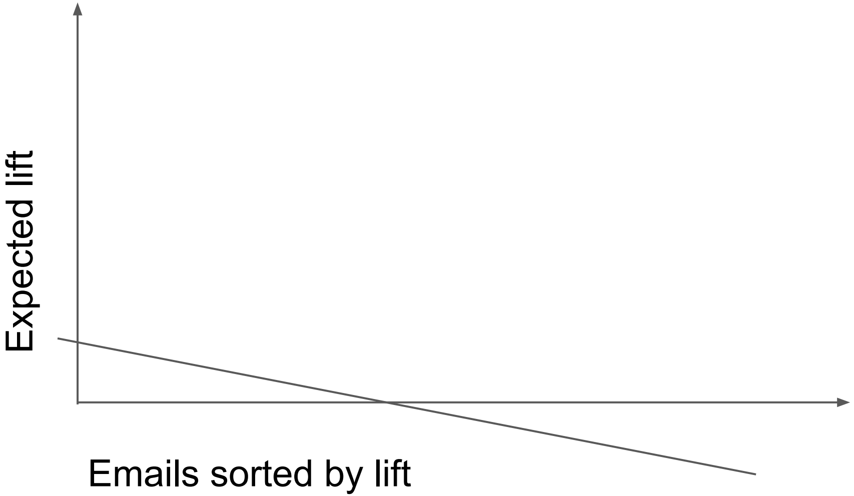
- Expect that customers whose treatment in the training data matched model recommendations to make more revenue
    - Revenue(DD) >? Revenue(DS)
    - Revenue(SS) >? Revenue(SD)
    - Differences were very noisy with small magnitude

Results

- Offline analysis
  - Interaction term significance
    - revenue(treatment, customerFeature) =
      - w0 + w1*treatment + w2*customer + w3*(treatment*customerFeature)
    - Expect w3 to be significant

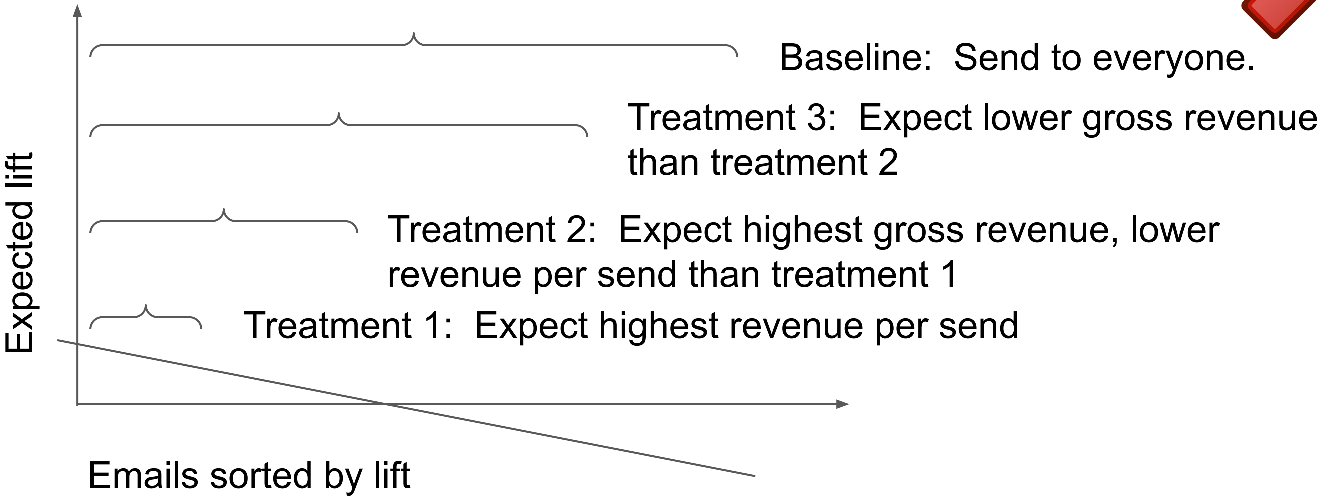
Results

Online test: Want to see monotonicity / sensical directionality in results...



Results

Online test: No monotonicity in revenue per send / directionality in gross revenue

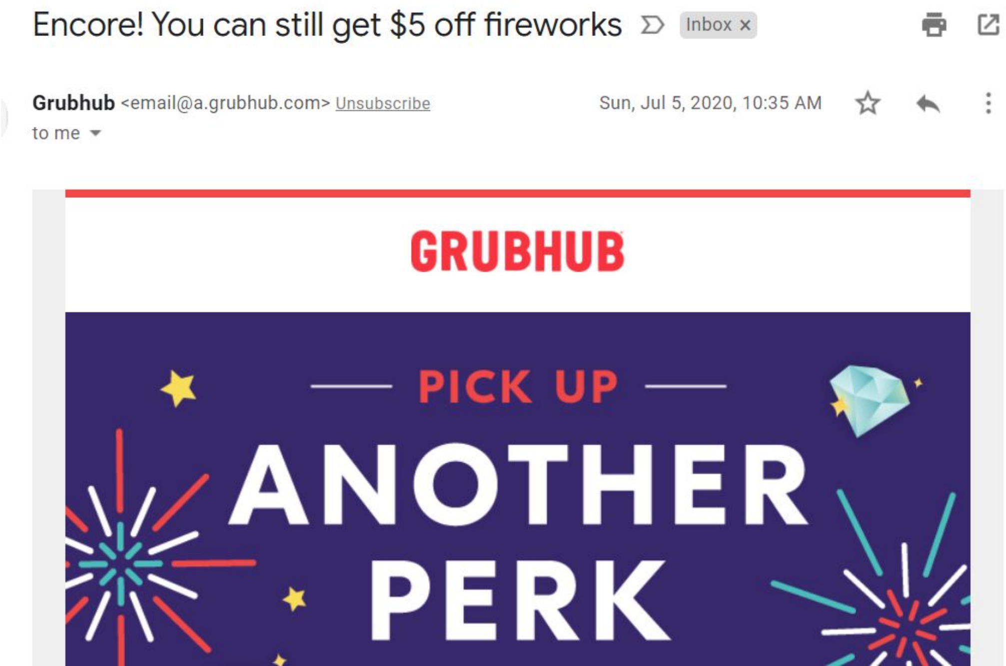


What To Do?

