Recipe Recommendation System

INSPIRED BY *Friends*

THE MOISTMAKER SANDWICH

MAKES 1 SANDWICH, PLUS ENOUGH FILLINGS TO SERVE 10

Monica Geller had a simple but genius idea: nestle a slice of gravy-soaked bread in the middle of a Thanksgiving leftovers sandwich. The result not only drove Ross to near madness, it became the first real windfall for *Binging with Babish* Reddit, in many ways, acts as a tastemaker for the rest of the internet—so when this episode began climbing the ranks of /r/videos, it wasn't long before it was being posted elsewhere, as well as written up by the Huffington *Post* and *Entertainment Weekly*. It was the first time I thought, Okay, maybe I've got something worth pursuing here. I was still working a full-time job (and would continue to for another year), but I resolved to put out an episode every Tuesday at nine a.m. And if the Moistmaker was any indication, this was another full-time job; its construction required the making of an entire Thanksgiving meal from scratch. Freshly tattooed and cooking in my pajamas, day 1 consisted of spatchcocking and dry-brining the turkey, a technique popularized by my personal food idol, J. Kenji López-Alt. Not only does dry-brining deeply flavor the meat and desiccate the skin (making it extra crisp), spatchcocking the bird allows it to cook in nearly half the time it would take whole. Yes, you heard me right, folks—half. Imagine cutting an extra two or so hours out of the whirlwind clusterfuck that is Thanksgiving prep, while being able to make turkey stock from scratch in the process. All you need is a strong pair of kitchen shears.

VERDICT: This recipe yields, well, a full Thanksgiving dinner. Each technique is a thoughtful and reliable, and if you're crazy enough to make the whole thing explicitly for a sandwich, you'll end up with a superlative one. The so-called "Moistmaker" does just that: introduces gravy to the sandwich without making it sloppy, keeping everything moist and flavorful.

3 slices hearty white sandwich bread
Turkey Gravy (recipe follows)
Sliced Roast Turkey and Sage-Sausage Stuffing (recipe follows)
Cranberry Sauce (recipe follows)

LIGHTLY TOAST the bread. Soak one piece of the toast in the gravy. Assemble a triple-decker sandwich with the roast turkey and some of the gravy, some stuffing, and some cranberry sauce, placing the gravy-soaked bread in the middle of the layers. Transfer to a plate and serve.

10

An unsupervised NLP model for discovering food recipes with the ingredients you have on hand

