

Recipe Recommender

Erin Cox | Oakland, CA



[coxem14/Recipe-Recommender](https://github.com/coxem14/Recipe-Recommender)



[in/coxem14](https://www.linkedin.com/in/coxem14)



Inspiration



I love trying to recreate my favorite restaurant dishes at home!

My partner and I are in the exploratory phase to add new recipes to our cookbook.

I thought it would be fun and helpful to make a content-based recipe recommender.

Data



recipe-search.typsense.org

- I used a subset of 250K recipes which were formatted to have better quality (source = Gathered)
- The dataset included 6 feature columns containing strings of lists and text:

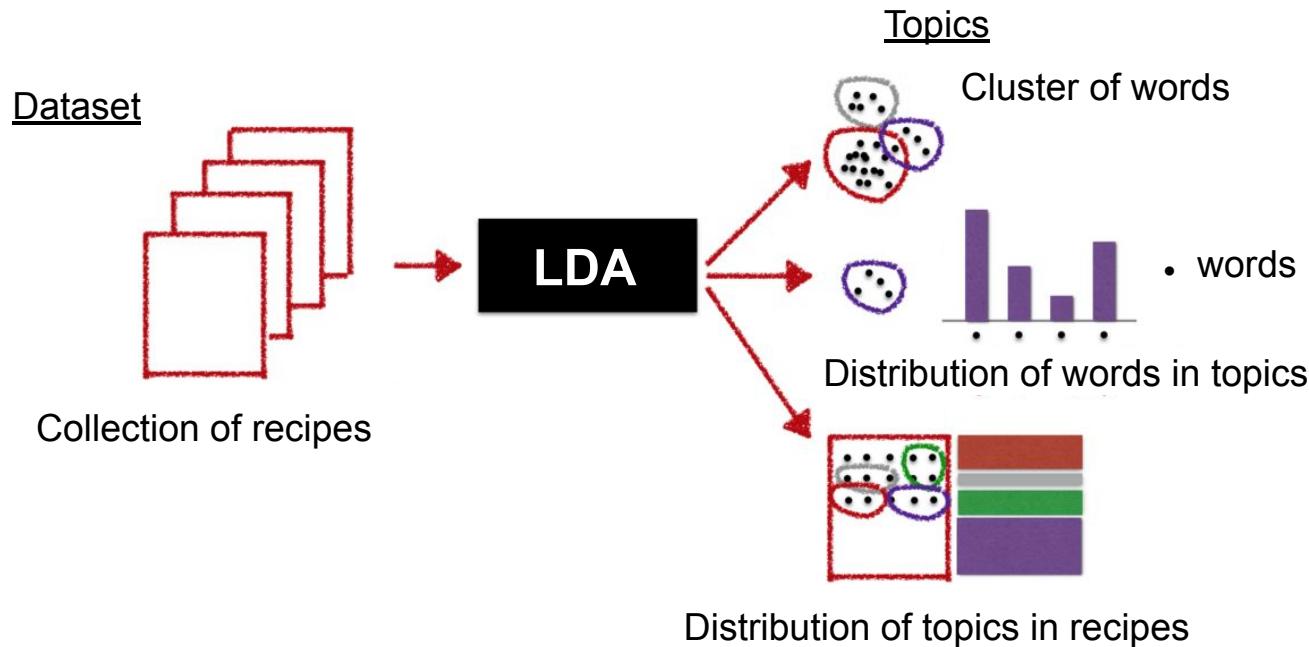
Unnamed: 0	title	ingredients	directions	link	source	NER
0	0 No-Bake Nut Cookies	["1 c. firmly packed brown sugar", "1/2 c. eva..."]	["In a heavy 2-quart saucepan, mix brown sugar..."]	www.cookbooks.com/Recipe-Details.aspx?id=44874	Gathered	["brown sugar", "milk", "vanilla", "nuts", "bu..."] ["["1 c. firmly packed brown sugar", "1/2 c. evaporated milk", "1/2 tsp. vanilla", "1/2 c. broken nuts (pecans)", "2 Tbsp. butter or margarine", "3 1/2 c. bite size shredded rice biscuits"]"]

Process



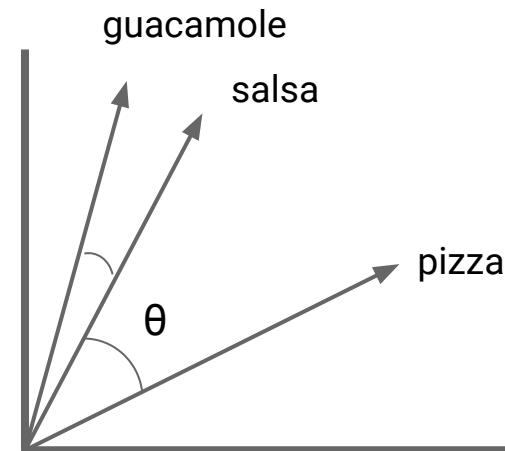
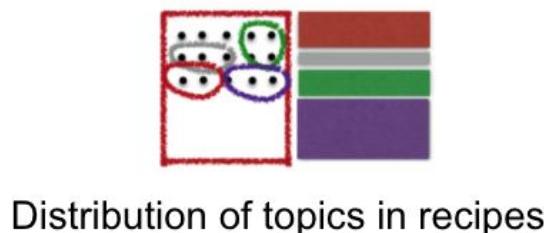
Model Selection