- Work with more complex algorithms
- Get better data ✅

Use A Coupon Email Campaign

- More than 50% greater engagement rates than other types of emails
- Varied dollar off and minimum purchase amount



Results

Treatment	Order Lift	Revenue Lift
$3 OFF $15	0.71%	1.63%
3 OFF 10	1.43%	-1.37%
5 OFF 15	3.13%	1.49%
5 OFF 10	4.52%	-7.67%
7 OFF 15	4.74%	-3.30%
7 OFF 10	7.72%	-13.52%
10 OFF 15	5.08%	-8.97%

Maximize revenue subject to not losing order volume

*Numbers are obfuscated but directionally consistent

Targeted Treatment - Heterogeneous Treatments

Order Lift								
Bucket	$3 OFF $15	3 OFF 10	5 OFF 15	5 OFF 10	7 OFF 15	7 OFF 10	10 OFF 15	holdout
D customers	**1.69%**	2.33%	2.34%	7.64%	6.10%	9.34%	9.03%	0.00%
C customers	**1.52%**	2.41%	3.43%	4.49%	4.29%	6.81%	2.84%	0.00%
B customers	-0.52%	0.40%	**1.00%**	1.05%	1.90%	3.96%	-0.61%	0.00%
A customers	-1.83%	-2.75%	4.49%	0.27%	4.37%	3.99%	2.05%	**0.00%**

Revenue Lift								
Bucket	3 OFF 15	3 OFF 10	5 OFF 15	5 OFF 10	7 OFF 15	7 OFF 10	10 OFF15	holdout
D customers	**<u>1.84%</u>**	-2.06%	1.76%	-9.75%	-5.96%	-21.37%	-17.75%	0.00%
C customers	**<u>1.60%</u>**	1.93%	1.18%	-11.23%	-4.13%	-19.10%	-12.37%	0.00%
B customers	0.92%	-4.03%	**<u>0.41%</u>**	-7.81%	-4.25%	-16.86%	-10.85%	0.00%
A customers	0.34%	-3.07%	-0.03%	-5.73%	-0.99%	-6.53%	-1.22%	**<u>0.00%</u>**

Tradeoff

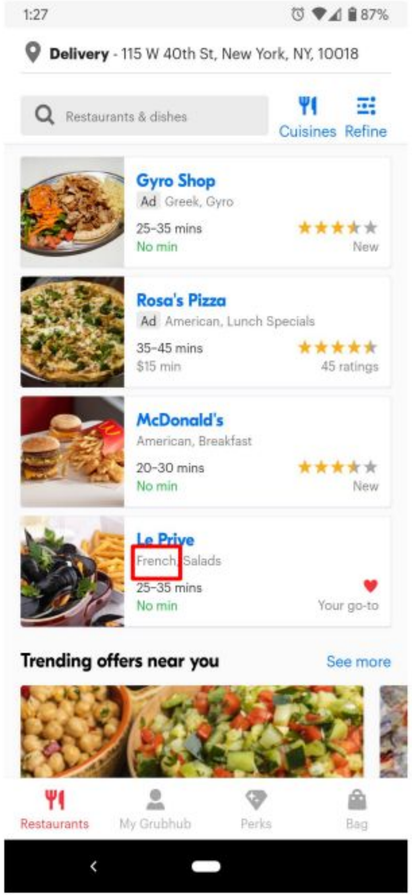
Lift vs. Holdout	Orders	Revenue
5off30	0.44%	1.07%
Targeted	1.15%	0.83%

Findings with Lift Modeling and Heterogeneous Treatment

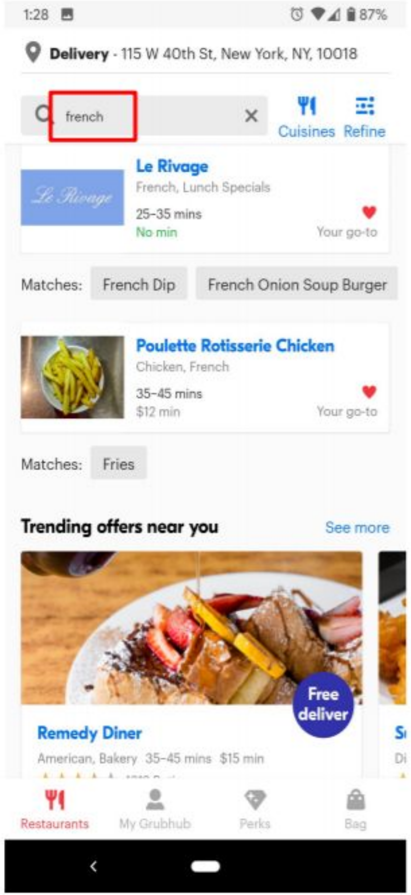
- Need data with sufficient signal
- Treatment effects can vary across populations
- Allows finer tuning of treatments depending on business goals

