Using Data Science to Transform OpenTable Into Your Local Dining Expert

Pablo Delgado @pablete

Sudeep Das @datamusing

OpenTable

the world's leading provider of online restaurant reservations

- Over 32,000 restaurants worldwide
- more than 760 million diners seated since 1998, representing more than $30 billion spent at partner restaurants
- Over 16 million diners seated every month
- OpenTable has seated over 190 million diners via a mobile device. Almost 50% of our reservations are made via a mobile device
- OpenTable currently has presence in US, Canada, Mexico, UK, Germany and Japan
- OpenTable has nearly 600 partners including Facebook, Google, TripAdvisor, Urbanspoon, Yahoo and Zagat.

At OpenTable

we aim to power

the best dining experiences!

Connecting the dots

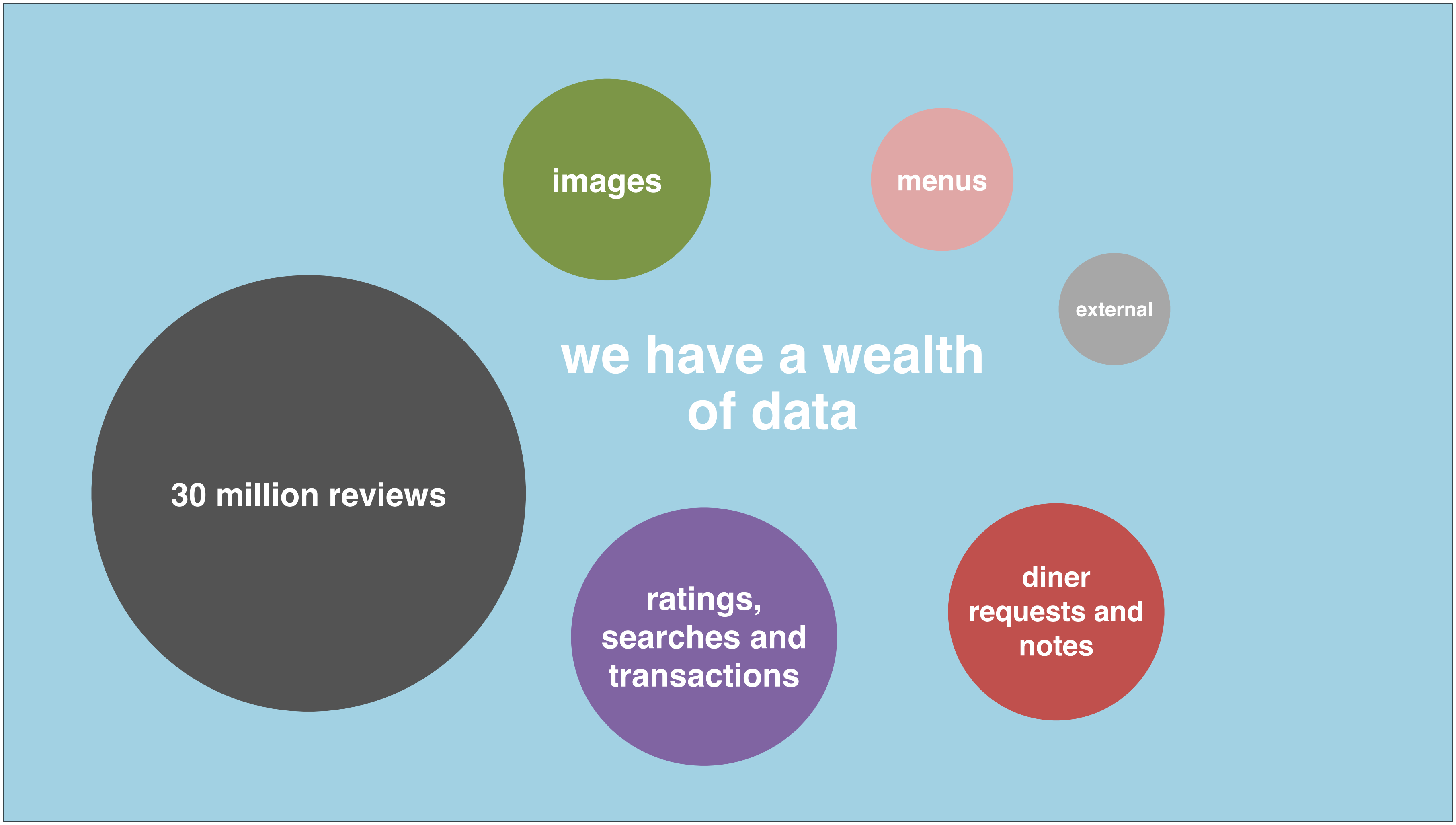
Understanding the diner

Building up a profile of you as a diner from explicit and implicit signals - information you have provided, reviews you have written, places you have dined at etc.

Understanding the restaurant

What type of restaurant is it?
What dishes are they known for?
Is it good for a date night/ family friendly/ has amazing views etc.
What's trending?

Ingredients of a magical experience


images
menus
external
we have a wealth of data
30 million reviews
ratings, searches and transactions
diner requests and notes

Making meaningful recommendation

The basic ingredients

diner-restaurant Interactions

ratings|searches|reviews

...

restaurant metadata

cuisine|price range|hours|topics

...

user metadata

user profile

Making recommendations starts with collecting the signal

Old data format: ETL with Spark

- We are mining log data all the way back to 2010 with Spark
- The result is to compute the number of times each user has searched for a restaurant.

```
2014-01-01 08:00:13 W3SVC1554668004 10.20.20.77 GET /opentables.aspx t=rest&r=11671&m=212&p=2&d=12/31/2013%2011:30:00%20PM&scpref=100 2020 -
64.145.88.235 Mozilla/5.0+(iPad;+CPU+OS+7_0_4+like+Mac+OS+X)+AppleWebKit/537.51.1+(KHTML,+like+Gecko)+Version/7.0+Mobile/11B554a+Safari/
9537.53GCSCU_1011230515163_H2=C=1011230515163.apps.googleusercontent.com:S=9cbd7893239881e8ac1f50262ff8508463c7b337.gSm24muLAXog5z3_.c2f1:I=
1388563131:X=1388649531;+__qca=P0-1372117488-1388563123763;+aCKE=e1722c4c-644b-413e-8173-
d9d5288b1682;+ftc=x=01%2f01%2f2014+11%3a00%3a10&p1=100&p1q=m%3d212%26mn%3d1075%26ref%3d13573%26sp
%3dppc_b_Anchorage_nontm&rf1=13573&rf2=13573&er=11671 302 0 0
```