Latent Dirichlet Allocation (LDA) for Topic Modeling



Model Selection

Latent Dirichlet Allocation (LDA) for Topic Modeling



Data Cleaning and Feature Selection

Unnamed: 0		title	ingredients	directions	link	source	NER
0	0	No-Bake Nut Cookies	["1 c. firmly packed brown sugar", "1/2 c. eva...]	["In a heavy 2-quart saucepan, mix brown sugar...	www.cookbooks.com/Recipe-Details.aspx?id=44874	Gathered	["brown sugar", "milk", "vanilla", "nuts", "bu...
1	1	Jewell Ball'S Chicken	["1 small jar chipped beef, cut up", "4 boned ...	["Place chipped beef on bottom of baking dish....	www.cookbooks.com/Recipe-Details.aspx?id=699419	Gathered	["beef", "chicken breasts", "cream of mushroom...
2	2	Creamy Corn	["2 (16 oz.) pkg. frozen corn", "1 (8 oz.) pkg...]	["In a slow cooker, combine all ingredients. C...	www.cookbooks.com/Recipe-Details.aspx?id=10570	Gathered	["frozen corn", "cream cheese", "butter", "gar...
3	3	Chicken Funny	["1 large whole chicken", "2 (10 1/2 oz.) cans...]	["Boil and debone chicken.", "Put bite size pi...	www.cookbooks.com/Recipe-Details.aspx?id=897570	Gathered	["chicken", "chicken gravy", "cream of mushro...
4	4	Reeses Cups(Candy)	["1 c. peanut butter", "3/4 c. graham cracker ...]	["Combine first four ingredients and press in ...	www.cookbooks.com/Recipe-Details.aspx?id=659239	Gathered	["peanut butter", "graham cracker crumbs", "bu...



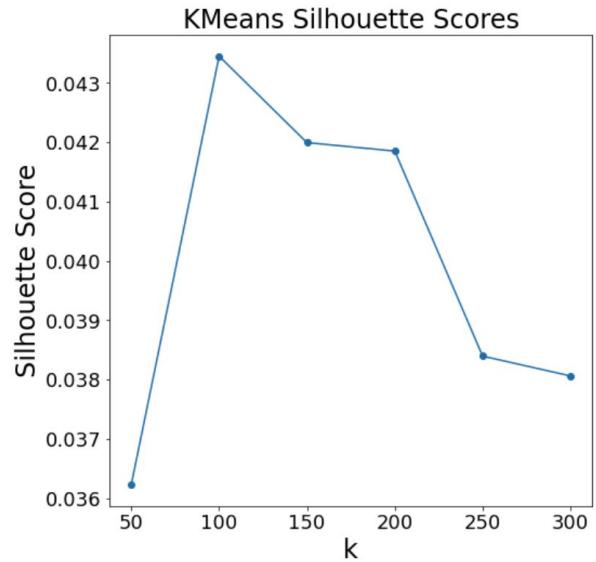
Unnamed: 0		title	NER	site	urls	bag_of_words	cleaned_bow
0	0	No-Bake Nut Cookies	[brown sugar, milk, vanilla, nuts, butter, bit...]	www.cookbooks.com	http://www.cookbooks.com/Recipe-Details.aspx?id=44874	No-Bake Nut Cookies 1 c. firmly packed brown s...	no-bak nut cooki 1 firml pack brown sugar 1/2...
1	1	Jewell Ball'S Chicken	[beef, chicken breasts, cream of mushroom soup...]	www.cookbooks.com	http://www.cookbooks.com/Recipe-Details.aspx?id=699419	Jewell Ball'S Chicken 1 small jar chipped beef...	jewel ball 's chicken 1 small jar chip beef cu...
2	2	Creamy Corn	[frozen corn, cream cheese, butter, garlic pow...]	www.cookbooks.com	http://www.cookbooks.com/Recipe-Details.aspx?id=10570	Creamy Corn 2 (16 oz.) pkg. frozen corn 1 (8 o...	creami corn 2 16 pkg frozen corn 1 8 pkg cream...
3	3	Chicken Funny	[chicken, chicken gravy, cream of mushroom sou...]	www.cookbooks.com	http://www.cookbooks.com/Recipe-Details.aspx?id=897570	Chicken Funny 1 large whole chicken 2 (10 1/2 ...	chicken funni 1 larg whole chicken 2 10 1/2 ca...
4	4	Reeses Cups(Candy)	[peanut butter, graham cracker crumbs, butter,...]	www.cookbooks.com	http://www.cookbooks.com/Recipe-Details.aspx?id=659239	Reeses Cups(Candy) 1 c. peanut butter 3/4 c....	rees cup candi 1 peanut butter 3/4 graham crac...