- Not covered: where this work fits into a bigger targeting framework

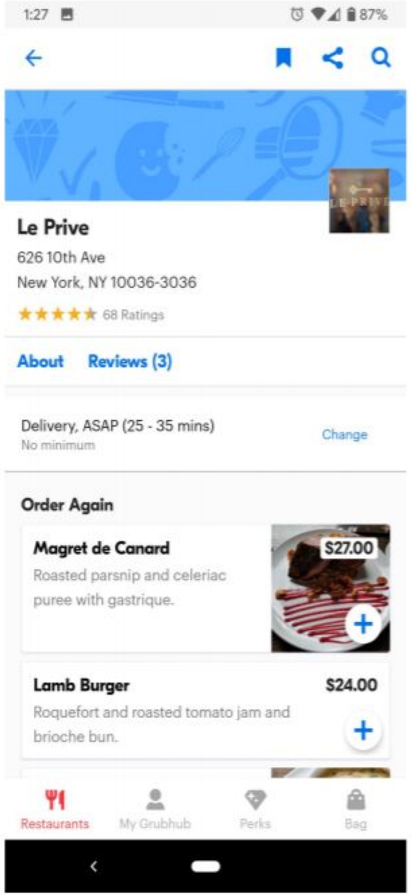
Restaurant Recommendations



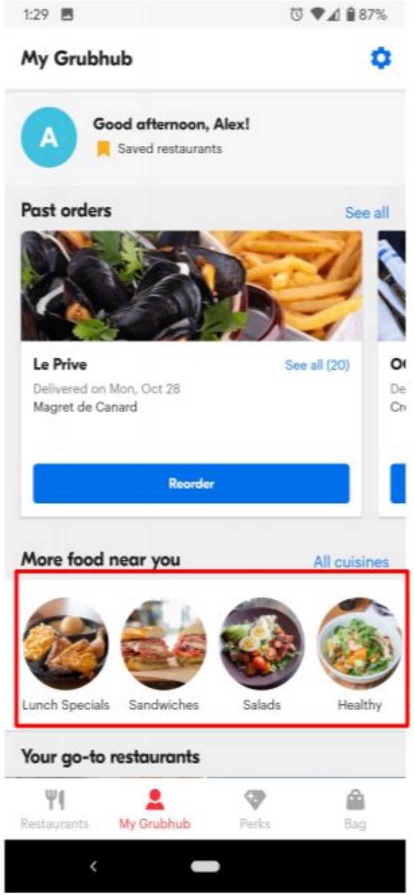
Rest/Dish/Cuisine Search



Menu/Dish Recommendations

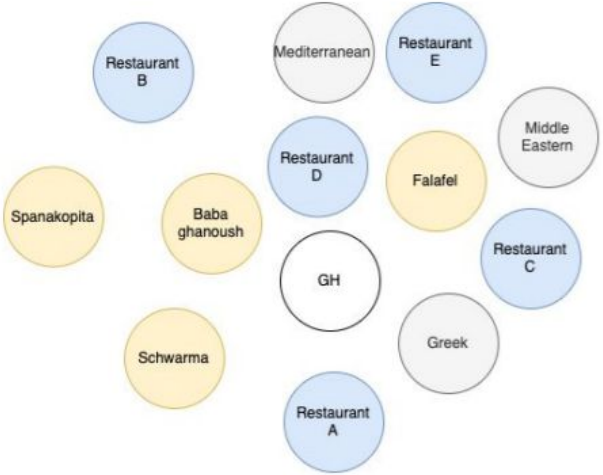


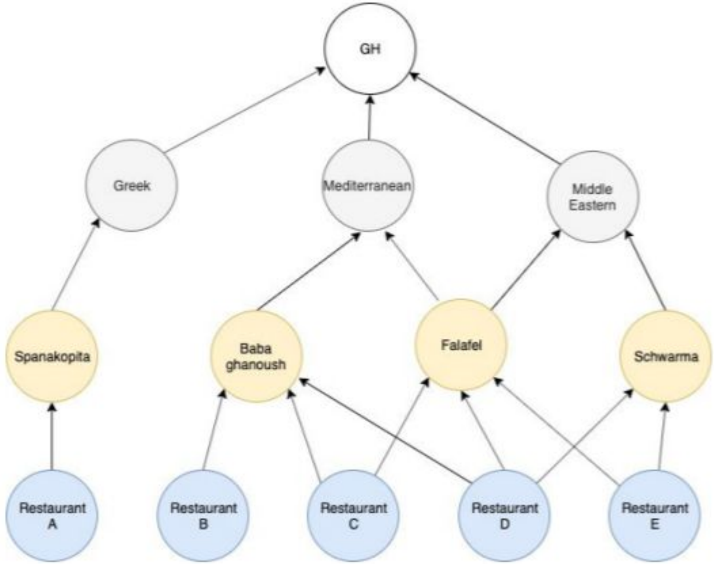
Cuisine Recommendations



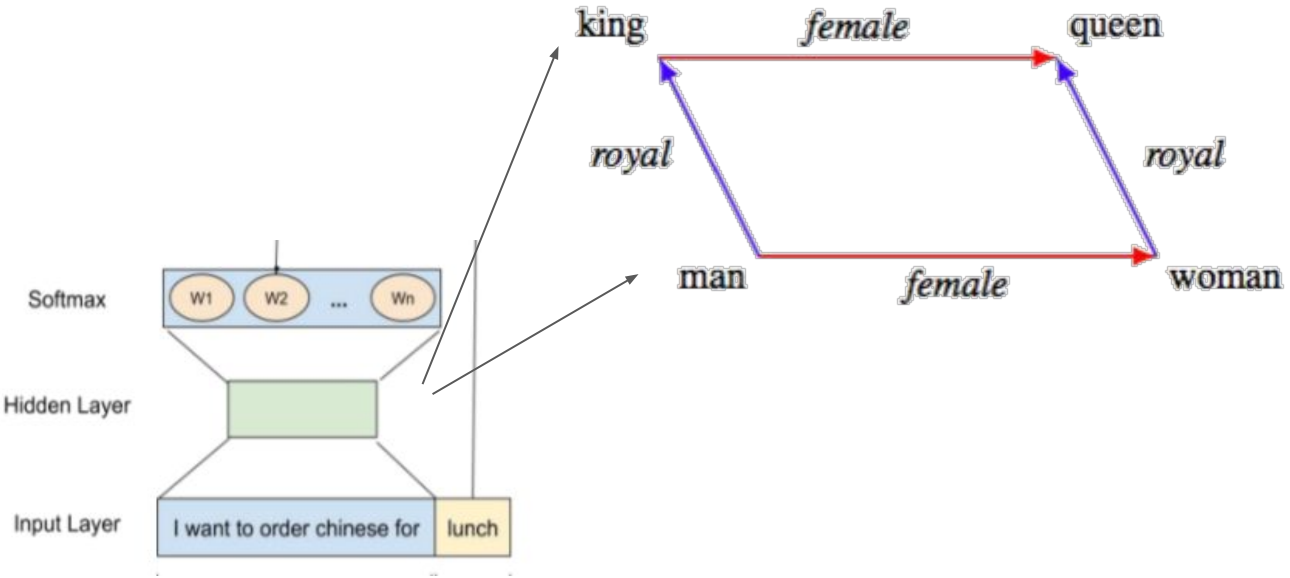
Ecommerce Dilemma

- Our catalog grows everyday
- Data is unstructured
- and unbounded
- *How can we understand it to drive: search & recommendations?*





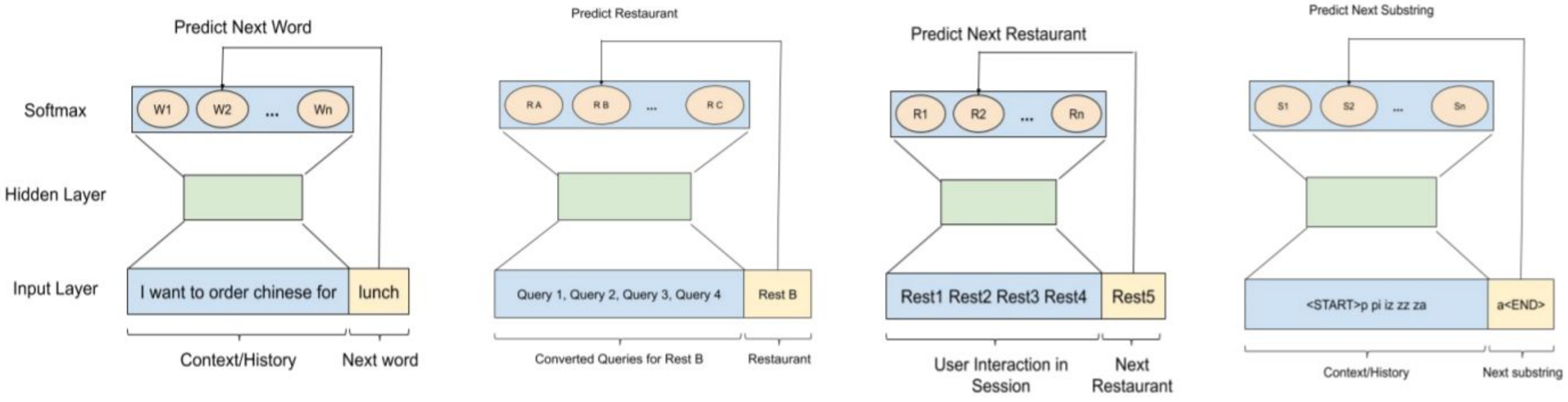
Embedding Primer



Representation Learning

1. Query2vec: understanding ***users***
2. Rest2vec: understanding ***restaurants***
3. FastMenu: understanding ***menus***