John Villanueva | Thinkful - Data Science | October 2019

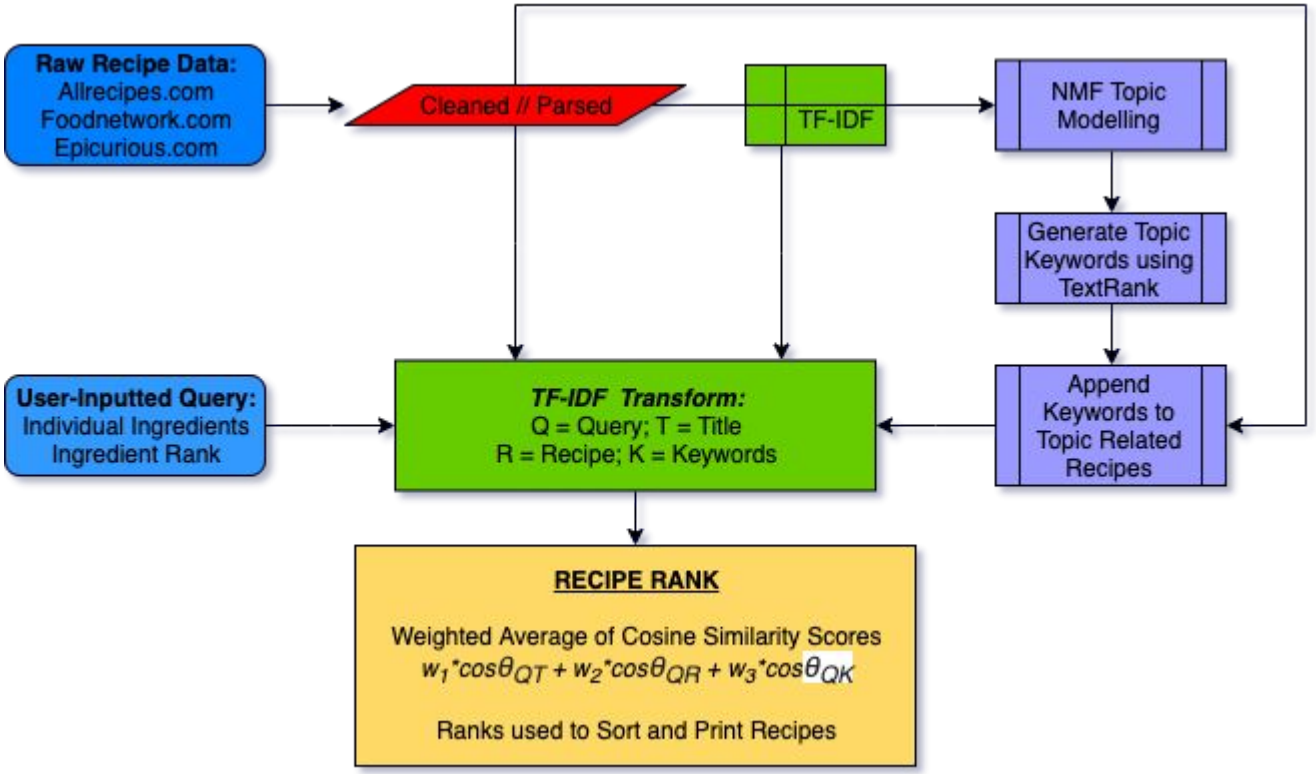
The Problem:
You Don’t Know What You Don’t Know

- Finding New Food Recipes without knowing the names of the dishes
- Searching for recipes that utilize disparate or unlikely ingredient pairings

Who can use it?

- Home cooks, new to the kitchen
- Scrappy college students with limited ingredients
- More experienced cooks seeking inspiration

The Model


Raw Recipe Data:
Allrecipes.com
Foodnetwork.com
Epicurious.com
Cleaned // Parsed
TF-IDF
NMF Topic Modelling
Generate Topic Keywords using TextRank
Append Keywords to Topic Related Recipes
User-Inputted Query:
Individual Ingredients
Ingredient Rank
TF-IDF Transform:
Q = Query; T = Title
R = Recipe; K = Keywords
RECIPE RANK
Weighted Average of Cosine Similarity Scores
$w_1 * cos\theta_{QT} + w_2 * cos\theta_{QR} + w_3 * cos\theta_{QK}$
Ranks used to Sort and Print Recipes

Unique Model Features

Query Ingredient Ranking:

- Optional Ranked Ingredient Search
- Ingredients are scaled by decreasing weights
- Only works with three ingredients or more
- [1], [.5,.5], [.5,.25,.25], [.5,.25,.125,.125]

Keyword Modeling

- Documents within Each Topic used to extract Topic Keywords
- Purpose: Topic Keyword in place of abstractive category assignment

Examples

Query 1:
- Cream
- Cinnamon
- Banana

Query 2:
- Wine
- Butter
- Cilantro
- (Ranked)

Query 3:
- Jelly
- Wine
- (Unlikely)

Query 4:
- Pepper
- Apple
- Pork
- (Unlikely)

Query 5:
- Cream
- Cinnamon
- Banana
- (No Weights)

Query 6:
- None

Query 7:
- Apple
- Blueberry
- (Only Category)

Query 8:
- Japanese
- (Only Category)