KMeans Clustering - number of topics

25,000 sample subset

```
vectorizer = TfidfVectorizer(max_df=0.85,  
                             min_df=10,  
                             ngram_range=(1,3),  
                             max_features=1000)  
  
docs_vec = vectorizer.fit_transform(cleaned_bow)
```

100 topics resulted in the best silhouette score

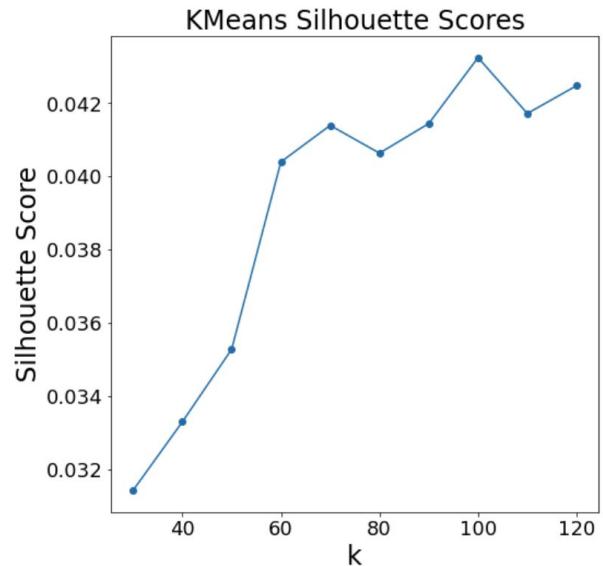


KMeans Clustering - number of topics

25,000 sample subset

```
vectorizer = TfidfVectorizer(max_df=0.85,  
                             min_df=10,  
                             ngram_range=(1,3),  
                             max_features=1000)  
  
docs_vec = vectorizer.fit_transform(cleaned_bow)
```

100 topics resulted in the best silhouette score



KMeans Clustering - top words by cluster

~200,000 sample subset

top features (words) for each cluster:

0, rib, sauc, barbecu, short, hour, pork, pepper, grill, garlic, brown
1, pancak, batter, egg, flour, milk, powder, bake, sugar, cook, butter
2, flour, sugar, egg, bake, cake, vanilla, add, butter, powder, cream

KMeans Clustering - recipes by cluster

top features (words) for each cluster:

0, rib, sauc, barbecu, short, hour, pork, pepper, grill, garlic, brown
1, pancak, batter, egg, flour, milk, powder, bake, sugar, cook, butter
2, flour, sugar, egg, bake, cake, vanilla, add, butter, powder, cream

cluster 0:

Korean Barbecued Short Ribs
Texas Style Fall Of The Bone Crock Pot Pork Ribs
Japanese Spicy Spare Ribs
Chinese Five Spice Spare Ribs
Hoisin-Ginger Spare Ribs With Coleslaw
Beef Short Ribs Polynesian
Pineapple-Glazed Spareribs
Cajun Back Ribs
Jalapeno And Beer Baked Short Ribs
Thai Barbecue Ribs

KMeans Clustering - recipes by cluster

top features (words) for each cluster:

0, rib, sauc, barbecu, short, hour, pork, pepper, grill, garlic, brown
1, pancak, batter, egg, flour, milk, powder, bake, sugar, cook, butter
2, flour, sugar, egg, bake, cake, vanilla, add, butter, powder, cream

cluster 0:

Korean Barbecued Short Ribs
Texas Style Fall Of The Bone Crock Pot Pork Ribs
Japanese Spicy Spare Ribs
Chinese Five Spice Spare Ribs
Hoisin-Ginger Spare Ribs With Coleslaw
Beef Short Ribs Polynesian
Pineapple-Glazed Spareribs
Cajun Back Ribs
Jalapeno And Beer Baked Short Ribs
Thai Barbecue Ribs

cluster 1:

Potato Pancakes
Cottage Cheese Latkes
Basic Pancakes
Pancakes
Caramel Cake Pancakes
Norwegian Pancakes
Blueberry Banana Pancakes
My Mother'S Buttermilk Pancakes
Banana Pancakes
Volcano Pancake

KMeans Clustering - recipes by cluster

top features (words) for each cluster:

0, rib, sauc, barbecu, short, hour, pork, pepper, grill, garlic, brown
1, pancak, batter, egg, flour, milk, powder, bake, sugar, cook, butter
2, flour, sugar, egg, bake, cake, vanilla, add, butter, powder, cream

cluster 0:

