Appetite for Destruction

Recipe for Disaster

Michael Schillawski, 10 April 2018
General Assembly, Data Science Immersive

Recipe to Rot

It’s Sunday morning. You’re in your kitchen, flipping through a cookbook, looking for what to make for the week.

You find a recipe, perfect.

But there’s a problem. You think a couple hours ahead, after you make that recipe. Crap, you didn’t use the entire quantity of raw ingredients that you’re going to buy at the store for this recipe.

Guess they’re going to sit in the fridge until you throw them out at the end of the week after they rot.

Now you also have to clean up the mess.

Just how big a problem are we talking?

- In 2016, the EPA found that Americans throw away $1,600 of waste food per year for a family of 4[1]
- In 2014, Americans disposed of more than 38 million tons of food waste, 95% of which ended up in landfills or combustion facilities[1]
- In 2010, the USDA ERS found consumer-losses of 90 billion pounds of the available food supply (21 percent of the available food supply)[2]
- Globally, UNFAO estimates nearly one-third of all food grown is lost or wasted, valued at nearly $3 trillion[1]

[1] https://www.theatlantic.com/business/archive/2016/07/american-food-waste/491513/

[2] https://www.ers.usda.gov/webdocs/publications/43833/43680_eib121.pdf

When Life Gives You Lemons…
… Make Lemonade and Lemon Meringue Pie

What if, instead, you could find other recipes that were similar to that one perfect recipe?

Some system would recommend other recipes that share the same ingredients -- so instead of watching those ingredients rot in the back of the fridge, you could make a second meal with those leftovers.

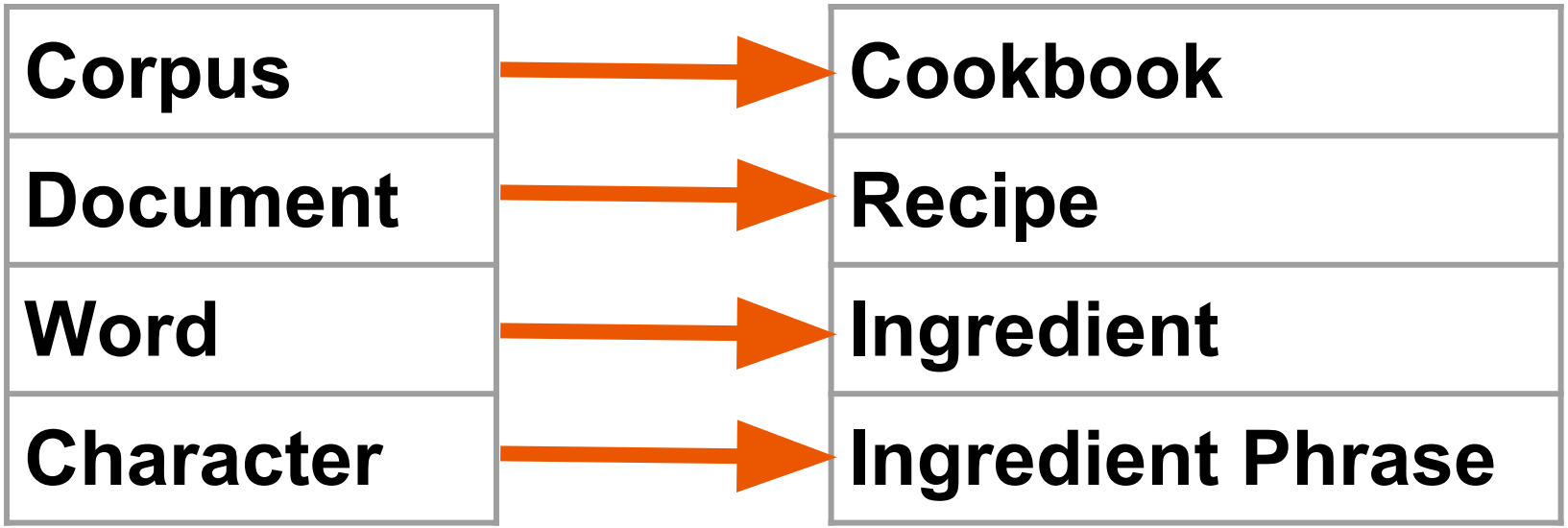
Of course, you wouldn't want to find an exact match for that perfect recipe. Don't want to make the same thing twice.

