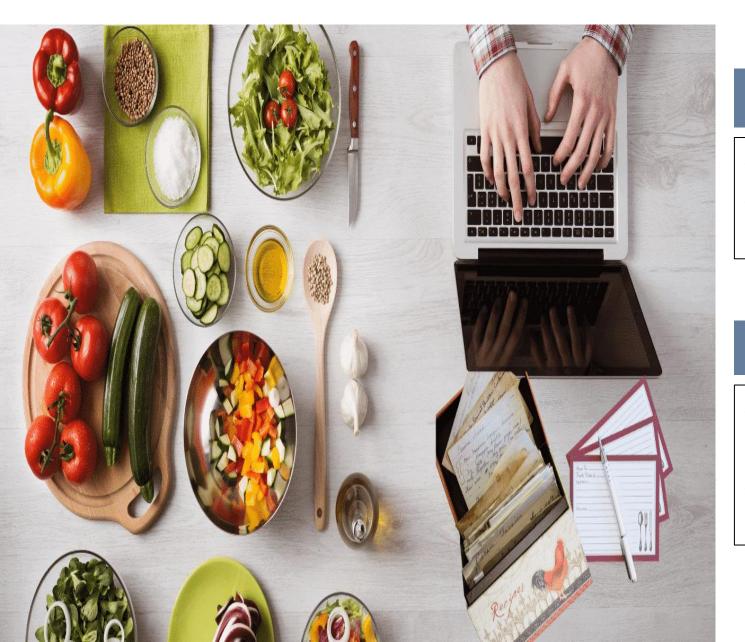
Food Recipes: Topic Modeling & Recommendations

Salma Katri February 23, 2022

Eind a Recipe

Background



Motivation

- Millions of recipes are available online
- We spend a lot of time thinking about what to cook and end up making the same recipes again

Objective

- Identify key ingredients across cuisines and recipe types
- Personalized recommendations for new recipes to try

Data Source



Scraped Food.com data available on Kaggle

	10.00		aset
Keci			aser
11001		MUL	uset

231,637 Recipes

4000 Ingredients

555 Cuisines/Tags

Users Ratings dataset

226,570 Users

1.1M Ratings

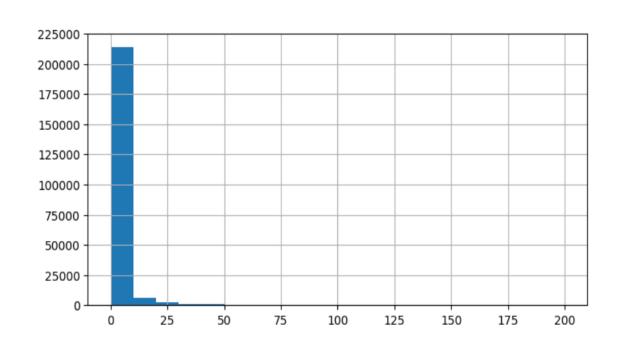
231,637 Recipes

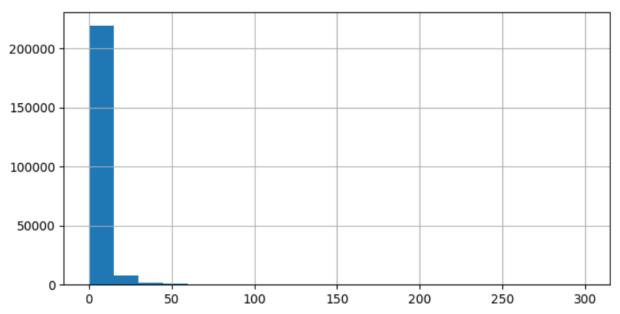
- :



Ratings per User Distribution

Ratings per Recipe Distribution





Methodology

Text Preprocessing

Topic Modeling

Recommender System

- Ingredients text cleaning:
 - Lowercase
 - Removing punctuation
 - Separating each ingredient
- Compound Term Extraction
- Lemmatization using WordNet
- Only keeping nous using speech tagging
- Removing stop words (English plus additional created ones)

- Using CountVectorizer and TF-IDF:
 - LSA
 - NMF
- LDA
- Corex:
 - Unsupervised
 - Semi-supervised

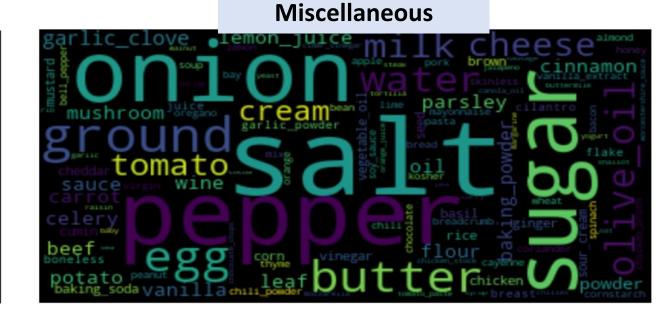
- Content-based Filtering
 - CountVectorizer, LSA
 - LDA
- Collaborative Filtering
- Hybrid: LightFM

Topic Modeling: Unsupervised

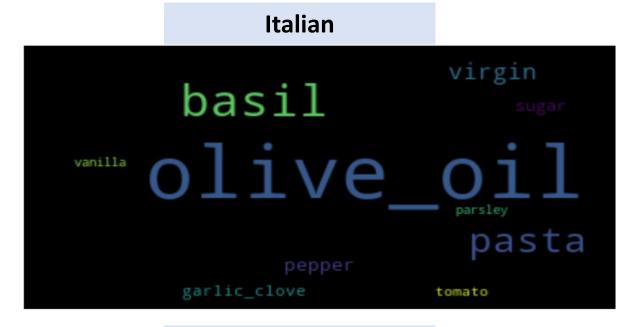




chocolate chicken tomato tortilla cube egg celery cracker mix lowfat powder pied pasta pas



Topic Modeling: Semi-supervised



cheese cheddar sour_cream beef

American

Chinese



Mexican

orange



Recommendation Results

 Texas style skillet All in the kitchen chili Texas beef skillet La bamba casserole CountVectorizer, LSA Collard greens with gruyere Cherry tomatoes on provolone garlic bread Alouette potatoes Olive garden bruschetta Jalapeno red bean bbq burger Left over pork and black bean taco soup All in the kitchen chili Torshi lift Rigatoni with bacon, spinach and cheese LDA Penne with sausage and yum sauce Alouette potatoes Capellini primavera Danish French bread Sugar shortbread Vegan caramels cookies Blueberry muffin mix

Future Work

- Further text preprocessing and parameters tuning
- Classify recipes into types to improve recommendations
- Test on new data set
- Create web app to recommend new recipes