Korean Barbecued Short Ribs
Texas Style Fall Of The Bone Crock Pot Pork Ribs
Japanese Spicy Spare Ribs
Chinese Five Spice Spare Ribs
Hoisin-Ginger Spare Ribs With Coleslaw
Beef Short Ribs Polynesian
Pineapple-Glazed Spareribs
Cajun Back Ribs
Jalapeno And Beer Baked Short Ribs
Thai Barbecue Ribs

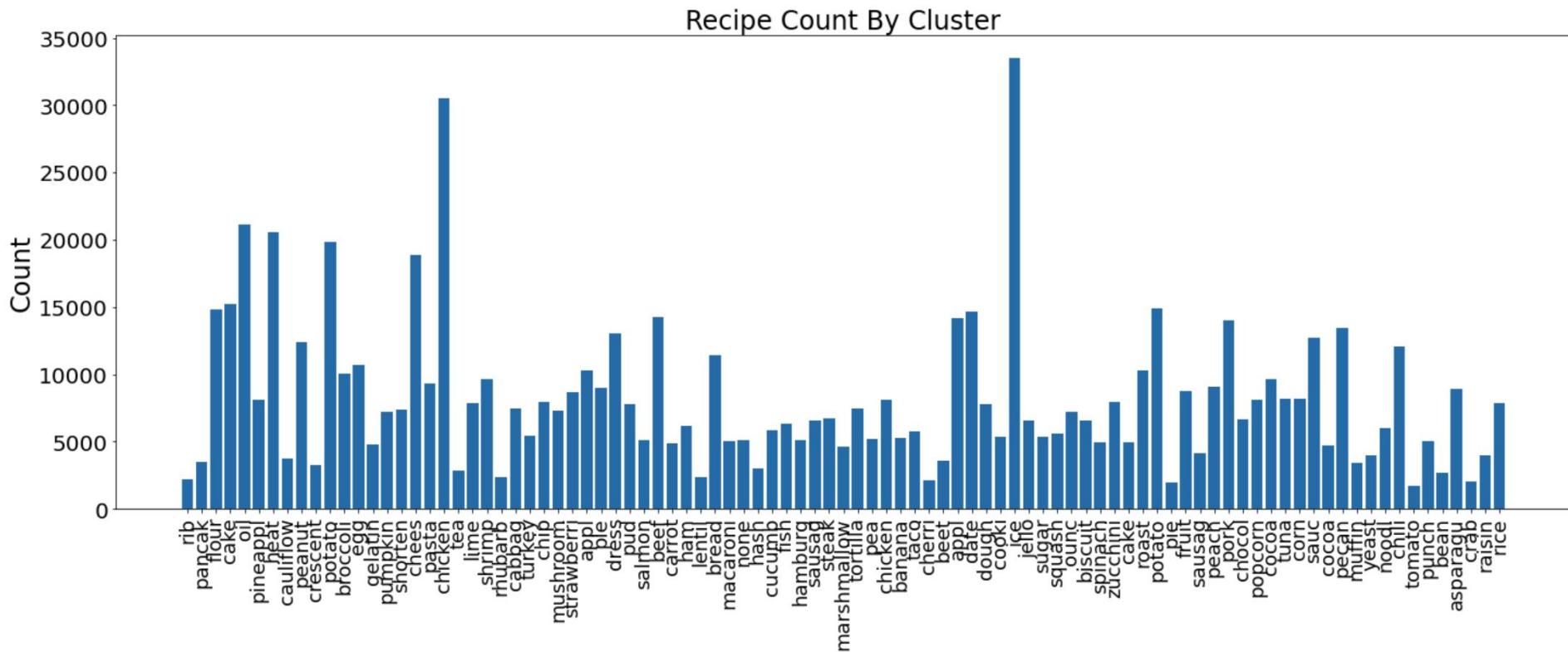
cluster 1:

Potato Pancakes
Cottage Cheese Latkes
Basic Pancakes
Pancakes
Caramel Cake Pancakes
Norwegian Pancakes
Blueberry Banana Pancakes
My Mother'S Buttermilk Pancakes
Banana Pancakes
Volcano Pancake