Users + Restaurants + Menus = Grubhub Food Universe



Search Pipeline

Query Understanding → **Query Building** → **Candidate Selection** → **Ranking** → **Enrichment**

Query Understanding
- Language Normalization
- Intent Classification

Query Building
- Filtering
- Query Expansion

Candidate Selection
- Phrase/Term Matching
- Semantic Matching

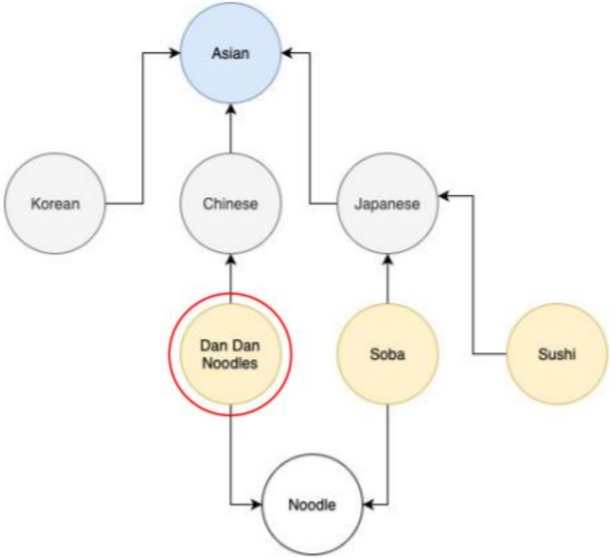
Ranking
- Revenue
- Relevance
- Personalization

Enrichment
- Pruning
- Hydration
- Pagination

Query Expansion w/ query2vec

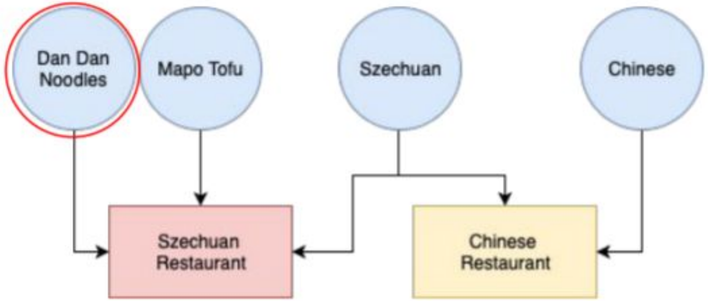
Classical Query Expansion

- Thesaurus/Synonyms:
  - Not Robust
  - “Dan Dan Noodles” -> cranium, brain, noggin, thinker
- Knowledge Graph
  - Difficult to build and maintain