Future Improvements

- Dense Word Embeddings (LDA2Vec, Word2Vec, GloVe)
  - Sophisticated Semantic Similarity Scoring
  - Generate Abstractive Keyword Summarizations (Topic Names) for NMF produced topics
- Robust Recipes Stopwords List
- Database
  - Recipes and Document Vectors currently stored as Global Variables
- Updates using custom scrapers, Expand Recipe Dataset
- Run Topic and Keyword generation on Spark in the cloud
- Negative Query Scoring
- Productionize Model Using Flask

Model Tuning and Evaluation

Parameters:

Keyword Extraction on Topics

- NMF vs. LDA for Topic Modeling
- Number of Topics (Arbitrary, 50)
- Number of Documents from which to Extract Keywords (NNMF Elbow Method, 200)
- Neighbor Window Size (Arbitrary, 4)
- Number of Keywords to Extract (Arbitrary, 25)
- Number of Documents to which keywords are assigned (Histograms of NNMF Loadings, 2000)

Search Parameters

- Title Weight
- Instructions Weight
- Keywords Weight

Miscellaneous

- Recipe Stop Words
- Query Ingredient Rank

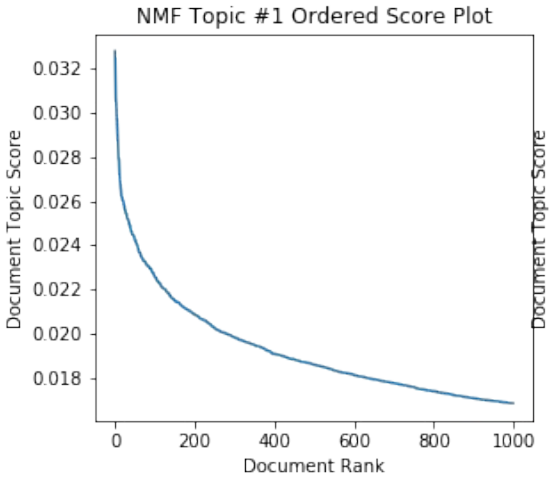
Model Tuning and Evaluation

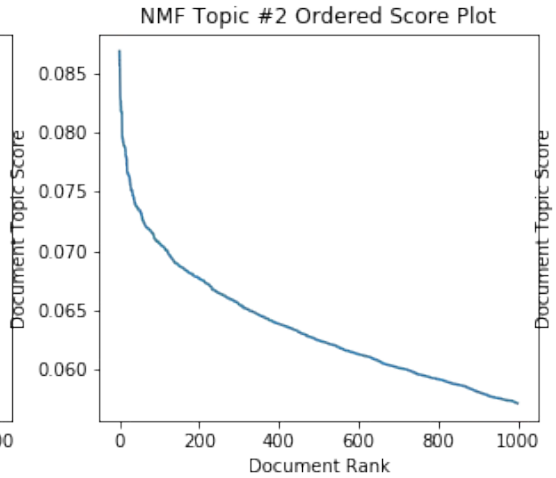
Choosing a Topic Modeling Method:
NMF vs. LDA

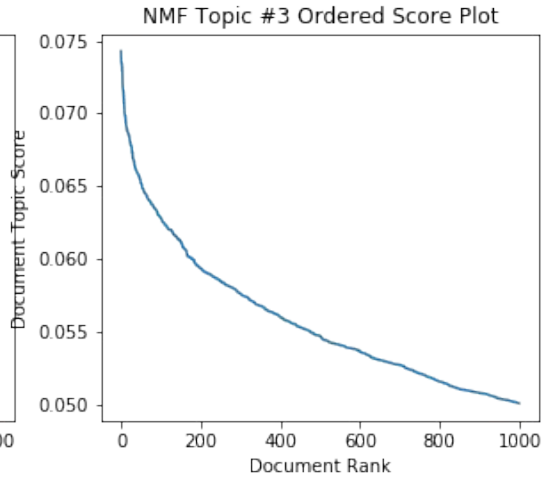
IMPRESSIONS		By Documents	By Words
LDA	Topic 1	Salad	Salad
	Topic 2	Variety Beef/Pork	Pasta
	Topic 3	Cooked Fruit?	Stopwords?
NNMF	Topic 1	Spreads	Dressing
	Topic 2	Cake/Bread	Baking
	Topic 3	Chicken	Chicken