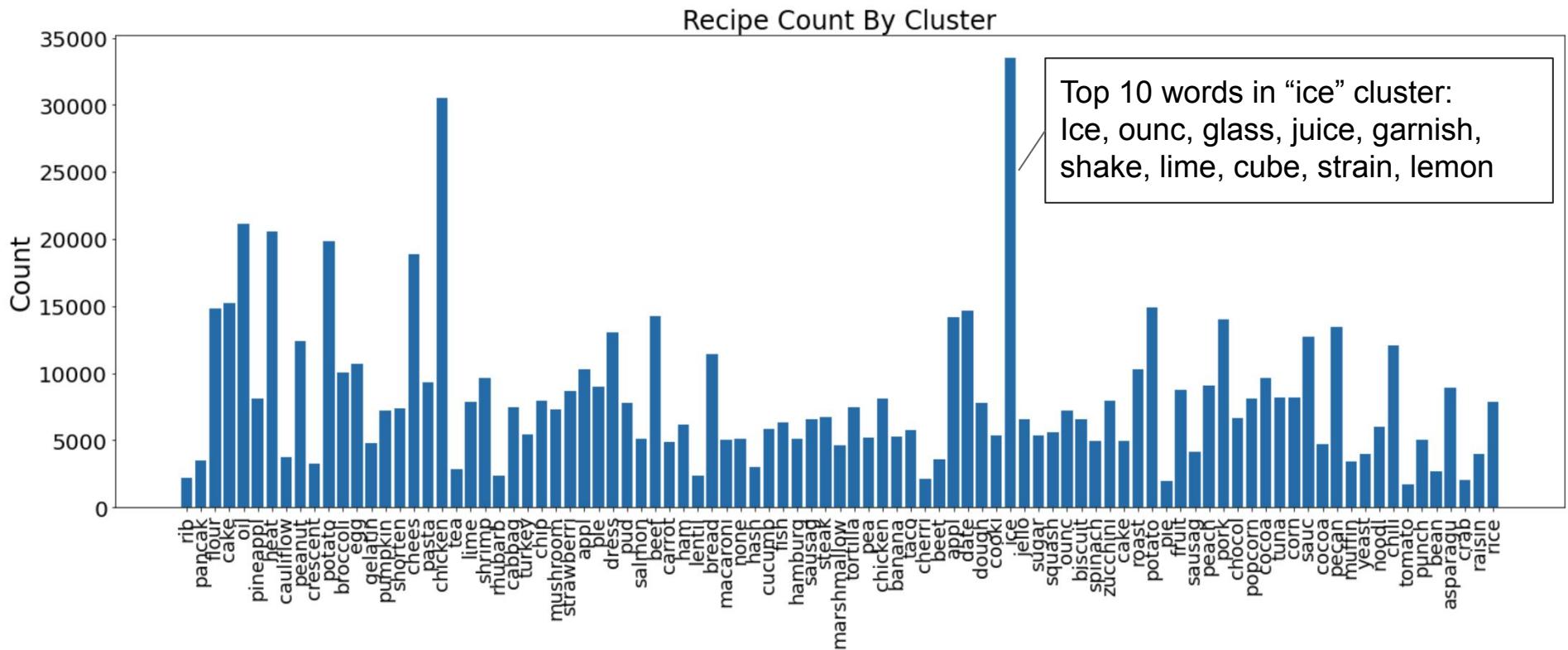
cluster 2:

Sour Cream Pound Cake
Italian Cream Cake
Blackberry Cake
Italian Cream Cake
Cajun Chew Cake
Brownies
Potato Cake
1-2-3-4 Cake
Lilly'S Pound Cake
Coffee Cake

KMeans - recipe count by cluster



KMeans - recipe count by cluster



LDA

Vectorizer

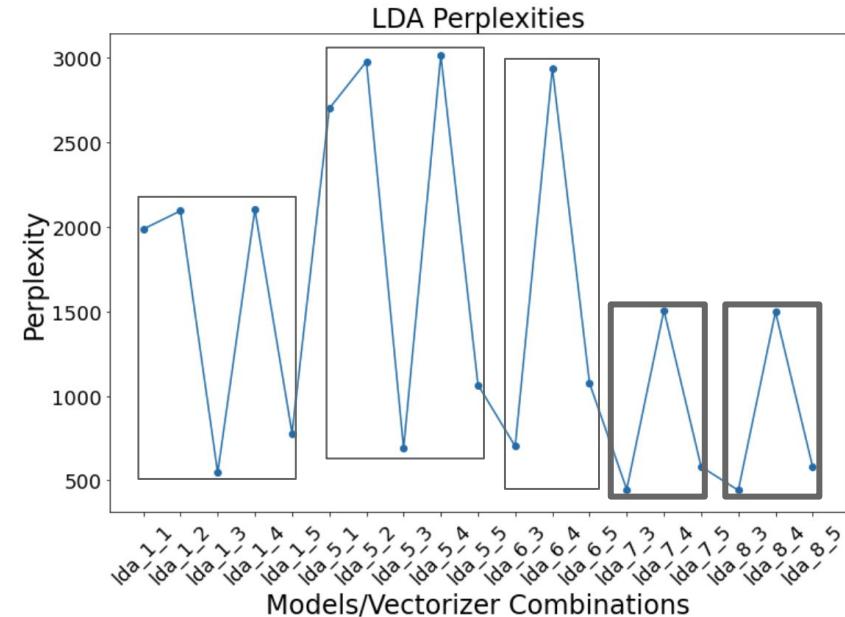
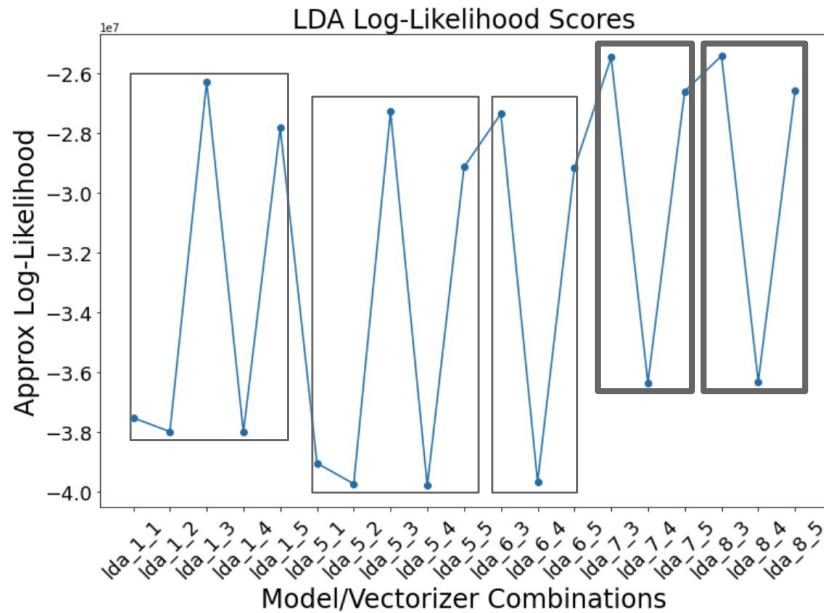
```
vec = CountVectorizer(max_df=0.85,  
                     min_df=2,  
                     ngram_range=(1,2),  
                     max_features=1000)  
  
term_frequency = vec.fit_transform(cleaned_bow)
```

