Centre for Data Analytics



Recipe Recommendation Reasoning on Features: Ensemble Topic Modeling Based Approach

Mansura A. Khan, Ellen Rushe, Barry Smyth and David Coyle University College Dublin











Why food recommendation is interesting ?

Food choice is multifaceted, multilevel, contextual, dynamic, integrated and diverse

-Jeffery Sobal



Core Challenges in Food/Recipe Recommendation Domain

- Capturing user's preference on food related aspects.
- Identifying coherent content and context information.
- Getting access to standardized food and nutrition data.
- Representation of Food Data
- Mapping health on recommendations.
 - Measuring healthiness of food item.
 - Measure user's nutrient requirement.
- Diversity
- Coverage
- Pursue user to choose healthier options.
- Achieve continuous and regular user engagement.

What is a recipe?

Low Carb Scalloped Zucchini



S 1hr15mins 🛞 4 Serving 😫 Nutrition

redients			
1:3 medium zu	ucchini: sliced int	o disks	
2:1 egg			
3:1 cup cottag	je cheese		
4:1/3 cup red (onion		
5: 1/4 lb diced	ham		
6:2 teaspoons	dried parsley		
7:1/2 teaspoor	n salt		
8: 1/4 teaspoor	n garlic powder		
9:6 slices bac	on		

Instruction

- Step 1 : Preheat oven to 350°F. Step 2 : Boil the zucchini in salted water for 5 minutes and then drain really well and pat dry.
- Step 3 : Mix the egg, cottage cheese, onion, diced ham, parsley, salt and garlic powder together.
- Step 4 : Spray a 2 quart casserole with nonstick cooking spray and then layer 1/3 of the zucchini in the bottom.
- Step 5 : Layer 1/2 of the cheesy ham mixture over the zucchini and
- continue layering until finished off with a top layer of zucchini.
- Step 6 : Spread the uncooked bacon out over the top and bake for 1 hour at 350°F.

Nutrition

- Calorie : 299.9 Calories from Fat : 187
- Total Fat: 20.8g
- Saturated Fat : 7.6g Total Carbohydrate : 8.2g
- Sugars : 3.4g
- Dietary Fiber: 1.9g
- Protein: 20.4g
- Cholesterol: 98.6mg
- Sodium: 1249.6mg

Recipe is a Story : Document

Title: Yummy mashed potatoes Ingredients: butter, cream, potatoes, salt and pepper Instruction : cook potatoes in bowling water until tender but still firm, In a small saucepan heat butter and cream at low heat. Using a potato masher slowly blend cream mixture into potatoes. Season with salt and pepper to taste.

How RS reads a recipe?

Frequency Count :

Occurrence of terms in each document

Terms	Frequency
Potatoes	4
saucepan	1
heat	2
title	1
in	2
salt	2
tender	1
а	2

Recipe 1

Title: Yummy mashed potatoes Ingredients: potatoes, butter, cream, salt and pepper Instruction : cook potatoes in bowling water until tender but still firm, In a small saucepan heat butter and cream at low heat. Using a potato masher slowly blend cream mixture into potatoes. Season with salt and pepper to taste.

Bag of word representation



а	in	potatoes	salt	low	Sauce pane	but		masher		blend
2	2	4	2	1	1	1		1		1

Hundreds of thousands of Documents : Recipe Corpus



Recipe 1, 2, 3

Bag of word representation





	а	in	potatoes	salt	low	saucepan	but	beef	lamb
R1	2	2	4	2	1	1	1	0	0
R2	9	5	0	4	0	4	0	2	0
R3	6	4	3	3	0	2	0	0	3

Recommending recipes

beef recipes

	а	in	potatoes	salt	low	saucepan	but	beef	lamb
R1	2	2	4	2	1	1	1	0	0
R2	9	5	0	4	0	4	0	2	0
R3	6	4	3	3	0	2	0	0	3

Understanding the relevance weight

	а	in	potatoes	salt	low	saucepan	but	beef	lamb
R1	2	2	4	2	1	1	1	0	0
R2	9	5	0	4	0	4	0	2	0
R3	6	4	3	3) 0	2	0	0	3

Document Frequency: Occurrence of terms over entire corpus

Title: Irish stew Ingredients: chuck lamb stew meat, parsnips, broth, russet potatoes Title: Gunnies beef pie Ingredients: flour, beef, gunnies, butter, salt , celery , beef stock , carrot Title: Yummy mashed potatoes Ingredients: potatoes, butter, cream, salt and pepper Instruction : cook potatoes in bowling water until tender but still firm, In a small saucepan heat butter and cream at low heat. Using a potato masher slowly blend cream mixture into potatoes. Season with salt and pepper to taste.