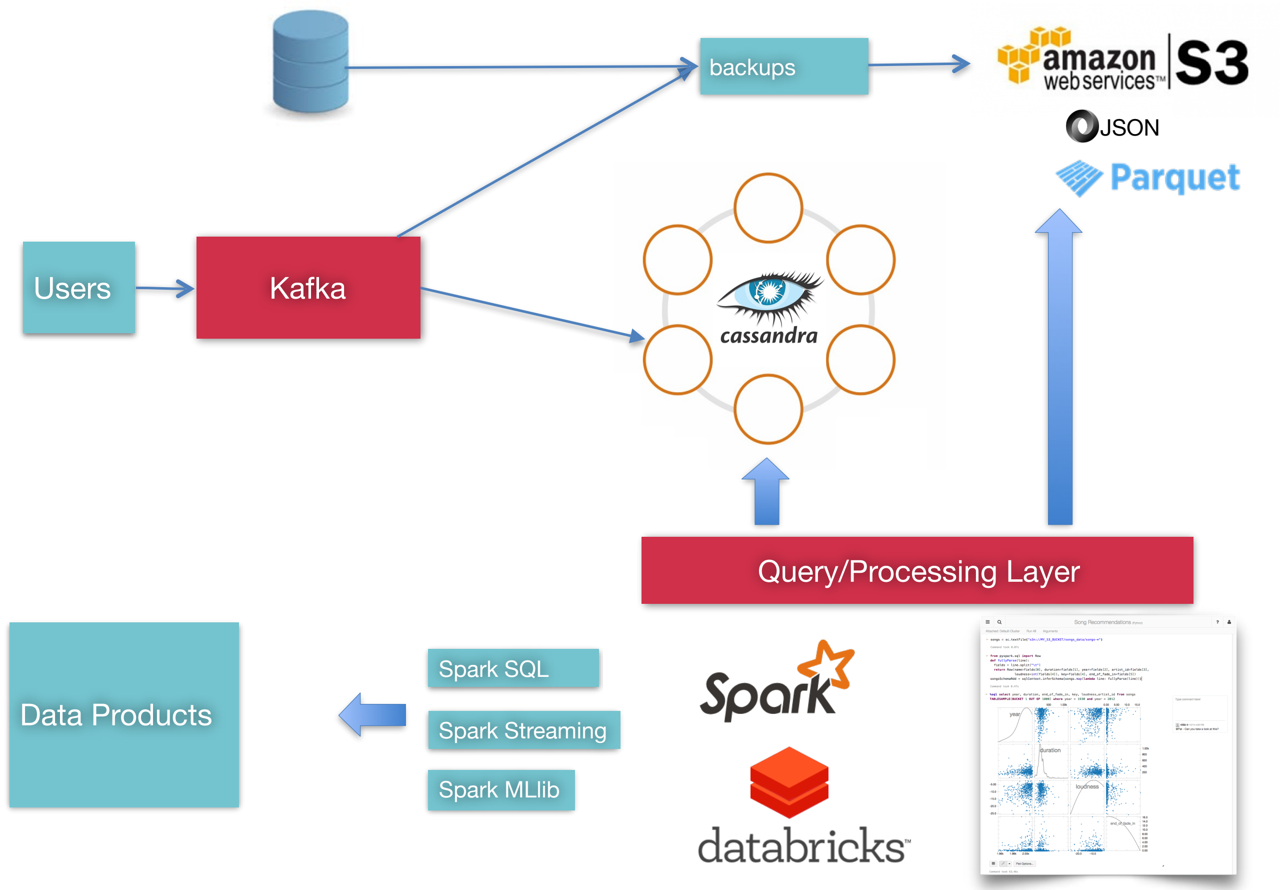
```
SingleSearches(timestamp, rid, uid, reservationSlot,
partySize)
```

New data format: ETL with Spark

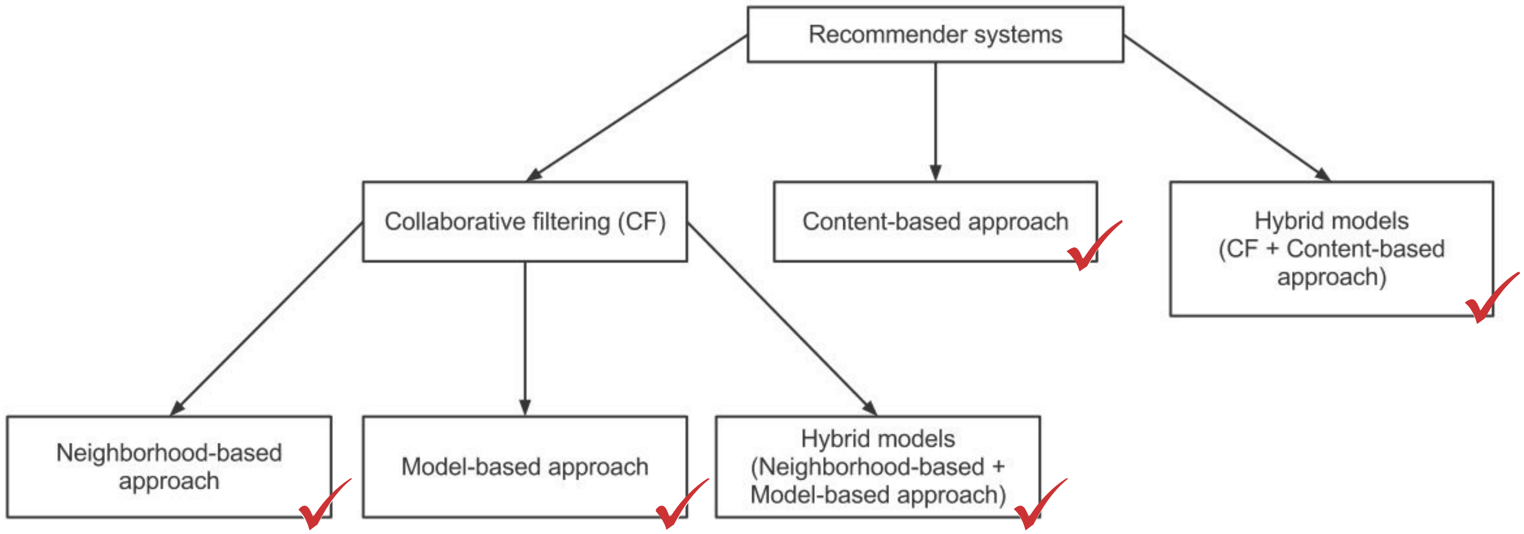
{“userId”:"xxxxxxxx","event":"personalizer_search","query_longitude":-77.16816,"latitude":38.918159,"req_attribute_tag_ids":
["pizza"],"req_geo_query":"Current Location”,"sort_by":"best","longitude":-77.168156,"query_latitude":38.91816,"req_forward_minutes":
30,"req_party_size":2,"req_backward_minutes":30,"req_datetime":"2015-06-02T12:00","req_time":"12:00","res_num_results":
784,"calculated_radius":5.466253405962307,"req_date":"2015-06-02"},"type":"track","messageId":"b4f2fafc-
dd4a-45e3-99ed-4b83d1e42dcd","timestamp":"2015-06-02T10:02:34.323Z"}

```
val single_searches = sqlContext.jsonFile("s3://single_searches/20150909.json")
single_searches.registerTempTable("searches")

val ss_df = sqlContext.sql(
  "SELECT timestamp, rid, uid, reservation_slot, party_size FROM searches"
)
```

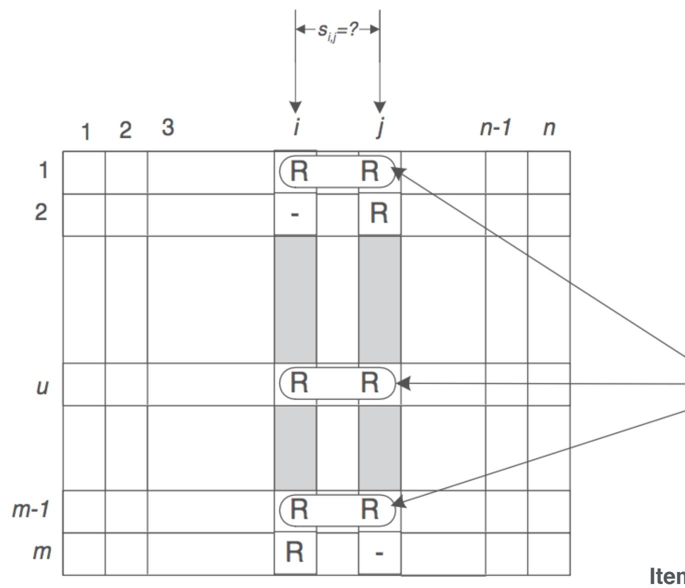

backups
amazon web services
S3
JSON
Parquet
Users
Kafka
cassandra
Query/Processing Layer
Data Products
Spark SQL
Spark Streaming
Spark MLlib
Spark
databricks

There are various approaches to making meaningful recommendations



http://en.wikipedia.org/wiki/File:Collaborative_Filtering_in_Recommender_Systems.jpg

Neighborhood-based CF
Item/Item or Restaurant Similarity



Item-item similarity is computed by looking into co-rated items only. In case of items *i* and *j* the similarity $s_{i,j}$ is computed by looking into them. Note: each of these co-rated pairs are obtained from different users, in this example they come from users *1, u* and *m-1*.