LDA

Model

```
lda = LatentDirichletAllocation(n_components=100,  
                                learning_method='online',  
                                learning_offset=10,  
                                doc_topic_prior=0.01,  
                                topic_word_prior=0.01,  
                                batch_size=128)  
  
lda.fit(term_frequency)
```

LDA Model Tuning



Model Tuning

Top 10 recipes for keyword “ribs”

Reference Recipe: Baked Pork Ribs With Hoisin Barbecue Sauce

Model Perplexity: 441	Model Perplexity: 1497	Model Perplexity: 580
Lite Spaghetti	Chinese Country-Style Pork Ribs	Marinated Brisket Roast
Heartwarming Beef Stew	Baked Barbecue Ribs	Barbecue Sauce For Spareribs
Barbecued Beef(For A Bunch, 24)	Country Ribs(Very Good)	Barbecued Beef(For A Bunch, 24)
Oh Meatballs!, Oh Meatloaf!	Honey Barbecue Spareribs	Boneless Korean Short Ribs
Oven-Baked Maple Barbecued Ribs	Chinese 5 Spice Spare Ribs	Wicked Chicken (Bbq New England Style)
Venison Jerky	Ribs With Apple Butter	Ken'S Incredible Barbecue Sauce
Sloppy Joes	Kansas City-Style Baby Back Ribs	Barbecue Sauce
Marinated Brisket Roast	Oven-Baked Maple Barbecued Ribs	Meat Loaf
Barbecue Sandwiches	Mahogany Glazed Spare Ribs	Daniel'S Quick & Easy Tartar Sauce
Maple Barbecued Caribou Ribs	Asian Ribs	Coke Chicken

Demo

250k Recipes! Home Explore About

Recipe Recommender

search by keyword(s)

[Explore Recipes](#)

250k Recipes! Home Explore About

Explore recipes by keyword(s)
try "pizza" or "guacamole" or "chocolate cake", etc.

[Get Results](#)

250k Recipes! Home Explore About

Curious about "indian"? Maybe these will hit the spot:

- Buddha'S Happy Belly Curried Lentils, Cauliflower And Tomatoes
- Cashewcaesar Con Chive Verde
- Indian Spiced Cauliflower And Potatoes
- Potato Chana Masala
- Potato And Pea Curry (Also Mater) Recipe
- Tomato Masala Rice (To Use Up Leftovers!)
- Kidney Bean Curry / Rajma Masala
- Eggplant In Spicy Tomato Sauce
- Unstuffed Cabbage

See more results for "indian" by refreshing the page.