Terms	Frequency
Potatoes	7
saucepan	7
heat	9
Title	3
in	11
salt	9
tender	1
а	17

Frequency Matrix

	а	in	potatoes	salt	low	saucepan	but	beef		lamb
R1	2	2	4	2	1	1	1	0		0
R2	9	5	0	4	0	4	0	2		0
R3	6	4	3	3	0	2	0	0	(3

Tf-Idf Matrix

	а	in	potatoes	salt	low	saucepan	but			beef			lamb
R1													
	0.493	0.871	2.296	1.045	1.477	0.632	1.477	0	0	0	0	0	0
R2													
	2.220	2.178	0.574	2.091	0	2.528	0	0	0	2.352	0	0	0
R3													
	1.480	1.742	1.722	1.568	0	1.264	0	0	0	0	0	0	3

- 10,661 ingredients
- 230,872 recipes



Considering only highly frequent ingredients ?



What information are relevant to Food Recommender?

Topic Modeling

Topic models aim to discover the latent semantic structure or themes within a corpus of documents, which can be derived from co-occurrences of words across the documents.

-Mark Belford

Topic Model

A topic model typically consists of k topics, each represented by a ranked list of strongly-associated terms (often referred to as a "topic descriptor").

		а	in	potatoes	salt	low	saucepan	but	beef	lamb			
	R1	2	2	4	2	1	1	1	0	0			
	R2	9	5	0	4	0	4	0	2	0			
	R3	6	4	3	3	0	2	0	0	3			
	R4	2	2	4	2	1	1	1	0	0			
	R5	9	5	0	4	0	4	0	2	0			
L	R6	6	4	3	3	0	2	0	0	3			
				-	Тор	ic N	Nodel	ing			Recip	e-Topic I	Matri
								•			lopic	lopic	Ιορι
										R_1	.5	0	0
rm M	atrix	X					,			R ₂	0	0	0
otatoes		sau	сера	an co	со	be	ef la	amb		R ₃	.7	0	0.00
6			.07	(C	-	7	0		R_4	0	0	0
					-						•		

Topic -Te

	potatoes	saucepan	сосо	beef	lamb
Topic ₁	6	.07	0	7	0
Topic ₂	6	1	0	0	4
Topic ₃	0	1	2	0	0

ix

	Topic ₁	Topic ₂	Topic ₃
R_1	.5	0	0
R_2	0	0	0
R_3	.7	0	0.002
R_4	0	0	0
R_5	0	0.3	0.9
R_6	0	0	0

Each document/recipe in the corpus can be associated with one or more topics.

Non-negative matrix factorization



$$V \approx W \times H$$

It can be rewritten column by column as,

a ~ $W \times b$



		а	in	potatoes	salt	low	saucepan	but	beef	lamb			
	R1	2	2	4	2	1	1	1	0	0			
	R2	9	5	0	4	0	4	0	2	0			
	R3	6	4	3	3	0	2	0	0	3			
	R4	2	2	4	2	1	1	1	0	0			
	R5	9	5	0	4	0	4	0	2	0			
L	R6	6	4	3	3	0	2	0	0	3			
				-	Тор	ic N	Nodel	ing			Recip	e-Topic I	Matri
								•			lopic	lopic	Ιορι
										R_1	.5	0	0
rm M	atrix	X					,			R ₂	0	0	0
otatoes		sau	сера	an co	со	be	ef la	amb		R ₃	.7	0	0.00
6			.07	(C	-	7	0		R_4	0	0	0
					-						•		

Topic -Te

	potatoes	saucepan	сосо	beef	lamb
Topic ₁	6	.07	0	7	0
Topic ₂	6	1	0	0	4
Topic ₃	0	1	2	0	0

ix

	Topic ₁	Topic ₂	Topic ₃
R_1	.5	0	0
R_2	0	0	0
R_3	.7	0	0.002
R_4	0	0	0
R_5	0	0.3	0.9
R_6	0	0	0

