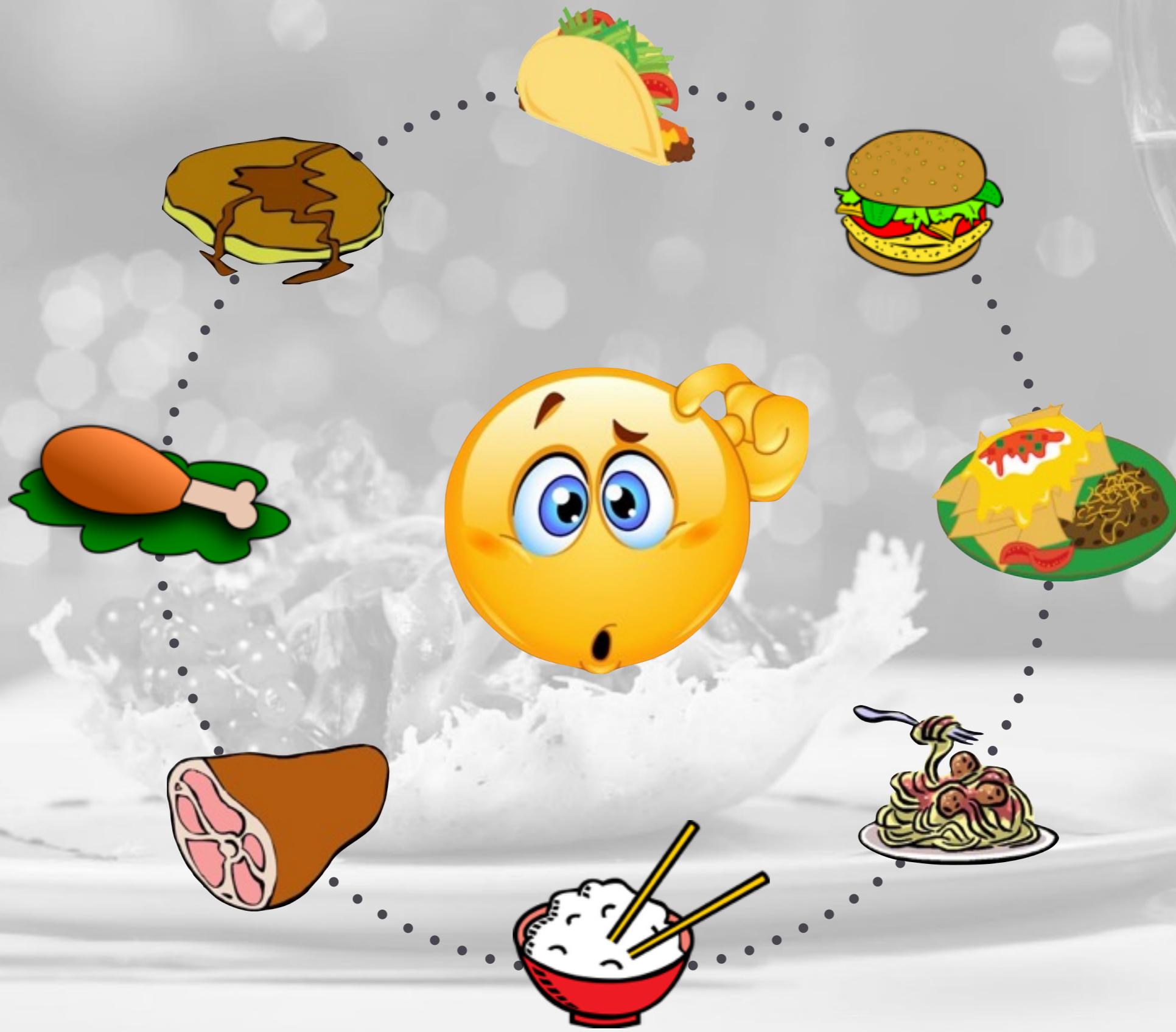


FOODIE FAVORITES

Vineet Abraham



PROBLEM



DATA SCIENCE SOLUTION – PART 1

Named Entity Recognition: Used a Conditional Random Field algorithm to determine whether text is **food** or not food.

Example Review:

- The **Butter Chicken** was delicious! The **Plain Naan** was good but the **Onion Naan** was even better. The **Mango Lassi** was also amazing and one of the better ones we have tried recently.