Modern Query Expansion

- Representation Learning
  - Click Pattern Mining: Cluster similar queries based on converting restaurant
  - query2vec à la word2vec builds a virtual product knowledge graph



Nearest Neighbors

Search: alcohol — by: label
neighbors: 100
distance: COSINE EUCLIDEAN

Nearest points in the original space:

wine	0.219
beer	0.337
liquor	0.503
grocery store	0.571
grocery items	0.584
corona	0.596
red wine	0.612
grocery	0.618
westville chelsea	0.635
freshies	0.638
adesi cafe	0.646
paniab cafe	0.649

BOOKMARKS (0)

Search: italian — by: label
neighbors: 100
distance: COSINE EUCLIDEAN

Nearest points in the original space:

maggianos	0.485
vongole	0.532
pappardelle	0.550
pomodoro	0.557
gnochi	0.562
rigatoni	0.565
spaghetti carbonara	0.568
cacio e pepe	0.570
chicken milanese	0.571
polenta	0.572
orecchiette	0.575
bruschetta	0.576

BOOKMARKS (0)

Search: ramen — by: label
neighbors: 100
distance: COSINE EUCLIDEAN

Nearest points in the original space:

raman	0.386
japanese ramen	0.394
miso ramen	0.407
vegetarian ramen	0.428
vegan ramen	0.430
shoyu ramen	0.451
jin ramen	0.455
ramen soup	0.466
chicken ramen	0.468
spicy miso ramen	0.473
ramen noodles	0.486
tonkotsu ramen	0.515

Search: kimchi jjiga — by: label
neighbors: 10
distance: COSINE EUCLIDEAN

Nearest points in the original space:

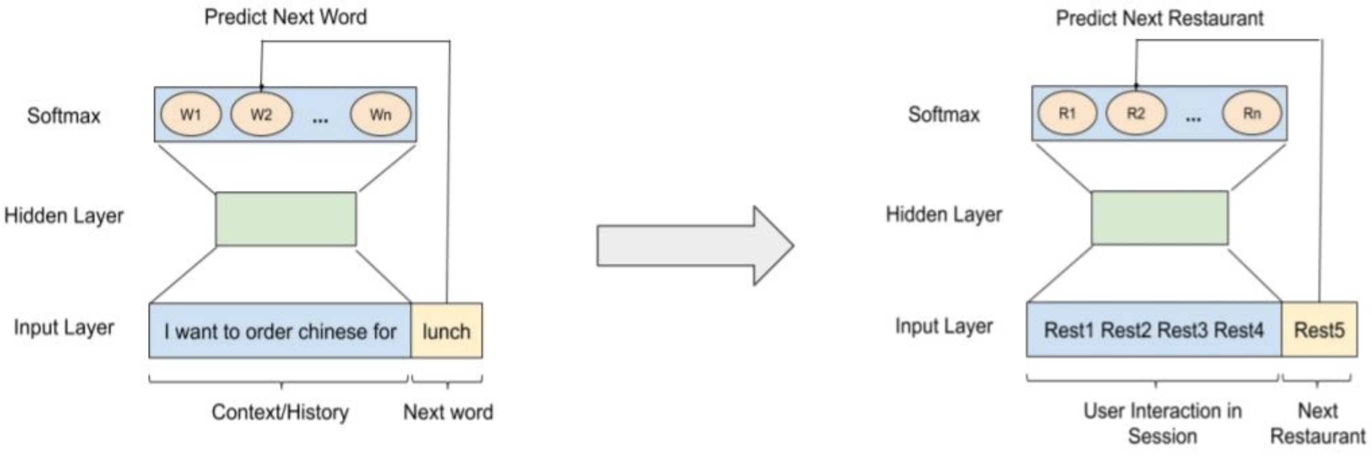
korean bbq	0.424
korean	0.431
korea	0.442
galbi tang	0.476
japcha	0.528
bibimbap	0.531
bulgogi	0.534
mokja	0.538
kimchi stew	0.542
galbi	0.545

Query2vec Latent Product Knowledge Graph

Rest2Vec

Creates numerical vector representation of restaurants from historical clickstream data using user's clicks/conversions

- Helps to understand Restaurants
- Helps to power Discovery
- Helps to power Personalization