[Return to Explore](#)

epicurious

INDIAN SPICED CAULIFLOWER AND POTATOES

SUBMITTED BY JPOINSCFHD | UPDATED: SEPTEMBER 28, 2015

Be the first to rate and review this recipe

0/4 0% MAKE IT AGAIN

... was pleasantly surprised with the flavor and texture of the entire dish. The peppers - added at the very end - give a nice raw pepper crispness and provide such a vibrant color to the entire dish.
Lifted from: <http://kastkitchentalk.blogspot.com/2008/01/indian-spiced-cauliflower-and-potatoes.html>

[f](#) [t](#) [p](#) [g](#) [e](#)

[SAVE RECIPE](#) [ADD TO MENU](#) [PRINT](#)



Erin Cox | Oakland, CA



[coxem14/Recipe-Recommender](https://github.com/coxem14/Recipe-Recommender)



[in/coxem14](https://www.linkedin.com/in/coxem14)



Appendix

Data

Instant Search 2 Million Recipes 🍜

powered by [typesense](#)

[View on GitHub](#)

Type in a recipe title

Try: Pizza, Pineapple, Kale, Oregano, Salad, Curry, Healthy, Keto, Low Carb, Steamed, Fried

recipe-search.typesense.org

RecipeNLG: A Cooking Recipes Dataset for Semi-Structured Text Generation

This is an archive of code which was used to produce dataset and results available in our INLG 2020 paper: [RecipeNLG: A Cooking Recipes Dataset for Semi-Structured Text Generation](#)

github.com/Glorf/recipenlg

[RecipeNLG \(Dataset\)](#)

- The dataset contains 2,231,142 recipes.
- ~1.6M of the recipes were formatted to have better quality (source = Gathered)
 - I used a subset (250K) of these recipes in my project.
- The dataset included 6 feature columns containing strings of lists and text.

Unnamed: 0	title	ingredients	directions	link	source	NER
0	0 No-Bake Nut Cookies	["1 c. firmly packed brown sugar", "1/2 c. eva..."]	["In a heavy 2-quart saucepan, mix brown sugar..."]	www.cookbooks.com/Recipe-Details.aspx?id=44874	Gathered	["brown sugar", "milk", "vanilla", "nuts", "bu..."] ["1 c. firmly packed brown sugar", "1/2 c. evaporated milk", "1/2 tsp. vanilla", "1/2 c. broken nuts (pecans)", "2 Tbsp. butter or margarine", "3 1/2 c. bite size shredded rice biscuits"]