Item-Based Collaborative Filtering Recommendation Algorithms (Sarwar et al., 2001)

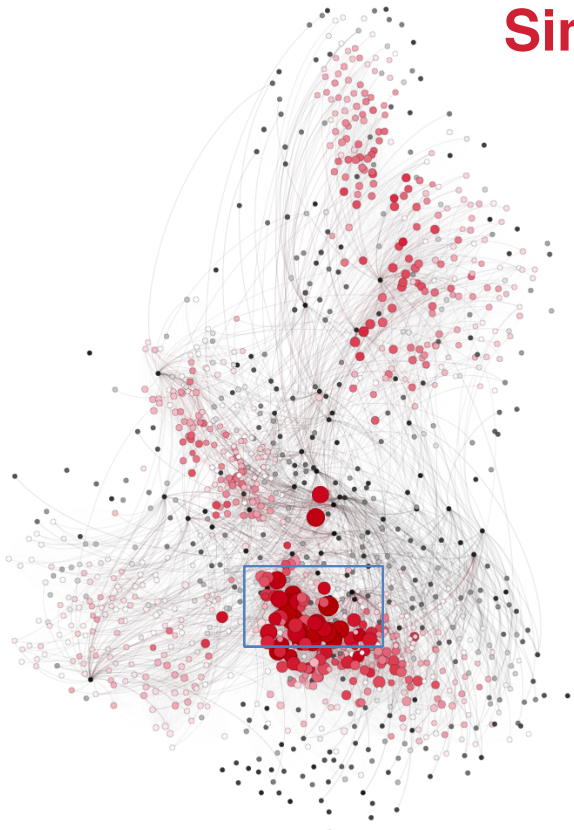
We use the Jaccard Index as a measure of similarity

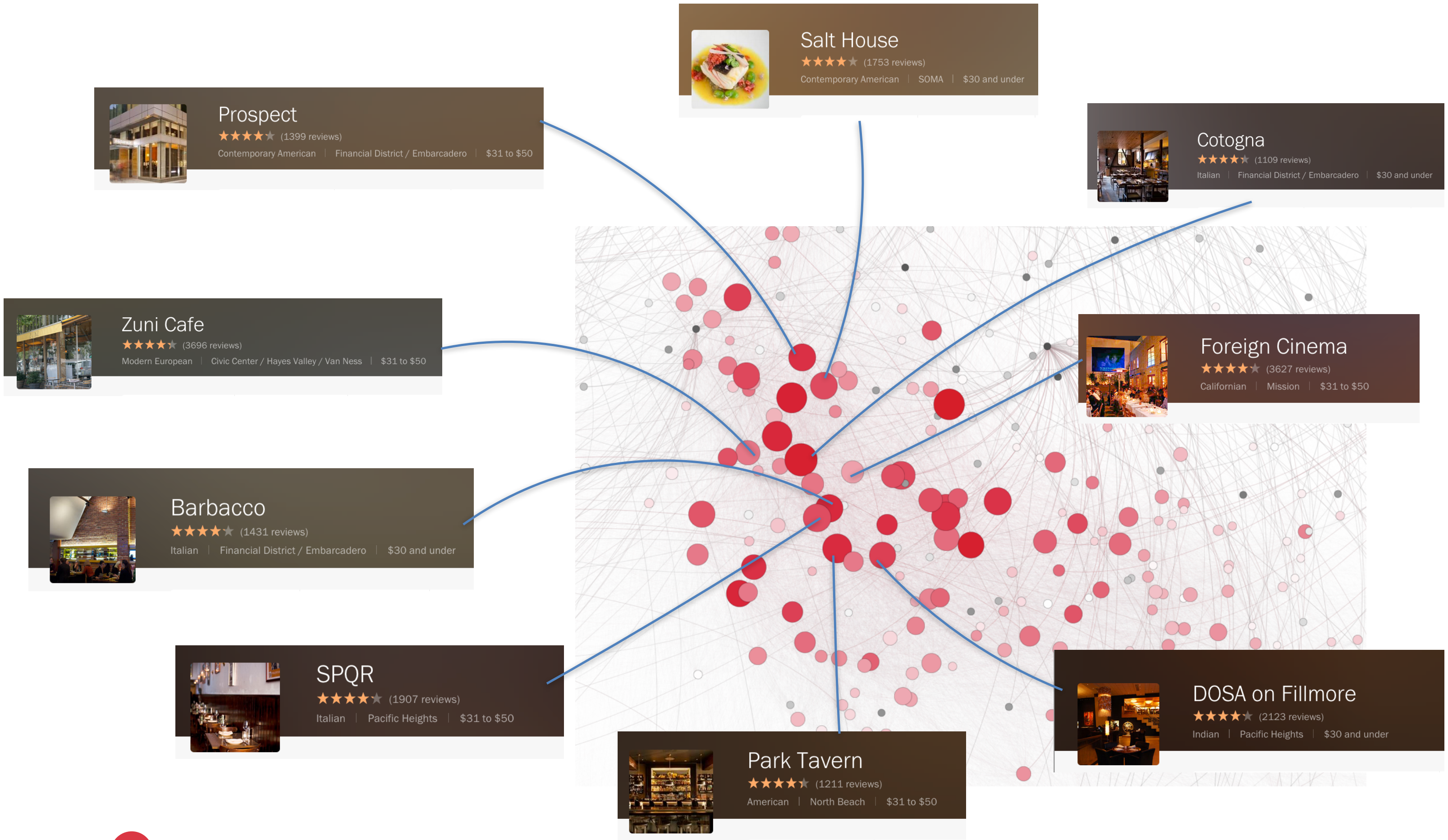
$$J(A, B) = \frac{|A \cap B|}{|A \cup B| + \beta}$$

where β is a shrinkage parameter to penalize item pairs that have too few users.

```
/**
 * The Jaccard Similarity between two sets A, B is
 *   |Intersection(A, B)| / (|Union(A, B)| + beta)
 */
def jaccardSimilarity(InCommon: Int, A: Int, B: Int, beta: Int) = {
  val union = A + B - InCommon
  InCommon / (union + beta)
}
```

Recommendations: Restaurant Similarity



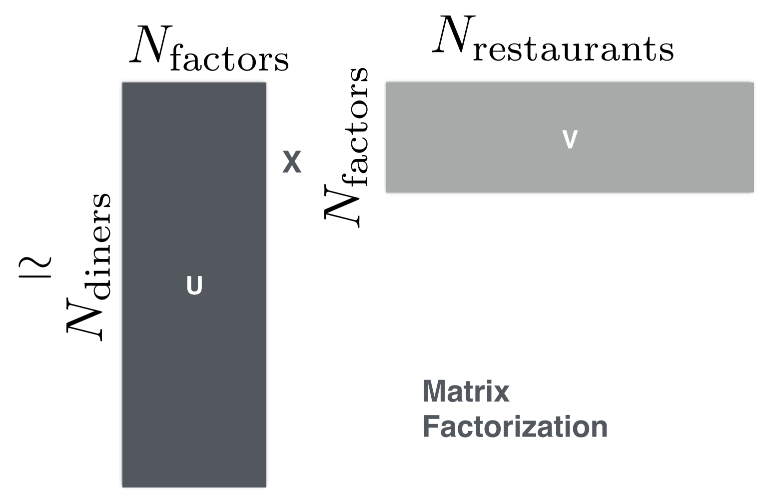


Salt House
(1753 reviews)
Contemporary American | SOMA | $30 and under
Prospect
(1399 reviews)
Contemporary American | Financial District / Embarcadero | $31 to $50
Cotogna
(1109 reviews)
Italian | Financial District / Embarcadero | $30 and under
Zuni Cafe
(3696 reviews)
Modern European | Civic Center / Hayes Valley / Van Ness | $31 to $50
Foreign Cinema
(3627 reviews)
Californian | Mission | $31 to $50
Barbacco
(1431 reviews)
Italian | Financial District / Embarcadero | $30 and under
SPQR
(1907 reviews)
Italian | Pacific Heights | $31 to $50
DOSA on Fillmore
(2123 reviews)
Indian | Pacific Heights | $30 and under
Park Tavern
(1211 reviews)
American | North Beach | $31 to $50