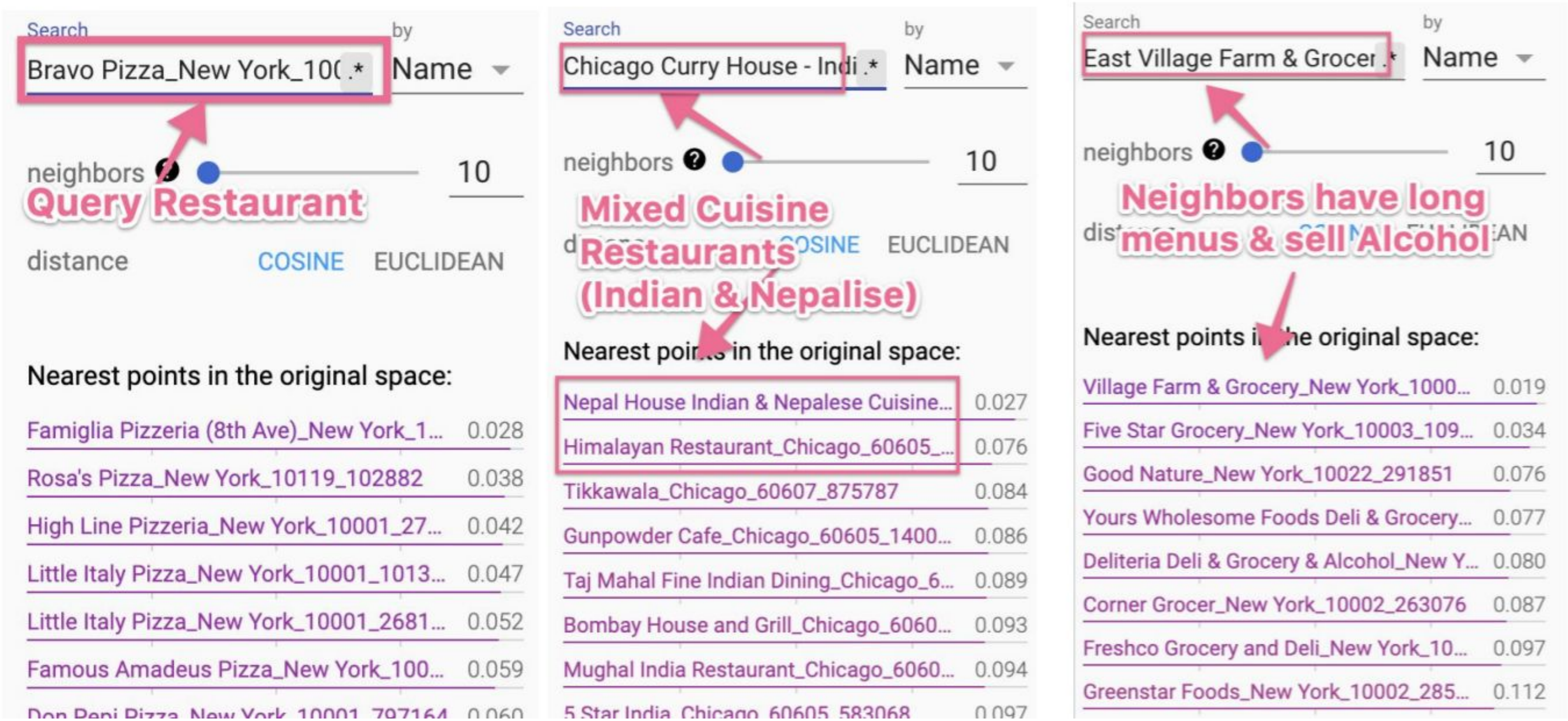


Word2Vec **Rest2Vec**

Tensorboard Visualization

- Each Market has its own Cluster
- Cluster size indicates how big the Market is

Nearest Neighbors



Menu Text Matching

Menu Item Description:

- mai fun
- blueberry pancake

String Matched Menu Items:

- mai fun, chow fun, shrimp mai fun
- blueberry smoothie, buttermilk pancake

Semantic Matched Menu Items

- stir fried noodles, thin rice noodles
- grand slam breakfast
- Increased recall

FastMenu

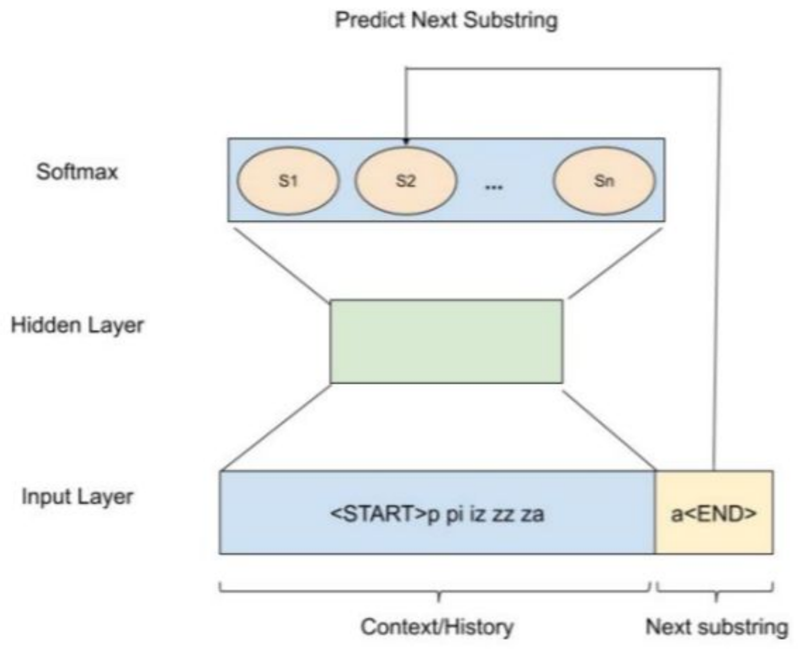
Creates numerical vector representation of **menu items** using associated textual data rather than diner behavior

- Helps to understand menus
- Helps to power semantic search
- Complete catalogue coverage