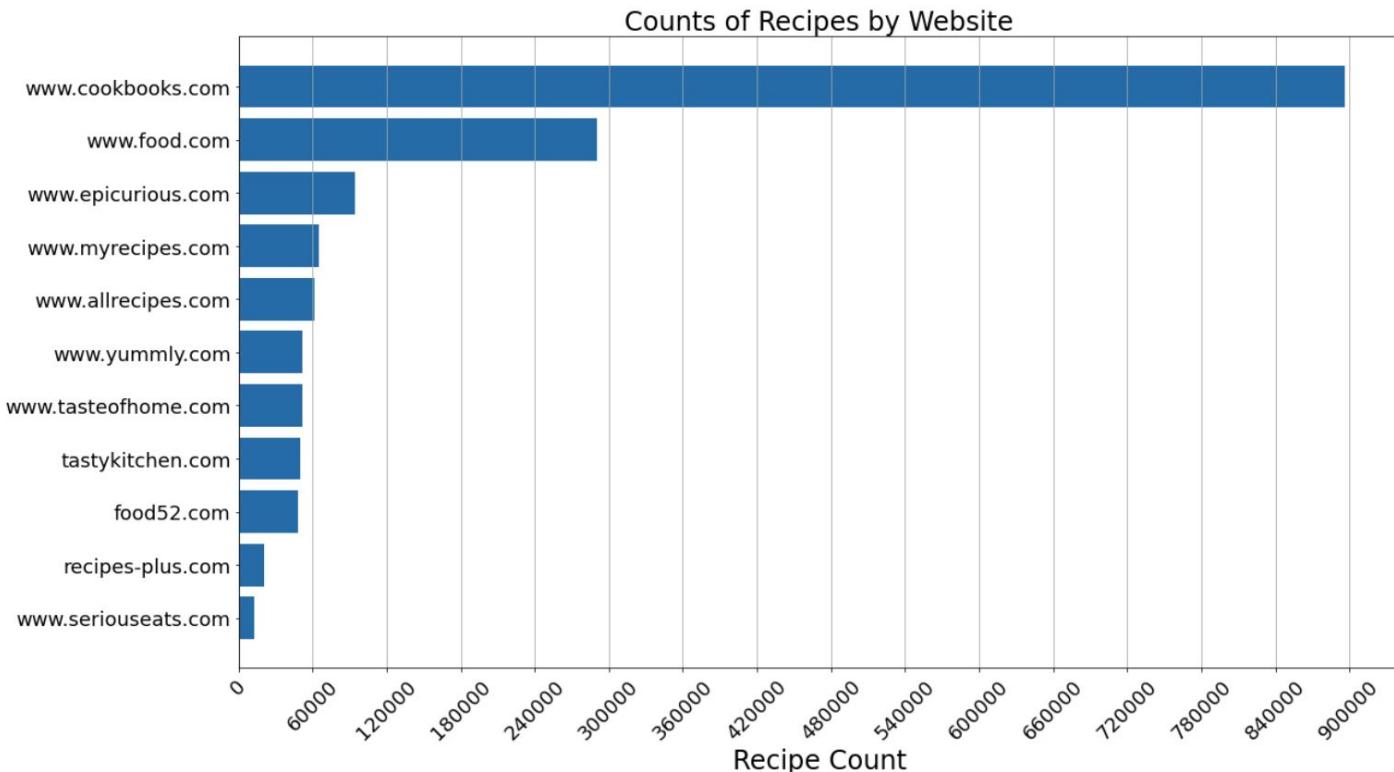
Data Cleaning and Feature Selection

Unnamed: 0		title	ingredients	directions	link	source	NER
0	0	No-Bake Nut Cookies	["1 c. firmly packed brown sugar", "1/2 c. eva...]	["In a heavy 2-quart saucepan, mix brown sugar...]	www.cookbooks.com/Recipe-Details.aspx?id=44874	Gathered	["brown sugar", "milk", "vanilla", "nuts", "bu...]
1	1	Jewell Ball'S Chicken	["1 small jar chipped beef, cut up", "4 boned ...]	["Place chipped beef on bottom of baking dish....]	www.cookbooks.com/Recipe-Details.aspx?id=699419	Gathered	["beef", "chicken breasts", "cream of mushroom...]
2	2	Creamy Corn	["2 (16 oz.) pkg. frozen corn", "1 (8 oz.) pkg...]	["In a slow cooker, combine all ingredients. C...]	www.cookbooks.com/Recipe-Details.aspx?id=10570	Gathered	["frozen corn", "cream cheese", "butter", "gar...]
3	3	Chicken Funny	["1 large whole chicken", "2 (10 1/2 oz.) cans...]	["Boil and debone chicken.", "Put bite size pi...]	www.cookbooks.com/Recipe-Details.aspx?id=897570	Gathered	["chicken", "chicken gravy", "cream of mushro...]
4	4	Reeses Cups(Candy)	["1 c. peanut butter", "3/4 c. graham cracker ...]	["Combine first four ingredients and press in ...]	www.cookbooks.com/Recipe-Details.aspx?id=659239	Gathered	["peanut butter", "graham cracker crumbs", "bu...]



Unnamed: 0		title	NER	site	urls	bag_of_words	cleaned_bow
0	0	No-Bake Nut Cookies	[brown sugar, milk, vanilla, nuts, butter, bit...]	www.cookbooks.com	http://www.cookbooks.com/Recipe-Details.aspx?i...	No-Bake Nut Cookies 1 c. firmly packed brown s...	no-bak nut cooki 1 firml pack brown sugar 1/2...
1	1	Jewell Ball'S Chicken	[beef, chicken breasts, cream of mushroom soup...]	www.cookbooks.com	http://www.cookbooks.com/Recipe-Details.aspx?i...	Jewell Ball'S Chicken 1 small jar chipped beef...	jewel ball 's chicken 1 small jar chip beef cu...
2	2	Creamy Corn	[frozen corn, cream cheese, butter, garlic pow...]	www.cookbooks.com	http://www.cookbooks.com/Recipe-Details.aspx?i...	Creamy Corn 2 (16 oz.) pkg. frozen corn 1 (8 o...	creami corn 2 16 pkg frozen corn 1 8 pkg cream...
3	3	Chicken Funny	[chicken, chicken gravy, cream of mushroom sou...]	www.cookbooks.com	http://www.cookbooks.com/Recipe-Details.aspx?i...	Chicken Funny 1 large whole chicken 2 (10 1/2 ...	chicken funni 1 larg whole chicken 2 10 1/2 ca...
4	4	Reeses Cups(Candy)	[peanut butter, graham cracker crumbs, butter,...]	www.cookbooks.com	http://www.cookbooks.com/Recipe-Details.aspx?i...	Reeses Cups(Candy) 1 c. peanut butter 3/4 c....	rees cup candi 1 peanut butter 3/4 graham crac...

Distribution of Websites



Final Thoughts

Recap of Findings:

- LDA is a useful tool for topic modeling, but tuning can be challenging as the standard scoring metrics in sklearn may not correlate with human judgement.
- LDA unearths hidden topics within recipes that might not be immediately obvious.

Future Steps:

- Use a topic coherence scoring metric and visualize topics with Gensim Library
 - Coherence measures the degree of semantic similarity between words in a topic. This helps distinguish between topics that are semantically interpretable, and those which result from statistical inference.
- Improve and deploy Flask App (filter by ingredients, custom number of recipe recommendations, etc.)

Model Tuning - Final Model

```
vec = CountVectorizer(max_df=0.85,
                     min_df=10,
                     ngram_range=(1,3),
                     max_features=1000)

tf = vec.fit_transform(cleaned_bow)

lda = LatentDirichletAllocation(n_components=100,
                                 learning_method='online',
                                 learning_offset=50,
                                 doc_topic_prior=0.9,
                                 topic_word_prior=0.9,
                                 batch_size=32,
                                 )

lda.fit(tf)
```