Matrix Factorization: Explicit ratings

Explicit Case: Use ratings

	Restaurant_1	Restaurant_2	…	Restaurant_M
Diner_1	5	?	…	3
Diner_2	?	3	…	?
…	…	…	…	
Diner_N	4	4	…	1



Matrix Factorization: Implicit preferences

	Restaurant_1	Restaurant_2	...	Restaurant_M
Diner_1	50	?	…	100
Diner_2	?	1	…	?
...	…	…	…	
Diner_N	3	30	…	1

$$p_{ij} = \begin{cases} 1, & r_{ij} > 0 \\ 0, & r_{ij} = 0 \end{cases}$$

Binary Preference Matrix

$$c_{ij} = 1 + \alpha r_{ij}$$

Confidence Matrix

Collaborative Filtering for Implicit Feedback Datasets (Hu, Koren, Volinsky 2008)

Matrix Factorization: Implicit preferences

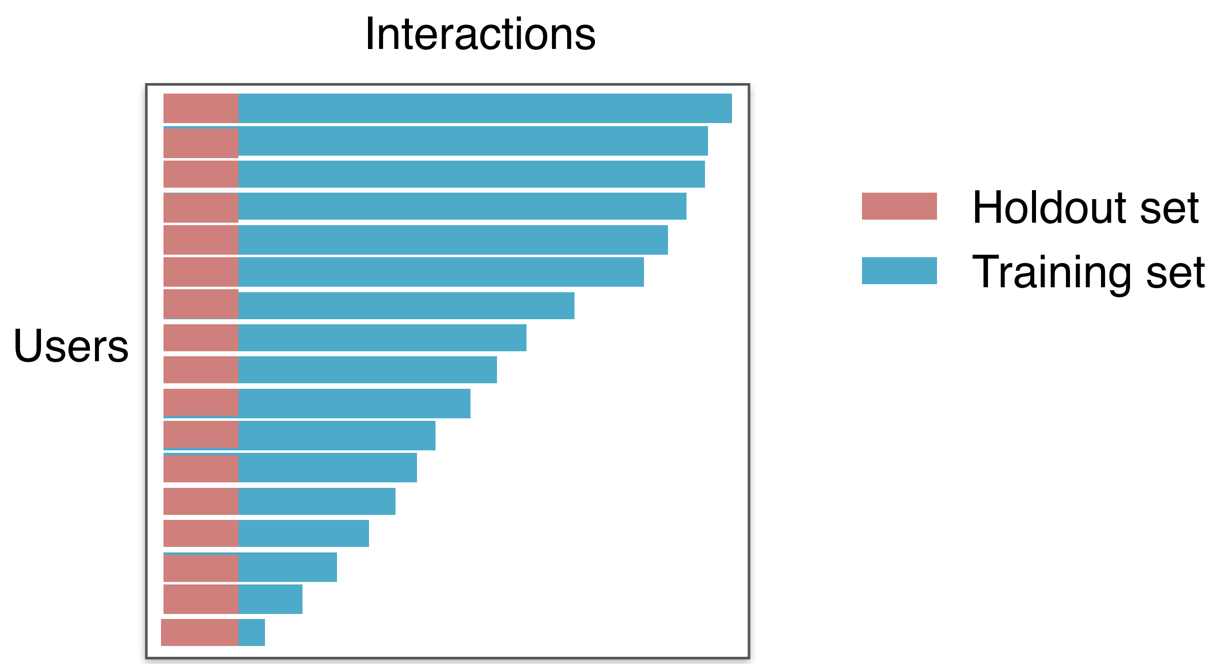
```
val model = ALS.trainImplicit(ratings, rank, numIterations, lambda, alpha)
```

$$\min_{x_\star, y_\star} \sum_{u,i} c_{ui}(p_{ui} - x_u^T y_i)^2 + \lambda \left(\sum_u \|x_u\|^2 + \sum_i \|y_i\|^2 \right) \quad (3)$$

The $\lambda \left(\sum_u \|x_u\|^2 + \sum_i \|y_i\|^2 \right)$ term is necessary for regularizing the model such that it will not overfit the training data. Exact value of the parameter λ is data-dependent and determined by cross validation.

Collaborative Filtering for Implicit Feedback Datasets (Hu, Koren, Volinsky 2008)

Evaluation / Hyperparameter tuning with Precison@K



Hyperparameter tuning using Precision@K

$$\text{Precision@K} = \frac{1}{K}|\text{topK} \cap \text{holdout}|$$

```
def precisionK(recommendation: Set[Int] , holdout: Set[Int], k: Int) = {
  (recommendation & holdout).size*1.0 / k
}
```

Precision@8 = 3/8

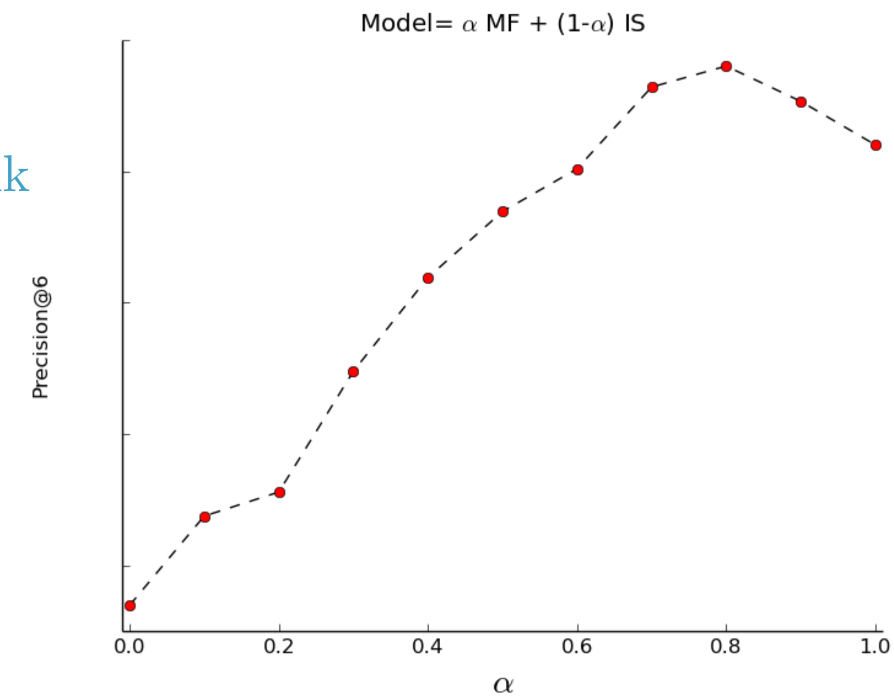
$$\frac{1}{m}\sum_{i=1}^{m} \text{Precision@K}(i)$$

```
def computePrecisionK(model: MatrixFactorizationModel, k: Int) = {
  var allrecommendations = recommendAll(model, k)
  val allprecisionk = allrecommendations.map({ x =>
     val user  = x._1
     val restaurants = x._2.map(y => y._1)
     val recommended_for_user = restaurants.toSet
     val holdout_for_user = Try(holdout_lookup.value(user)).getOrElse(Set())
     precisionk(recommended_for_user, holdout_for_user, k)
  }).mean
  allprecisionk
}
```