Model Tuning and Evaluation

Choosing the Number of Documents
per Topic from which to Pull Keywords

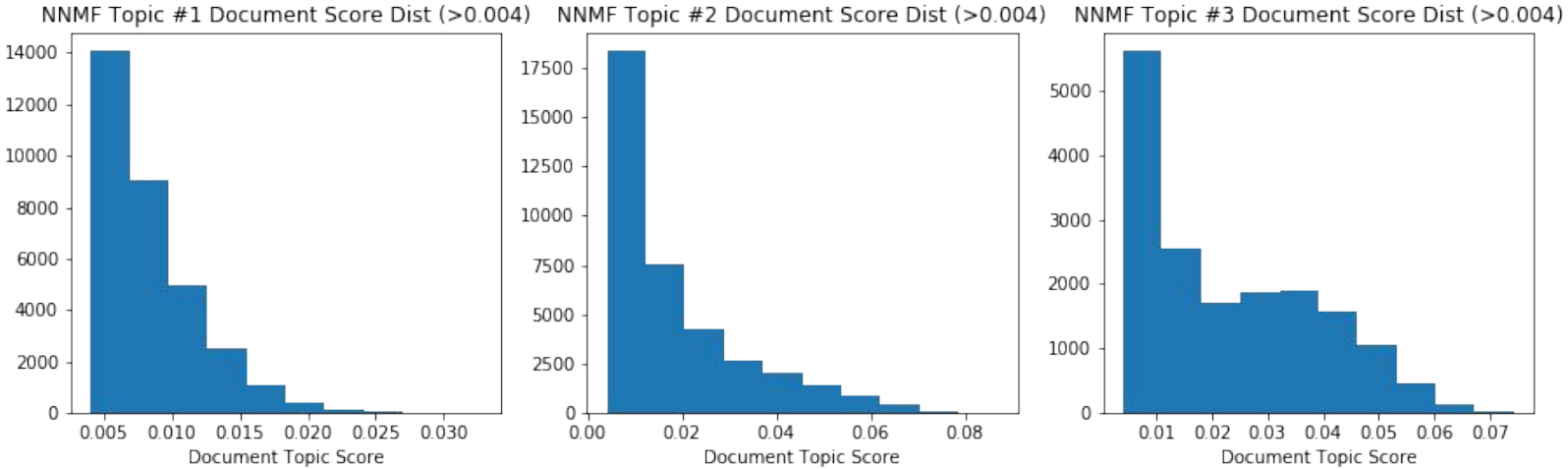






Model Tuning and Evaluation

Choosing the number of documents to which Keywords are assigned



Model Tuning and Evaluation

Peering into the Keywords

Max Number of Recipes with Categories: 50,000 ⇒ At least 70,000 Recipes uncategorized
The querying algorithm penalizes recipes without category keywords.

Model Tuning and Evaluation

Query Similarity Weights
Query Ingredient Rank
Recipe Stopwords

Similarity Weight Defaults:

- W_Title = 0.2
- W_Instructions = 0.3
- W_Keywords = 0.5

Query Ingredient Rank:

- Ex: [Cream, Cinnamon, Banana, Sugar]
- [0.5, 0.25, 0.125, 0.125]

Recipe Stopwords:

['cup', 'cups', 'ingredient', 'ingredients', 'teaspoon', 'teaspoons', 'tablespoon', 'tablespoons'' C', 'F', ‘degree’, ‘degrees’, ‘add’, ‘heat’, ‘minute’, ‘ingredient’, ‘ingredients’]

Stopwords compiled manually as needed