DATA SCIENCE SOLUTION - PART 1

Text Featurization

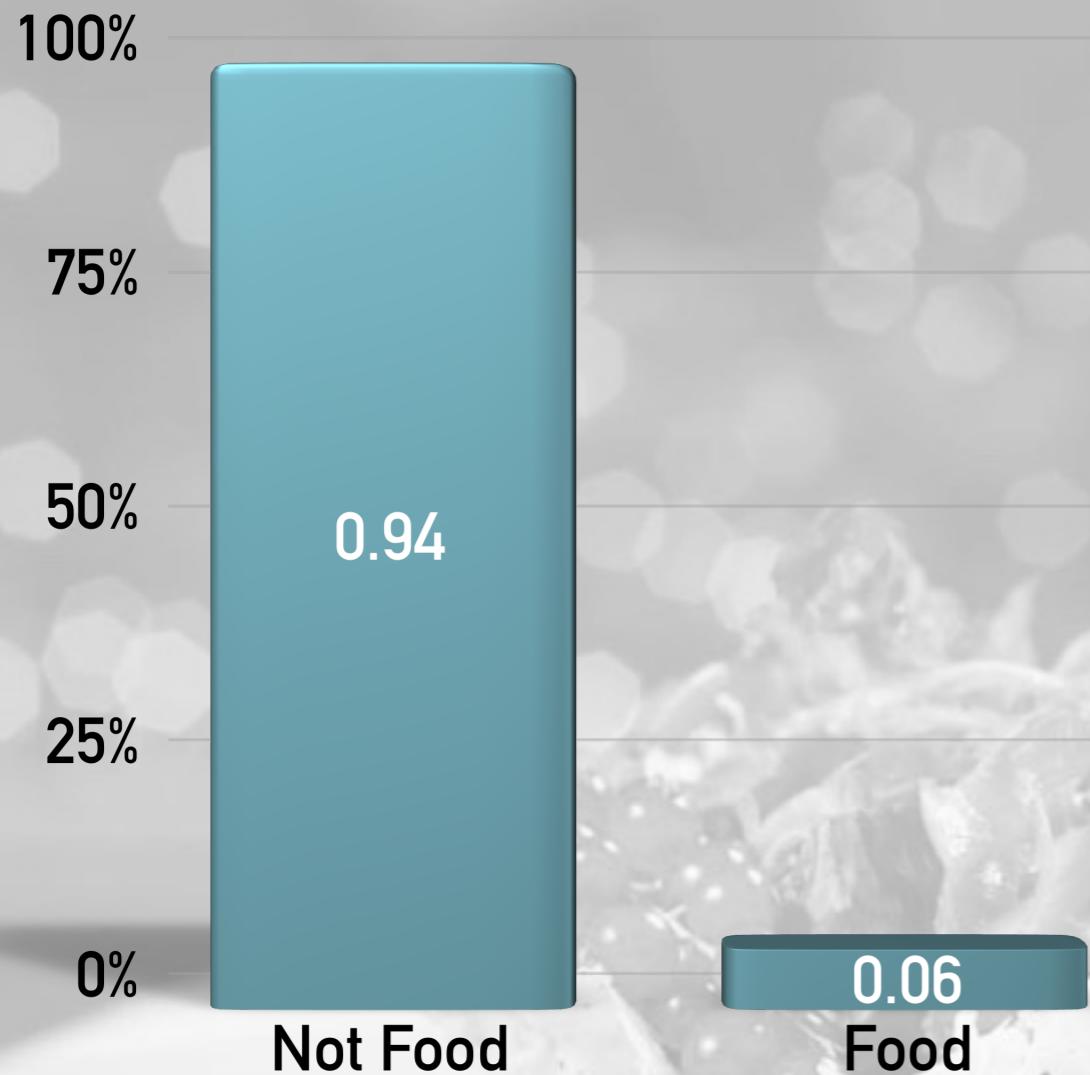
"The Butter Chicken was delicious!"