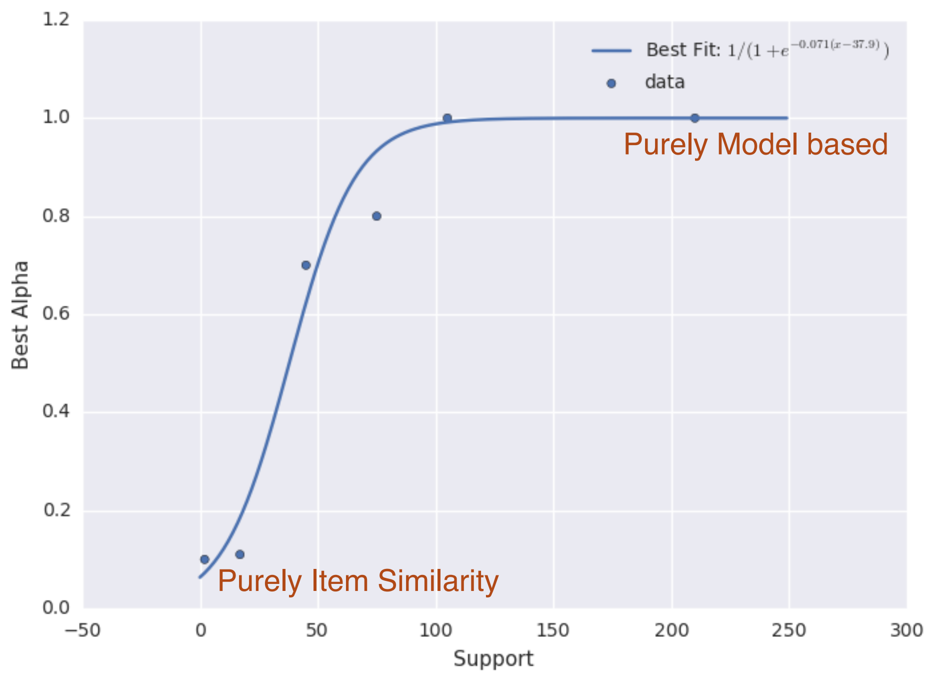
Ensemble of item-similarity and MF-based recommendations

Weighted mean inverse rank

$$\bar{a} = \alpha \frac{1}{r_1} + (1 - \alpha) \frac{1}{r_2}$$



Ensemble parameter is a function of the user support



OpenTable

Sunday, May 10
is fast approaching…

Did you know Mother's Day tables go like hotcakes? The time has come to snag yours.

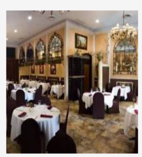
Book a Table

Pick the perfect place for Mother's Day…

1300 on Fillmore »
American
Pacific Heights

A16 Rockridge »
Italian
Oakland

Farina »
Italian
Mission



L'Olivier »
French
Financial District /
Embarcadero

OpenTable

At long last — it's summer!

With a new season upon us, Memorial Day weekend is the perfect time to celebrate summer and honor the spirit of the holiday with a relaxed meal.

Book a Table

Foreign Cinema »
Californian
Mission

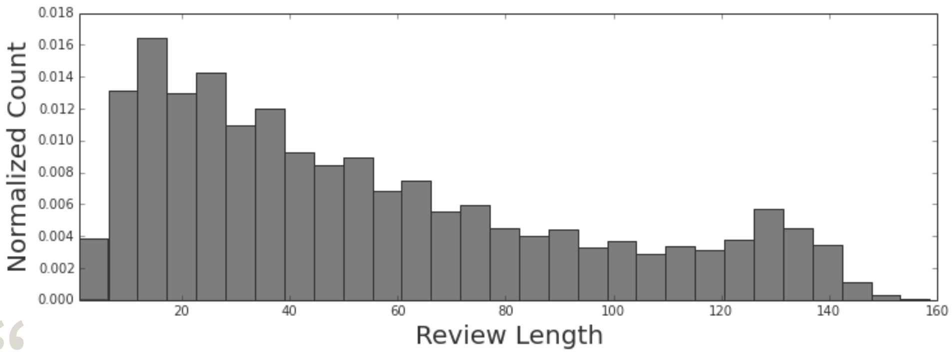
Cotogna »
Italian
Financial District /
Embarcadero

Bar Tartine »
Contemporary American
Mission

Coqueta »
Spanish
Financial District /
Embarcadero

Mining Content for Recommendations

Our reviews are rich and verified, and come in all shapes and sizes



Superb!

Many restaurants have thousands of reviews.

This really is a hidden gem and I'm not sure I want to share but I will. :) The owner, Claude, has been here for 47 years and is all about quality, taste, and not overcharging for what he loves. My husband and I don't often get into the city at night, but when we do this is THE place. The Grand Marnier Souffle' is the best I've had in my life - and I have a few years on the life meter. The custard is not over the top and the texture of the entire dessert is superb. This is the only family style French restaurant I'm aware of in SF. It also doesn't charge you an arm and a leg for their excellent quality and that also goes for the wine list. Soup, salad, choice of main (try the lamb shank) and choice of dessert - for around $42 w/o drinks.

We expect diner reviews to be broadly composed of a handful of broad themes

- Food & Drinks
- Ambiance
- Service
- Value for Money
- Special occasions

This motivated diving into the reviews with topic modeling

We approached the problem from the point of view of summarizing each restaurant using its reviews

Analyze the corpus of reviews in a geographic region to learn **topics**

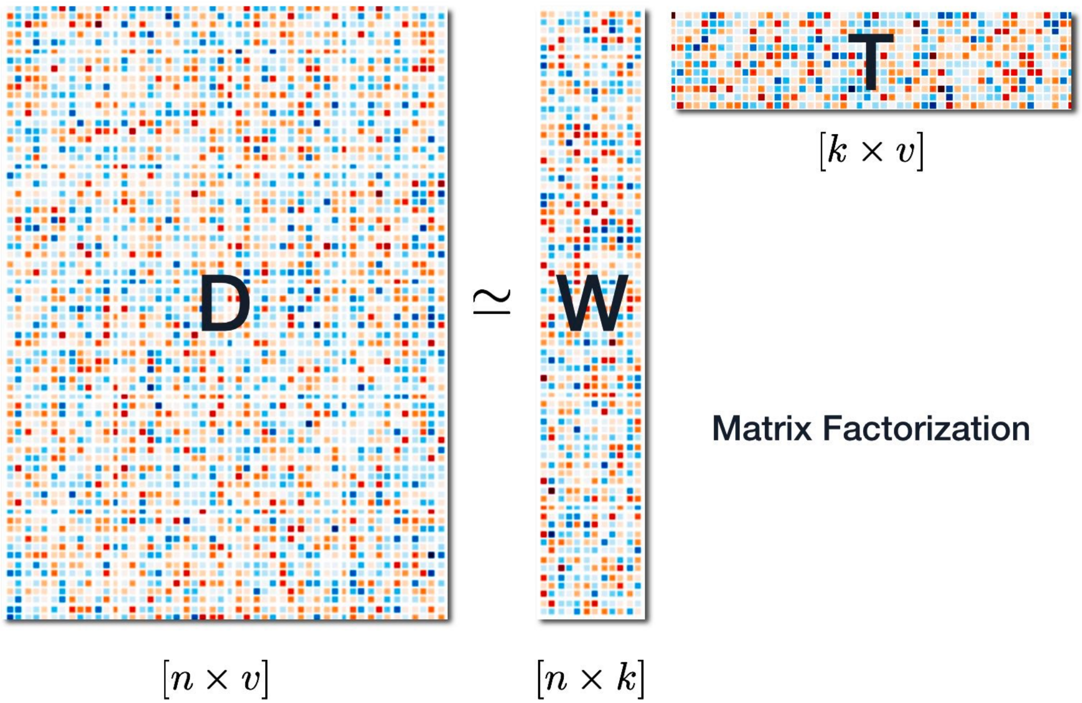
Classify topics into **categories** (food, ambiance, service etc…)