Machine to Table

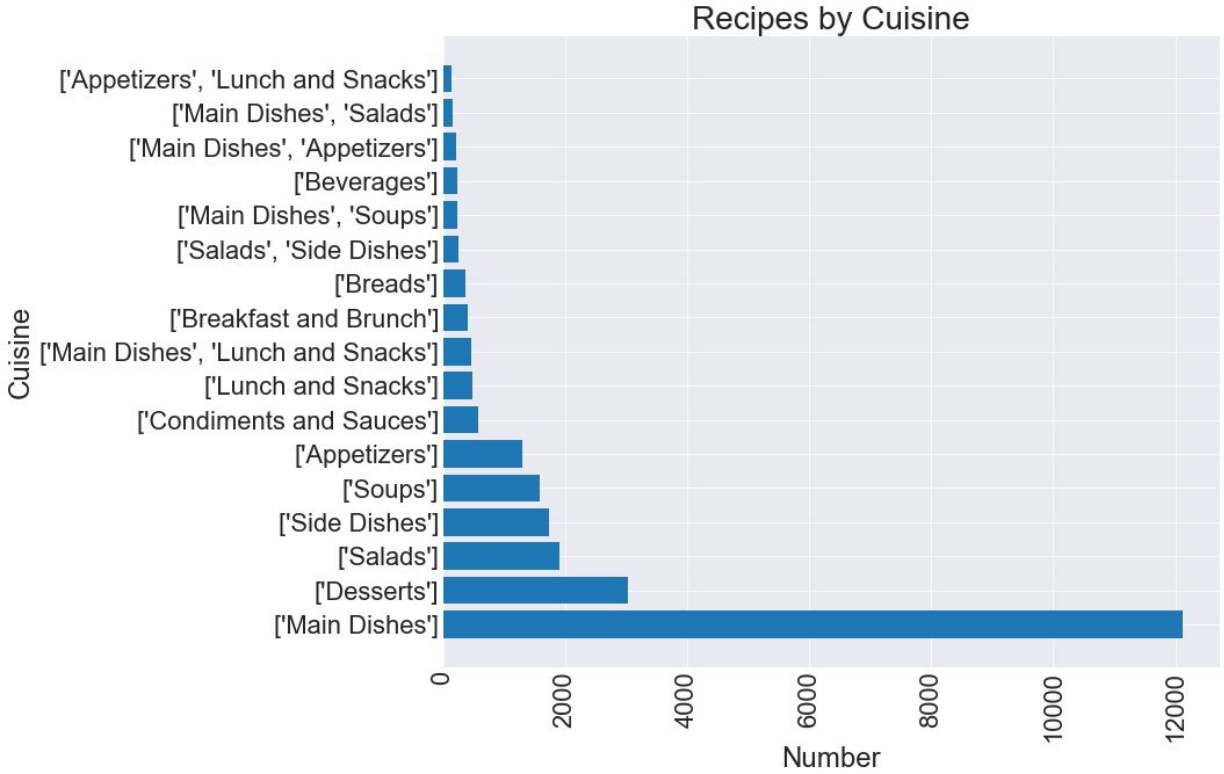
Solution: develop a recommendation system that finds *similar* recipes, based on ingredient commonality, to minimize food waste and develop a comprehensive shopping list.

Data: get a dataset of recipes, from a cooking site - like Yummly - to build a database of recipes and ingredients. Build a dynamic classifier that takes an input recipe and, after subtracting basic ingredients, find recipes that share core ingredients but are not carbon copies.