	The	Butter	Chicken	was	delicious	!
word (lc)	the	butter	chicken	was	delicious	!
POS	DT	NN	NN	VBD	JJ	!
is title?	True	False	False	False	False	False
is capital?	True	True	True	False	False	False
is digit?	False	False	False	False	False	False
is punct?	False	False	False	False	False	True
is EOS?	False	False	False	False	False	True

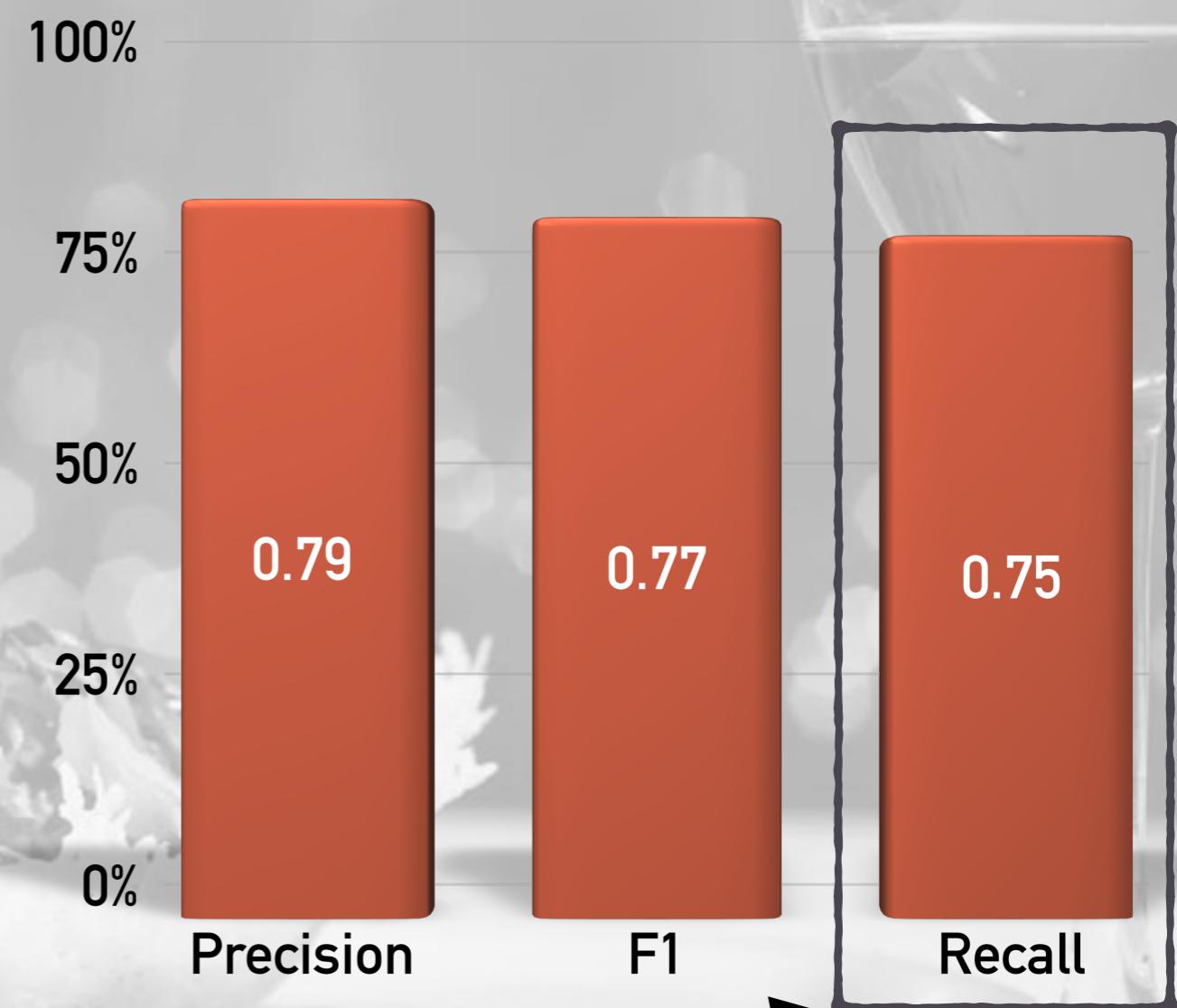
***Not an exhaustive list of features; POS - Part of Speech; EOS - End of Sentence*

DATA SCIENCE SOLUTION – PART 1

Text Distribution



Model Scores



Most important metric for this model - accurately classifies sentences that contain “food” mentions 75% of the time