Map topics back to restaurants

For each restaurant and a topic, surface **relevant reviews**

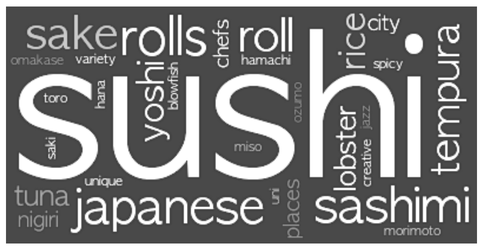
We applied non-negative matrix factorization to learn topics …

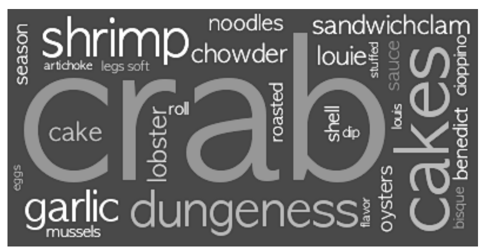
- stopword removal
- vectorization
- TFIDF
- NNMF



Matrix Factorization

… and obtained a pretty clean set of topics



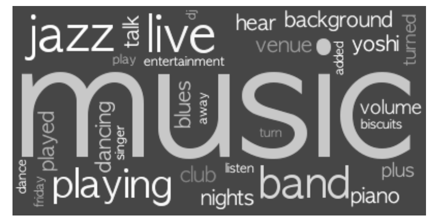


Topics fell nicely into categories

Food

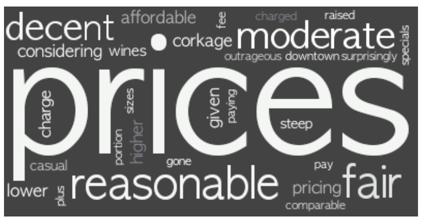
Drinks

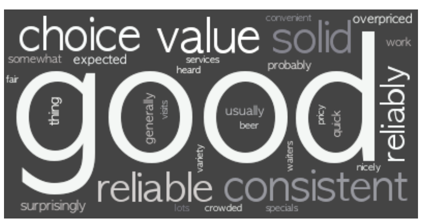
Ambiance



Topics fell nicely into categories

Value





Service

Occasions

Topics

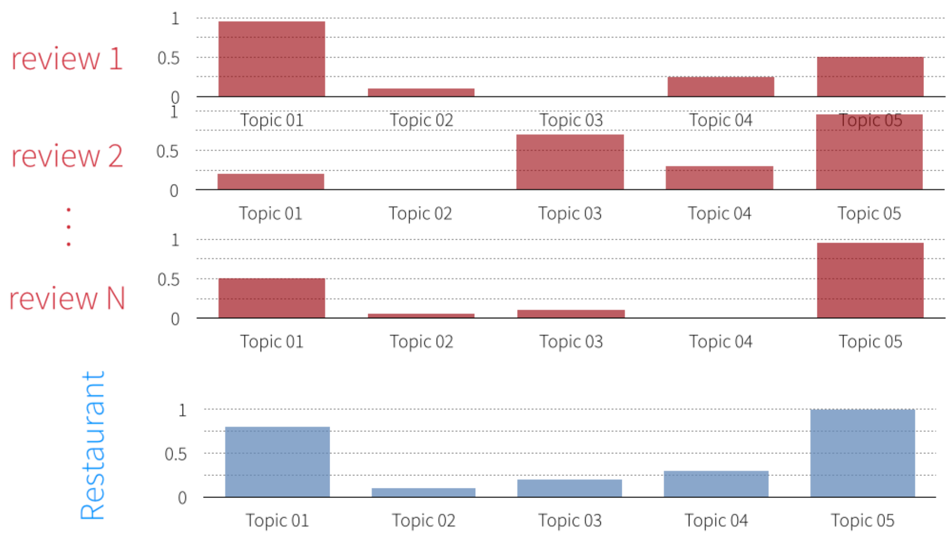
to

Insights

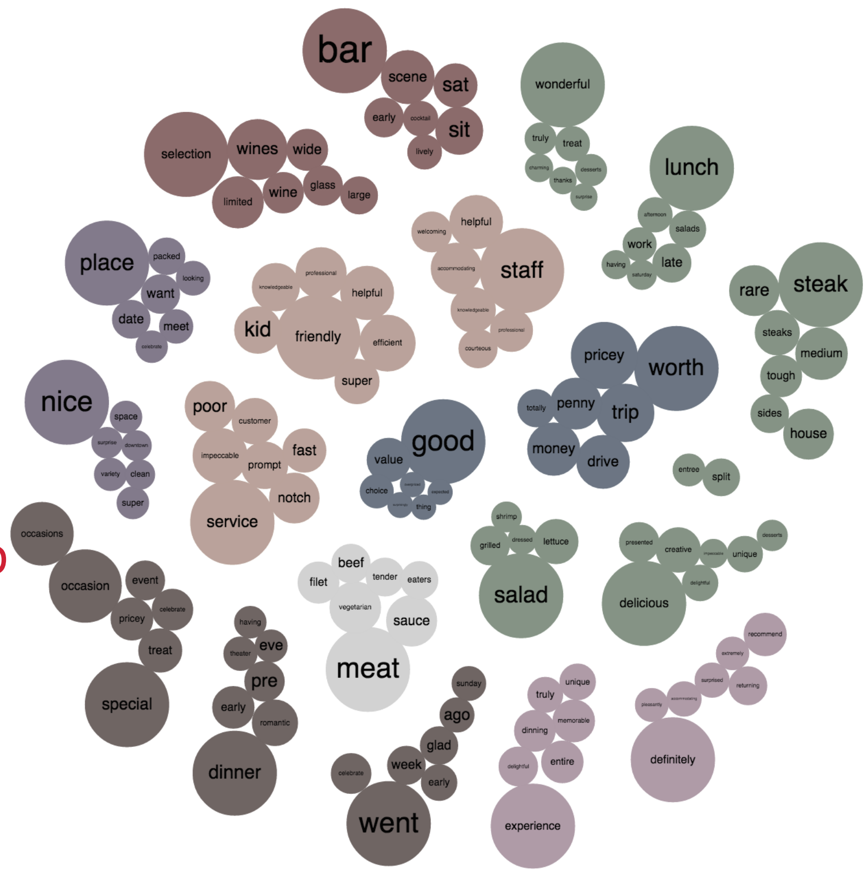
Our topics reveal the unique aspects of each restaurant without having to read the reviews …

Each **review** for a given restaurant has certain **topic distribution**

Combining them, we identify the **top topics for that restaurant.**



Looking at the topics and the top reviews associated with it , we know **Espetus Churrascaria** is not just about meat and steak, but has good salad as well! The service is top notch, its kid friendly, and people go for special occasions, …



Espetus Churrascaria - San Francisco

$50 and over

SF Bay Area

★★★★☆

If you like **meat,** you should eat here

The selection of **meat** was great and so was the service.

Great for **meat** lovers - a **meat** palace and a great, fun experience. Be sure to come hungry and not be in a hurry. The **meat** is delicious, varied, and plentiful - great for a special occasion.