Static Sequence Embeddings

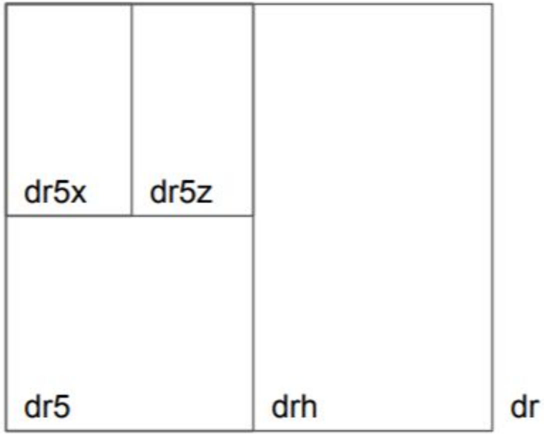
Fasttext = sub-words

- Handles out of vocabulary words
- “pizza”
- <START>p, pi, iz, zz, za, a<END>



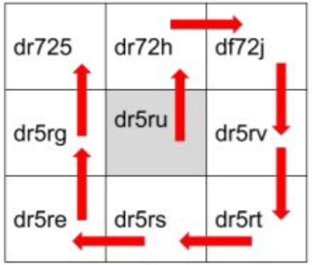
Geohashes

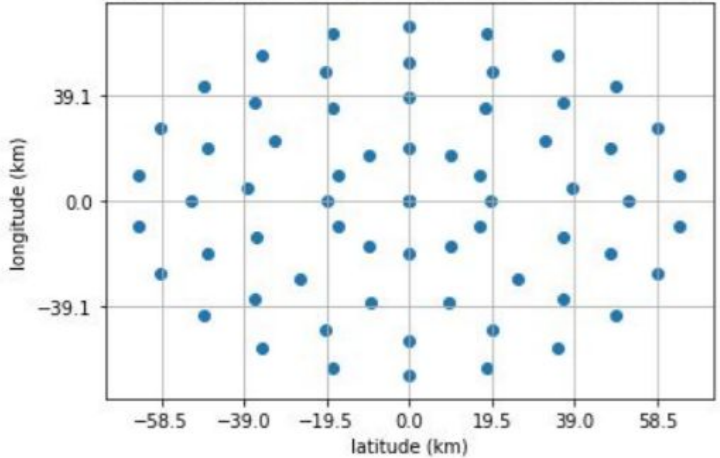
- Covers the surface of the earth
- Denotes rectangular area
- Alphanumeric string
- ~32 bit lat-long specification
- Nested precision levels



Geohash Embedding

Geohashes: same representative characters as language





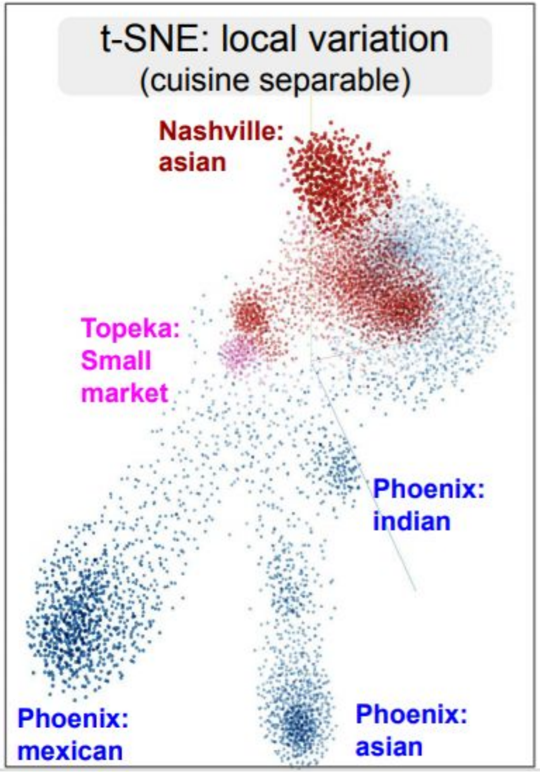
language	geohash
Word “pizza”	Geohash location “dr5ru”
Sentence “margarita **pizza** adorned simply with basil mozzarella tomato sauce pizza subs italian american lunch specials ”	Traversal pattern “dr725 dr72h dr72j dr5rg **dr5ru** dr5rv dr5re dr5rs dr5rt”

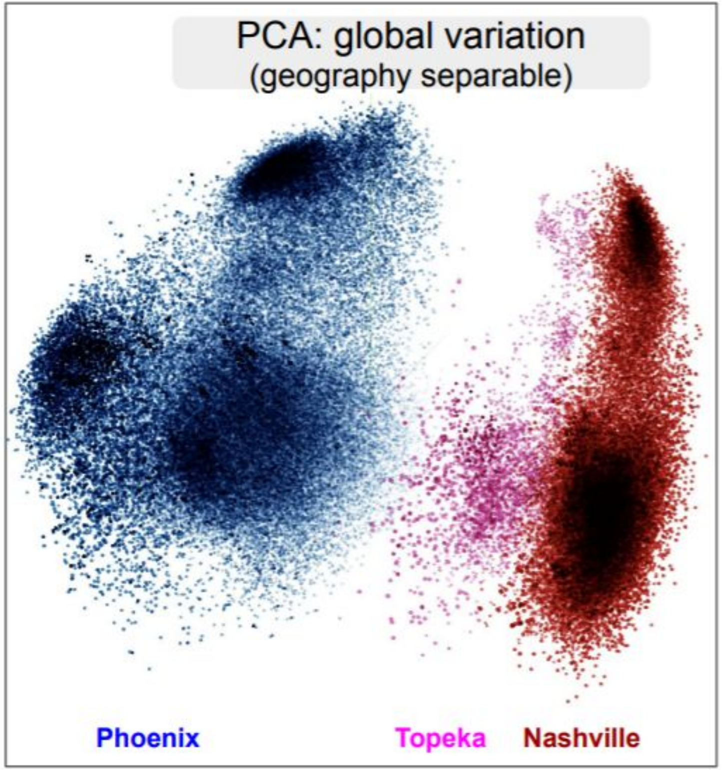
- Concat geohash sentence to menu text

“margarita pizza adorned simply with basil mozzarella tomato sauce pizza subs italian american lunch specials dr725 dr72h dr72j dr5rg dr5ru dr5rv dr5re dr5rs dr5rt”

- Expand “word” vocabulary, but still 26 chars and 10 numbers
- **Menu item text now knows about location**

Visualization: TensorBoard





- Topeka -> Nashville: 600 mi
- Topeka -> Phoenix: 1k m
- Topeka = few restaurants

Global trends
- Menu items separable by geography
- geography is strongest feature
- Embedding distance ~ physical distance
  - Topeka cluster closer to Nashville than Phoenix