DATA SCIENCE SOLUTION – PART 2

String Matching: Determined if any of the strings classified as **food** were menu items using a combination of the following Levenshtein distance methods:

- Partial String Ratio
- Token Set Ratio

Example:

My favorite dish was the **Wild Rice Salad with Salmon Pork.**

Menu Item	Match?
Rice	No
Black Rice Topped with Ice Cream	No
Wild Rice Salad with Salmon Pork	Yes

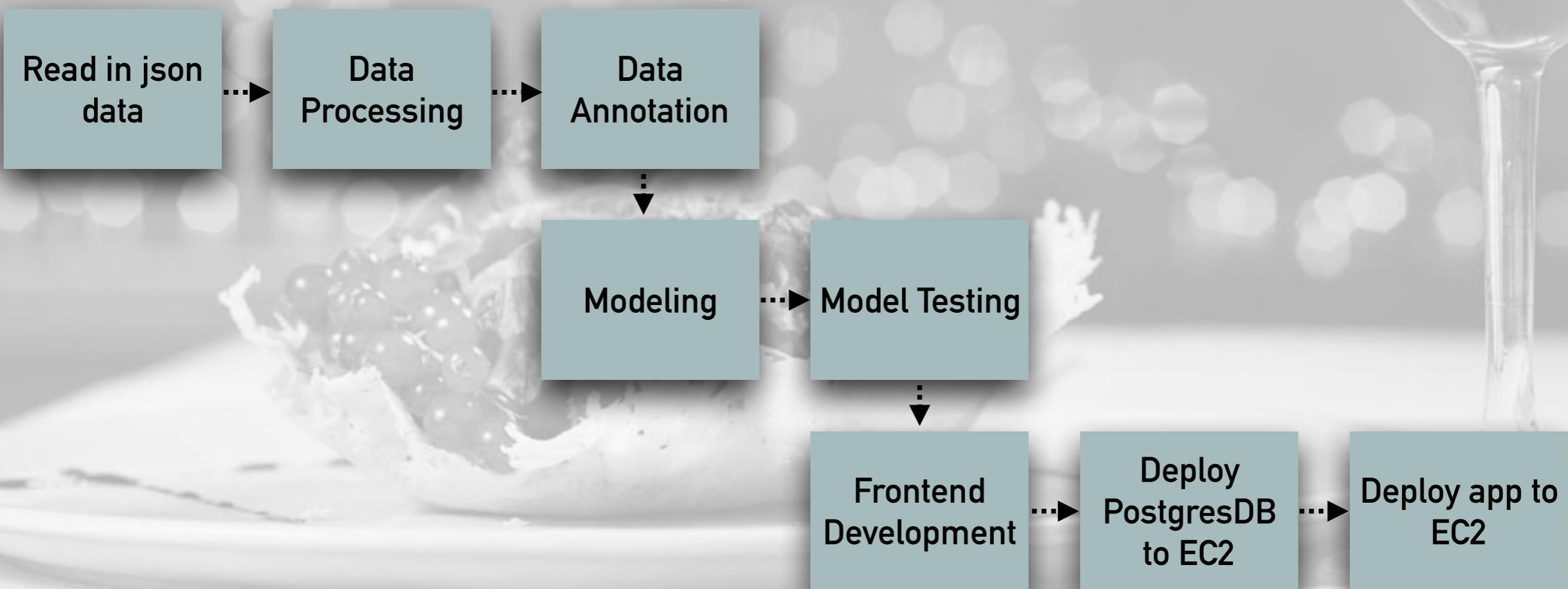
PROCESS FLOW



PostgreSQ^L



Bootstrap





Foodie Favorites

This application uses Natural Language Processing on Yelp reviews to help users find out what dishes are most popular.