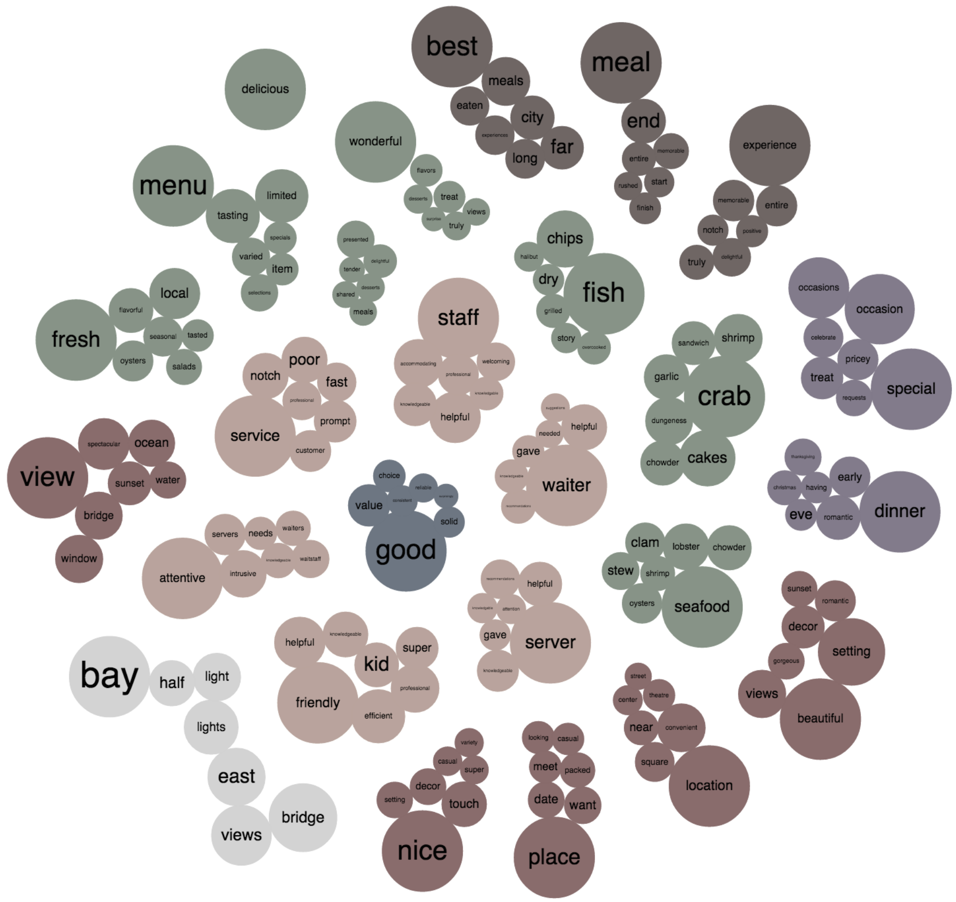
good for lots of steak and **meat**

I took my boyfriend here on his birthday. It was a great experience considering he is a **meat** lover. He kept commenting on how well done the **meat** is cooked. I was very happy with the salad bar since I need something more than just **meat. The** waiters were very attentive and helpful. They also brought him ice cream and a candle because it was his birthday. I really recommend taking your beloved ones to this place if they are **meat** lovers.

The **meat** is excellent. The bacon wrapped **filet** was wonderful. The buffet is also great. It is just a **meat** eater's delight.

love this place, and the **meat** servers are so friendly!

… while **Fog Harbor Fish House** is not just about fish, crab and local seafood, but boasts an excellent view of the bay, specially during sunset!



Fog Harbor Fish House

$30 and under

SF Bay Area

★★★★☆

GREAT waterfront dining on San Francisco **Bay.........**

Nice time, good food. Great view of the **bay** if you don't have fog.

Busy restaurant as you might imagine, overlooking the **bay** and the seals. Sat by the window and experiences sunset over the **bay.** Swordfish and filet mignon wonderful, and friendly staff even under pressure

Fabulous **views** over the **bay.** The food was ver well prepared and tasty, and the service was attentive. Would return.

The food was great and the server friendly, but it's all about the setting. The Pier 39 scene on one side and the **bay, bridge** and boats on the other seal the deal here.

The restaurant was very nice and had wonderful **views** of the **bay.** The food was Devine and would recommend the restaurant to anyone who likes seafood.

Despite being in touristy area the restaurant was quite the opposite. The food was great with beautiful view of the **bay.**

WORK IN PROGRESS

Including everything + context: Factorization Machines

Feature vector **x**

	User				Movie					Other Movies rated					Time	Last Movie rated					Target y	
	A	B	C	...	TI	NH	SW	ST	...	TI	NH	SW	ST	...	Time	TI	NH	SW	ST	...		
$\mathbf{x}^{(1)}$	1	0	0	...	1	0	0	0	...	0.3	0.3	0.3	0	...	13	0	0	0	0	...	5	$y^{(1)}$
$\mathbf{x}^{(2)}$	1	0	0	...	0	1	0	0	...	0.3	0.3	0.3	0	...	14	1	0	0	0	...	3	$y^{(2)}$
$\mathbf{x}^{(3)}$	1	0	0	...	0	0	1	0	...	0.3	0.3	0.3	0	...	16	0	1	0	0	...	1	$y^{(2)}$
$\mathbf{x}^{(4)}$	0	1	0	...	0	0	1	0	...	0	0	0.5	0.5	...	5	0	0	0	0	...	4	$y^{(3)}$
$\mathbf{x}^{(5)}$	0	1	0	...	0	0	0	1	...	0	0	0.5	0.5	...	8	0	0	1	0	...	5	$y^{(4)}$
$\mathbf{x}^{(6)}$	0	0	1	...	1	0	0	0	...	0.5	0	0.5	0	...	9	0	0	0	0	...	1	$y^{(5)}$
$\mathbf{x}^{(7)}$	0	0	1	...	0	0	1	0	...	0.5	0	0.5	0	...	12	1	0	0	0	...	5	$y^{(6)}$