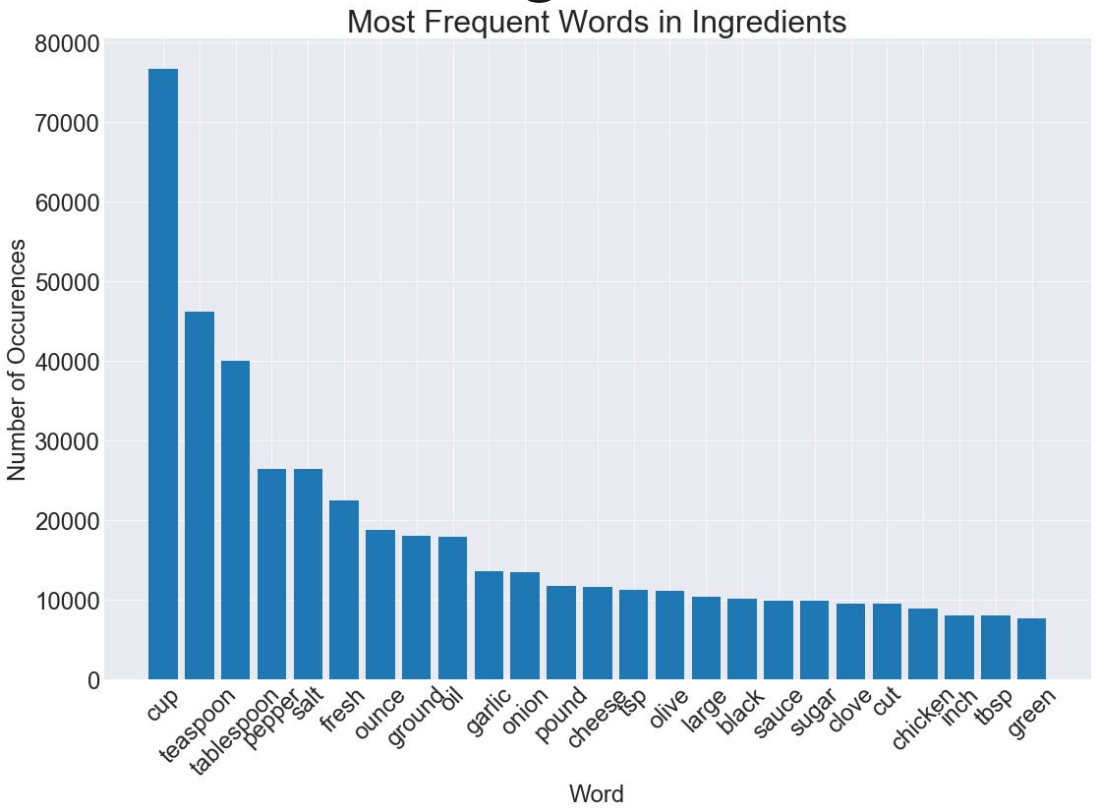
Cookbooks, Recipes, Ingredients, Oh My



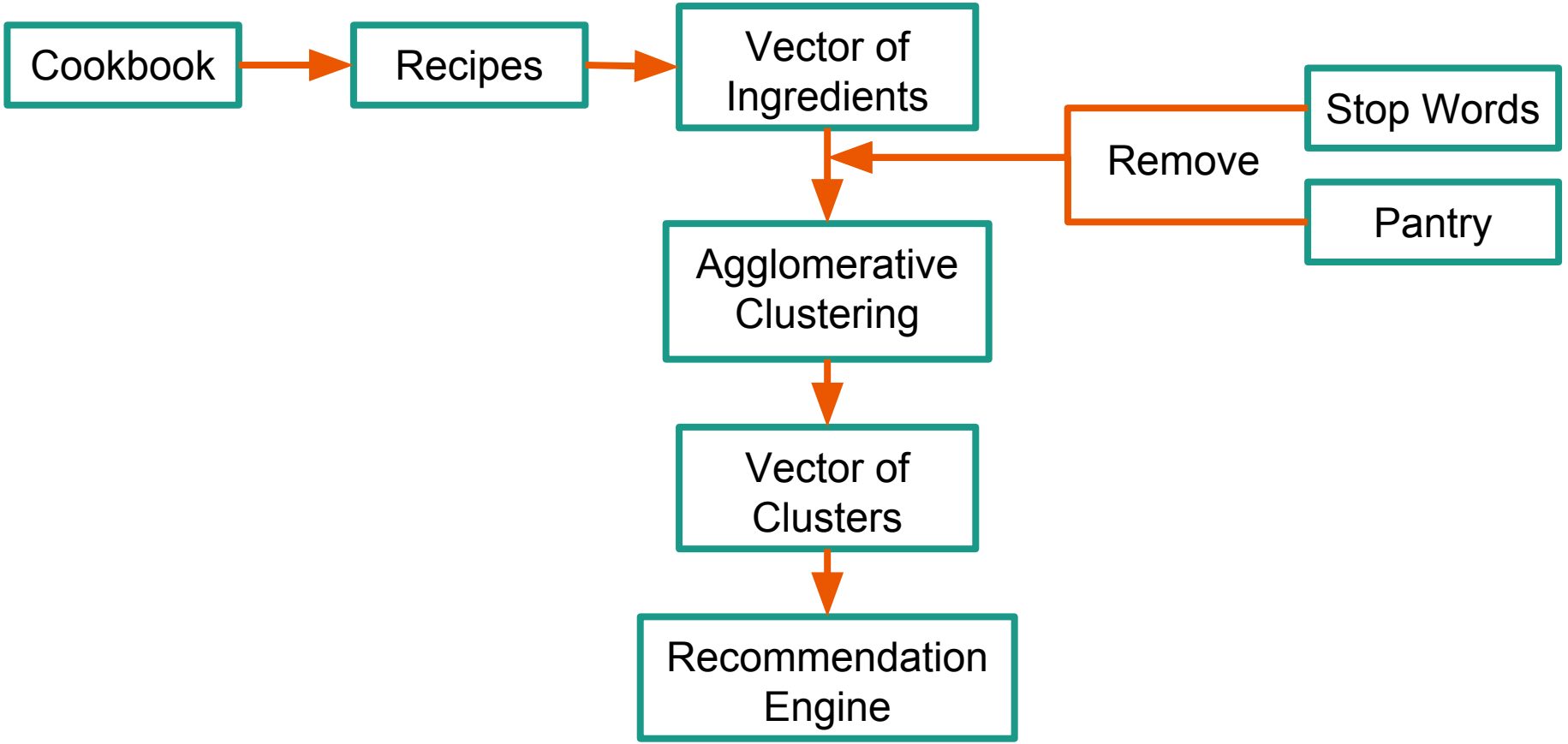
Recipes spanned 221 cuisines and encompassed 78,710 unprocessed ingredients



The most common words that described ingredients did not *actually* describe the ingredients



From Recipe to Recommendation



Step by Step

```
['9 ounces farro perlato (semi-pearled emmer), about 1 1/3 cups uncooked
 '1 teaspoon chia seed or ground flaxseeds',
 '3 medium zucchini, trimmed and halved lengthwise',
 '1/2 red bell pepper, seeded',
 '12-16 ounces cherry tomatoes, halved',
 '2 green onions, green parts only, thinly sliced',
 '2 tablespoons white balsamic or white wine vinegar',
 '1 tablespoon lemon juice',
 '2 teaspoons spicy or Dijon mustard, or to taste',
 '1 large clove garlic, minced or pressed',
 '1/4 teaspoon salt (optional)',
 '1/8 teaspoon freshly ground black pepper',
 '1/4 cup chopped fresh basil',
 'Salt and black pepper, to taste',
 'Additional lemon juice, to taste']
```

Unprocessed Ingredient List

Step by Step

```
[['farro perlato emmer',
  'chia ground',
  'zucchini',
  'bell pepper',
  'cherry tomatoe',
  'green onion green',
  'white balsamic white wine vinegar',
  'lemon juice',
  'spicy dijon mustard',
  'clove garlic',
  'salt',
  'ground black pepper',
  'basil',
  'salt black pepper',
  'lemon juice']]
```

Semi-Processed Ingredient List: No Stop Words

Step by Step

```
[['farro perlato emmer',
  'chia ground',
  'zucchini',
  'cherry tomatoe',
  'green onion green',
  'lemon juice',
  'spicy dijon mustard',
  'clove garlic',
  'basil',
  'lemon juice']]
```