To begin, select a City

City ▾





Foodie Favorites

Select a Restaurant to see what's popular

Restaurant ▾

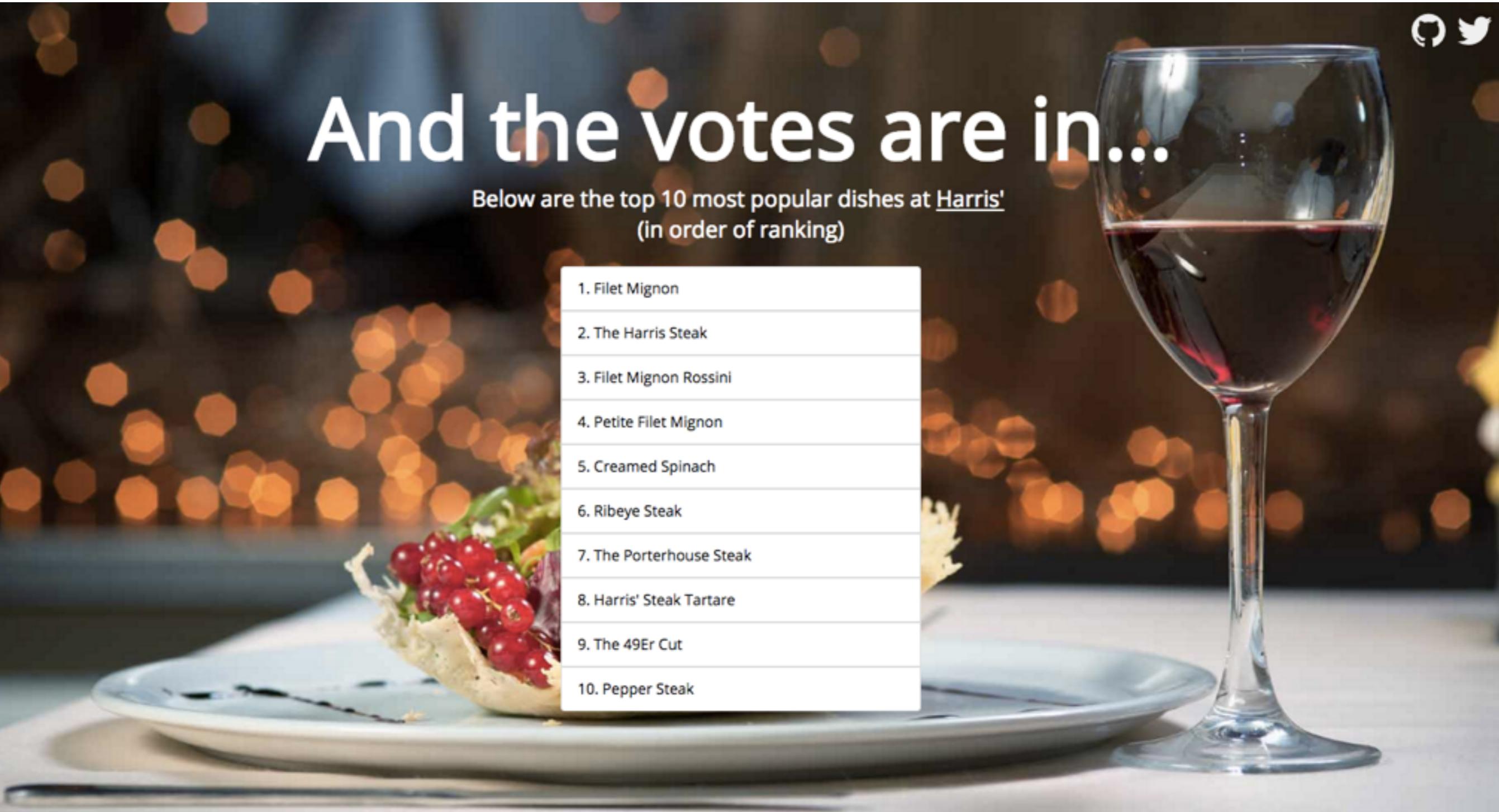




And the votes are in...

Below are the top 10 most popular dishes at [Harris'](#)
(in order of ranking)

1. Filet Mignon
2. The Harris Steak
3. Filet Mignon Rossini
4. Petite Filet Mignon
5. Creamed Spinach
6. Ribeye Steak
7. The Porterhouse Steak
8. Harris' Steak Tartare
9. The 49Er Cut
10. Pepper Steak



FUTURE WORK

- Use sentiment analysis to include text from more than just positive reviews.
- Separate rankings by part of menu (appetizers, lunch, dinner, dessert, etc.).
- Build crawler to parse restaurants on demand.
- Try and improve algorithm performance.

THANK YOU!

www.foodiefavorites.co

github.com/vabraham/foodie_favorites

vabraham24@gmail.com

Vineet Abraham