• Rendle (2010) www.libfm.org

Sentiment

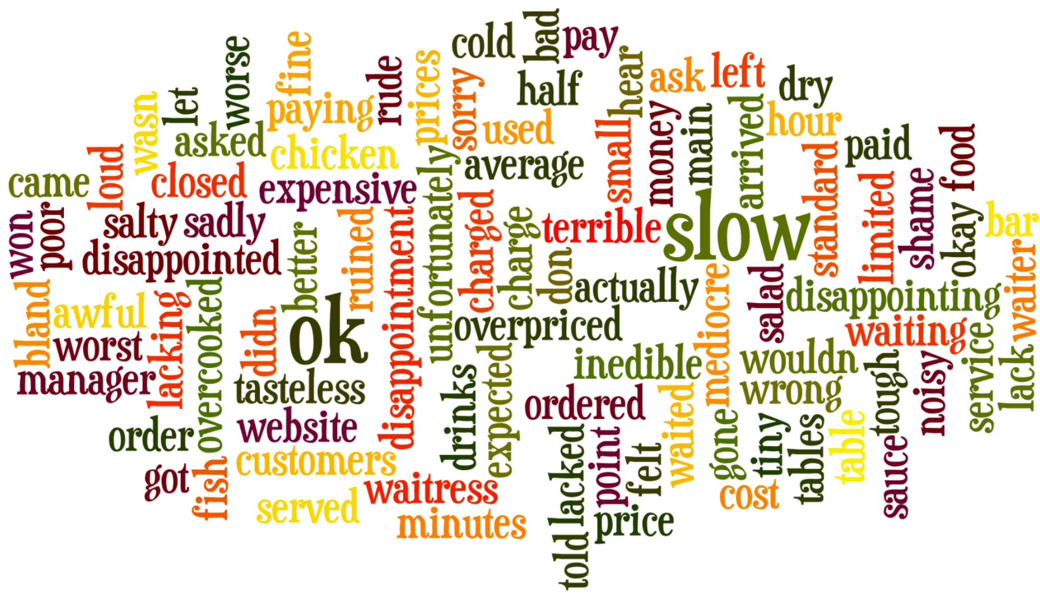
&

Snippets

Sentiments - we use ratings as labels for positive and negative sentiments

Ingredients of a stellar experience

Sentiments - we use ratings as labels for positive and negative sentiments



Ingredients of a terrible experience

Training on Spark

```
// Configure an ML pipeline
val tokenizer = new Tokenizer().setInputCol("text").setOutputCol("words")
val hashingTF = new HashingTF().setInputCol(tokenizer.getOutputCol).setOutputCol("features")
val logisticR = new LogisticRegression().setMaxIter(10)

val pipeline = new Pipeline().setStages(Array(tokenizer, hashingTF, logisticR))

val crossval = new CrossValidator()
  .setEstimator(pipeline)
  .setEvaluator(new BinaryClassificationEvaluator)

val paramGrid = new ParamGridBuilder()
  .addGrid(hashingTF.numFeatures, Array(100000, 150000))
  .addGrid(lr.regParam, Array(0.2, 0.1, 0.01))
  .build()
crossval.setEstimatorParamMaps(paramGrid)
crossval.setNumFolds(5)

// Run cross-validation, and choose the best set of parameters.
val cvModel = crossval.fit(training)
```

```
> display(result_phrases)
```

text	probability	prediction
The lamb was moist and tasty	{"type":1,"size":null,"indices":null,"values":[0.42399680426159214,0.5760031957384079]}	1
Swordfish was too salty	{"type":1,"size":null,"indices":null,"values":[0.6127395325547409,0.38726046744525916]}	0
The place was too noisy for me and my girlfriend	{"type":1,"size":null,"indices":null,"values":[0.5485815988019299,0.4514184011980701]}	0
The service was too rushed	{"type":1,"size":null,"indices":null,"values":[0.5708101545291011,0.4291898454708989]}	0
The lobster was melt in your mouth tender	{"type":1,"size":null,"indices":null,"values":[0.37750851234152727,0.6224914876584727]}	1
It went downhill from there	{"type":1,"size":null,"indices":null,"values":[0.5897037184348852,0.41029628156511483]}	0

We are using nlp+sentiments to surface relevant snippets around tags

dish	reviews	candidate sentences	proposed snippet
Goat Stew	• Both Evvia in Palo Alto and Kokkari in San Francisco share the same menu but the size of this location allows that wonderful bar and rotisserie in the front section. The menu has all the Greek basics and you can make a meal out of the appetizers- which are not so basic. The seasonal menu was exceptional as well. Lamb is an exceptional surprise as you might expect. I tried the **goat stew** this time and it was so delicious. I don't know how many times I've been here but it is my go to place in a city that has many wonderful options. It is a beautiful place with great service and exceptional food. Take visitors. • We took the advice of locals and Fodors to make sure we came here while in San Francisco. We were not disappointed. My husband and I both loved our meals (the **goat stew** was particularly good) and the wine. Our waiter - George - was fabulous. We loved the ambiance too. Recommend asking for a table in the back room, we the front seemed pretty noisy and crowded. • My husband and I enjoyed several appetizers (3 because we couldn't decide on just 2!) and all were great - the beets, the brussel sprouts and the the stuffed calamari. He had the fresh halibut which was excellent and I had the **goat stew** which was wonderful then and just as good for lunch the next day! The chocolate was rich and yummy and the special dessert that I had was excellent too; I'm not sure you can go wrong there. The service was very friendly and attentive without being intrusive. I'm really looking forward to going back ... • I was given a gift certificate to Kookari this Christmas, and decided to go when my sister-in-law was in town. We had an amazing white wine from Santorini -- an assortment of delicious appetizers, a yummy **goat stew**, and really well prepared sole. We will definitely return. • My boyfriend and I came here to celebrate his birthday while vacationing in San Francisco. I had never been to a Greek 'fine-dining' restaurant before and was excited to see what the menu had to offer. We had the prawn appetizer--delicious! For mains, I had the vegetable ravioli and my boyfriend had a special, a **goat stew**. While my dish was good (and I would definitely recommend to vegetarians), it didn't even compare to the goat stew which was absolutely outstanding! The only thing better than the food was the service, which was top notch. I can't wait to return the next time I'm in the bay area.	• I tried the goat stew this time and it was so delicious. (senitment: 1.00, other tags: 0) • My husband and I both loved our meals (the goat stew was particularly good) and the wine. (senitment: 0.70, other tags: 0) • He had the fresh halibut which was excellent and I had the goat stew which was wonderful then and just as good for lunch the next day! (senitment: 0.60, other tags: 1) • We had an amazing white wine from Santorini -- an assortment of delicious appetizers, a yummy goat stew, and really well prepared sole. (senitment: 0.36, other tags: 0) • For mains, I had the vegetable ravioli and my boyfriend had a special, a goat stew. (senitment: 0.36, other tags: 0) • We tried the prawns, the duck stuffed grape leaves, the roasted eggplant, the goat stew and the grilled whole fish. (senitment: 0.20, other tags: 1)	... I tried the goat stew this time and it was so delicious. I don't know how many times I've been here but it is my go to place in a city that has many w ...

The model knows that “to die for”, “crispy”, “moist” are actually indicative of positive sentiment when it comes to food!

- The lobster and avocado eggs Benedict are **to die for**.
- We finished out meal with the their blackberry bread pudding which was **so moist and tasty**.
- The pork and chive dumplings were **perfectly crispy and full of flavor**.
- I had the Leg of Lamb Tagine and it was **"melt in-your-mouth"** wonderful.
- … we did our best with the **scrumptious** apple tart and creme brulee.
- My husband's lamb porterhouse was a novelty and **extremely tender.**
- We resisted ordering the bacon beignets but gave in and tried them and were glad we did---**Yumm**! …

Some early exploration with Word2Vec


Find synonyms for “cioppino”
bouillabaisse
muscles
diavalo
linguini
clams
mussels
diavlo
pescatore
risotto
linguine
pescatora
seafood
rissoto
diabolo
mussles
ciopino
swordfish
mussel
fettuccine
gumbo
brodetto
ciopinno
capellini
cockles
langostines
cannelloni
rockfish
bisques
diavolo
cockle
stew
shrimp
prawns
fettucine
cardinale
bouillabaise
pasta
jambalaya
chippino

Wine pairings with Word2Vec !

Halibut: Chardonnay

Lamb: ?

Wine pairings with Word2Vec !

Halibut: Chardonnay
Lamb: Zinfandel

Translating restaurants via concepts

Harris'
Steakhouse in
Downtown area

$\vec{v}(\text{Harris'}) + \vec{v}(\text{patio})$

$\vec{v}(\text{Harris'}) + \vec{v}(\text{jazz})$

$\vec{v}(\text{Harris'}) + \vec{v}(\text{scenic})$

Patio at Las Sendas
Steakhouse with amazing patio

Broadway Jazz Club
Steakhouse with live jazz

Celestial Steakhouse
Steakhouse with a view

A restaurant like your favorite one but in a different city.

Find the “synonyms” of the restaurant in question, then filter by location!

Downtown upscale sushi experience with sushi bar

San Francisco

Akiko’s, SF

$\simeq$

Maui

Sansei Seafood Restaurant & Sushi Bar, Maui

Chicago

Masaki Sushi Chicago

New York

Sushi of Gari, Gari Columbus, NYC


keep in touch
@pablete
@datamusing
thank you!
CONFIDENTIAL