Local trends
- Geohash strongest feature
- Cuisine features highly unique
- Small market size obscures separability

Now we preserve the semantic comparison but don’t recommend pizzas Phoenix to a diner in Nashville -> decreases latency at run time

Nearest Neighbors

1096076,57175260,2 Buttermilk Biscuits,Topeka KS

label 1096076,57175260,2 Buttermilk Biscuits,Topeka KS

Nearest points in the original space:

1095970,57175400,2 Buttermilk Biscuit...	0.005
1076808,41975502,Biscuits and Gravy,T...	0.018
1076808,41975510,Trainwreck,Topeka ...	0.026
1096076,57175264,Hash Browns,Topek...	0.030
1096076,57175309,Hash Browns,Topek...	0.030
1095970,57175384,Blueberry Pancake ...	0.031
1095970,57175396,1 Hearty Breakfast ...	0.033
1076808,41975503,Breakfast Biscuit,To...	0.034
1076808,41975514,Breakfast Bowl,Top...	0.034
1096076,57175256,1 Hearty Breakfast ...	0.035
1095970,57175449,Hash Browns,Topek...	0.035
1095970,57175404,Hash Browns,Topek...	0.035
1096076,57175244,Blueberry Pancake ...	0.035
1095970,57175373,Lumberjack SlamÂ...	0.035

Thin Rice Noodle Tray

cust_id	302844
menu_item_id	46097558
name	Thin Rice Noodle Tray
location	Nashville TN

Nearest points in the original space:

Flat Rice Noodle Tray	0.002
Vegetarian Fried Noodles Lunch Special	0.007
House Special Mei Fun	0.007
Vegetable Mai Fen	0.008
40. House Mai Fen	0.008
54. Vegetable Mai Fun	0.008
Pork Mei Fun	0.008
Shrimp Mei Fun	0.009
58. House Special Mei Fun	0.009
53. Vegetable Mei Fun	0.009
58. House Special Mai Fun	0.009
55. Chicken Mai Fun	0.009
53. Vegetable Mai Fun	0.009
Lo Mein Noodle Tray	0.009
59. House Special Mai Fun	0.009
Combination Pan Fried Noodle	0.010

Nearest Neighbors

921564,55189279,Grande Nachos Box - Seasoned Beef,Nashville TN

label 921564,55189279,Grande Nachos Box - Seasoned Beef,Nashville TN

distance COSINE EUCLIDEAN

Nearest points in the original space:

923311,55181646,Grande Nachos Box - ...	0.000
923311,55181642,Grande Nachos Box - ...	0.000
922837,55185452,Grande Nachos Box - ...	0.000
922837,55185454,Grande Nachos Box - ...	0.000
922837,55185456,Grande Nachos Box - ...	0.000
921216,55189327,Grande Nachos Box - ...	0.000
921216,55189325,Grande Nachos Box - ...	0.000
921216,55189323,Grande Nachos Box - ...	0.000
921568,55191370,Grande Nachos Box - ...	0.000
921568,55191372,Grande Nachos Box - ...	0.000
921568,55191374,Grande Nachos Box - ...	0.000
922249,55185666,Grande Nachos Box - ...	0.000

37. Beef

cust_id	289964
menu_item_id	12216021
name	37. Beef
location	Nashville TN

Nearest points in the original space:

56. Beef Pad Thai	0.007
46b. Pad Thai Beef	0.007
32. Chicken Chop Suey	0.007
49. Roast Pork with Black Bean Sauce	0.007
31. House Special Fried Rice	0.007
45. Sweet and Sour Pork	0.007
Beef Fried Rice	0.007
103. Hunan Beef	0.007
104. Szechuan Beef	0.007
101. Quart of Shredded Beef with Scalli...	0.007
32. Chicken Chow Mein	0.007
34. House Fried Rice	0.007
72. Sweet and Sour Chicken	0.007
17. Kung Pao Beef Dinner	0.007

Conclusions and Future Work

AMAZING!!! blueberry pancakes

Are pierogis really empanadas?!?

10 Asian noodles near you

Try this French restaurant instead

- But lots more to go to tune models to optimize relevance and lifetime values

Thanks

- Contact info: waigen at ieee

- Thanks, too, to Alex, Emily, Parin, Sahil, Weiwen, Yong, Renata

References

- Changji Shi and Adriel Fuad, Air Traffic Controller: Member-First Notifications at LinkedIn, 2018-03-01
- Mark White, EXPLICITLY OPTIMIZING ON CAUSAL EFFECTS VIA THE CAUSAL RANDOM FOREST: A PRACTICAL INTRODUCTION AND TUTORIAL, 2018-10-14
- Eva Ascarza, Retention Futility: Targeting High-Risk Customers Might be Ineffective, Journal of Marketing Research, 2018-02-01
- Skyward, How the Marketing Funnel Works From Top to Bottom, https://www.skyword.com/contentstandard/how-the-marketing-funnel-works-from-top-to-bottom/, 2020-10-02
- Sören R. Künzel, Jasjeet S. Sekhon, Peter J. Bickel, Bin Yu, Meta-learners for Estimating Heterogeneous Treatment Effects using Machine Learning, Proceedings of the National Academy of Sciences 116(10), June, 2017
- Alex Egg, Emily A Ray, Parin Choganwala: Discover your latent food graph with this 1 weird trick, PyData New York 2019
- T. Mikilov, W. Yih, G. Zweig, Linguistic Regularities in Continuous Space Word Representations, Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies
- Grbovic et. al., E-commerce in Your Inbox: Product Recommendations at Scale, KDD, 2015
- Bojanowski et. al., Enriching Word Vectors with Subword Information, ACL, 2017
- Alex Egg, Query2vec: Search query expansion with query embeddings, Grubhub Bytes Blog, 2019