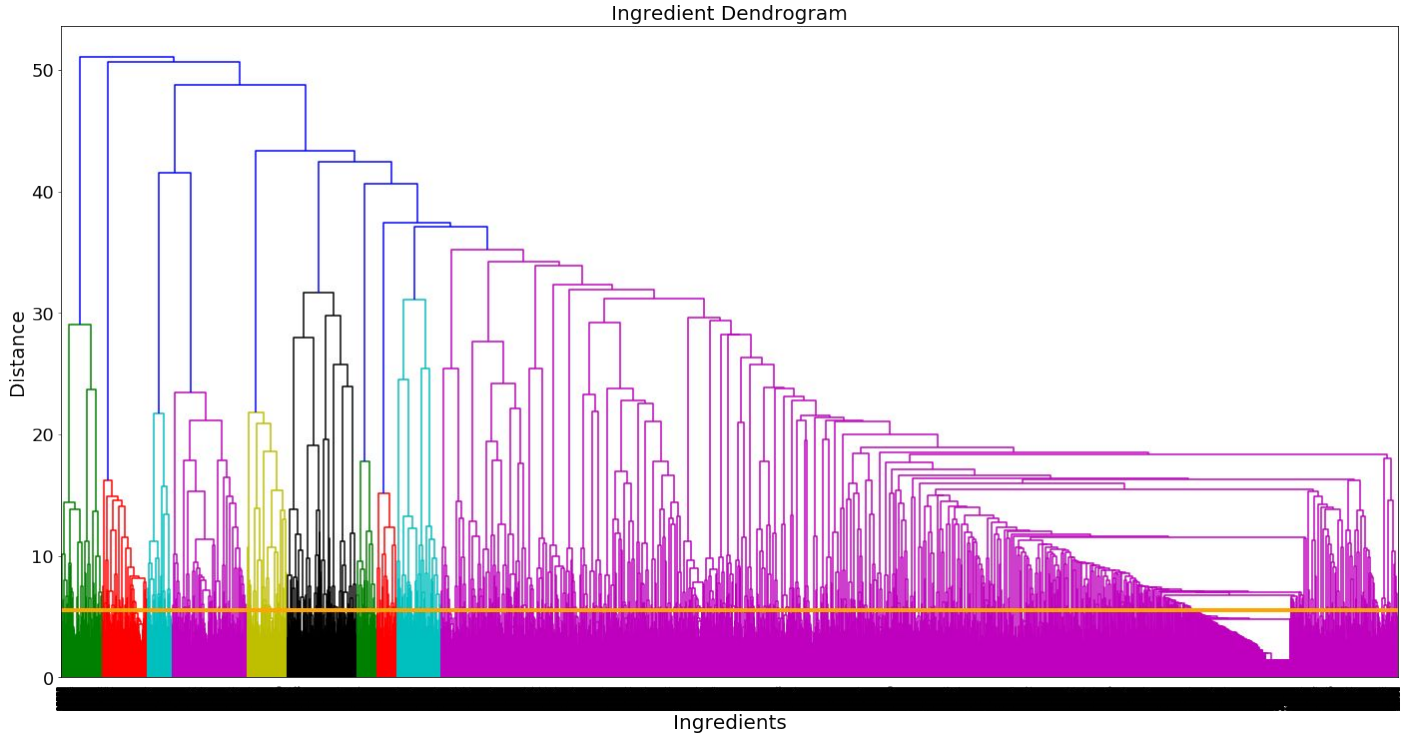
Processed Ingredient List:
No Pantry Items

Step by Step

```
[947, 366, 977, 688, 321, 433, 1021, 549, 436, 433]
```

Clustered Ingredient Vector

What was imagined as a recommendation engine quickly devolved into an unsupervised learning problem



So, what’s cooking?

Future Extensions

- Rebuild the clustering algorithm so that the clusters can be applied to the ingredients of an unknown recipe
- Expand the cookbook
- More robust pantry management: subtract ingredients as they are consumed
- Predictive grocery lists: advise the user when to replace pantry staples as stores are depleted
- Real-time pantry replenishment: Text-recognition for grocery receipts or integration into online grocery platforms.