Outcome of different topic modeling session is only approximately equal

Торіс	Terms
Topic 0	beginner-cook, easy, inexpensive, kid-friendly, appetizer, eggs-dairy, holiday-event, snack, toddler-friendly, technique, served-cold, potluck, egg-free, finger-food, european
Topic 1	cheese, eggs-dairy, cheddar, parmesan, appetizer, mozzarella, tortilla, high-calcium, low-carb, cheesecake, macaroni, mayonnaise, feta, goat, lasagna
Topic 2	spinach, olive, feta, artichoke, vegetable, strawberry, parmesan, low-carb, pine, main-dish, salad-dressings, lasagna, tofu, vinaigrette, phyllo
Topic 3	casserole, one-dish-meal, main-dish, sausage, cheddar, noodle, tuna, side-dishes, macaroni, potluck, bubbly, cheese, weeknight, ritz, less-than-half-hour
Topic 4	apple, dessert, low-protein, tart, cranberry, pies-and-tarts, walnut, oat, caramel, applesauce, pecan, maple, pear, batter, salad
Topic 5	dessert, cookies-and-brownies, for-large-groups, cocoa, holiday-event, pecan, almond, bar-cookies, pudding, drop-cookies, christmas, pies-and-tarts, pumpkin, savoury, mixer
Topic 6	vegetable, carrot, side-dishes, zucchini, soups-stews, celery, cabbage, stove-top, asparagus, carrots, cauliflower, pea, pumpkin, mushrooms, green-yellow-beans
Topic 7	corn, tortilla, cornstarch, cornmeal, salsa, cilantro, chowder, cornbread, jalapeno, chilies, cob, ear, popcorn, soups-stews, pecan
Topic 8	rice, side-dishes, soy, white-rice, brown-rice, chinese, basmati, long-grain, pea, one-dish-meal, risotto, pilaf, stove-top, stir-fry, celery
Topic 9	low-cholesterol, low-protein, low-calorie, low-carb, easy, side-dishes, soups-stews, appetizer, strawberry, smoothy, peach, cucumber, technique, savory-sauces, salsa
Topic 10	broccoli, floret, cauliflower, vegetable, mayonnaise, soy, cornstarch, stir-fry, cheddar, chinese, side-dishes, bisques-cream-soups, florets, low-carb, sunflower

Topic	Terms
Topic 0	vegetable, side-dishes, broccoli, spinach, zucchini, asparagus, stove-top, cauliflower, olive, mushrooms, green-yellow-beans, served-hot, soy, beans, cucumber
Topic 1	dessert, cookies-and-brownies, for-large-groups, cocoa, pecan, bar-cookies, holiday-event, drop-cookies, pudding, cooky, walnut, pies-and-tarts, mixer, christmas, hand-formed-cookies
Topic 2	bread, quick-breads, loaf, wheat, rolls-biscuits, low-cholesterol, muffins, small-appliance, for-large-groups, cornmeal, batter, weeknight, high-calcium, coffee-cakes, easy
Topic 3	dessert, apple, low-protein, tart, cranberry, pies-and-tarts, walnut, saur, caramel, applesauce, pecan, maple, pear, batter, salad
Topic 4	chicken, poultry, meat, chicken-breasts, main-dish, low-carb, one-dish-meal, high-protein, chicken-thighs-legs, olive, soy, wing, whole-chicken, easy, weeknight
Topic 5	low-cholesterol, low-protein, low-calorie, low-carb, easy, side-dishes, soups-stews, appetizer, strawberry, smoothy, peach, cucumber, technique, savory-sauces, salsa
Topic 6	beginner-cook, easy, inexpensive, kid-friendly, eggs-dairy, holiday-event, toddler-friendly, snack, technique, served-cold, potluck, egg-free, european, served-hot, south-west-pacific
Topic 7	orange, citrus, cranberry, marmalade, mandarin, holiday-event, tea, oranges, easy, christmas, served-cold, olive, apricot, eel, grapefruit
Topic 8	banana, tropical-fruit, bananas, quick-breads, smoothy, yogurt, muffins, smoothie, walnut, pudding, kid-friendly, wheat, oat, blueberry, batter
Topic 9	almond, blueberry, cherry, raspberry, almonds, peach, coffee, pitted-fruit, yogurt, batter, apricot, yolk, cranberry, european, maraschino
Topic 10	tortilla, appetizer, salsa, cilantro, avocado, jalapeno, mayonnaise, for-large-groups, lettuce, chilies, finger-food, snack, easy, olive, chips

Ensemble Topic Modeling

Ensemble procedures seek to encourage diversity with a view to improving the quality of the information available in the integration phase - Alexander Topchy



Recipe Recommendation Reasoning on Features : Ensemble Topic Modeling Based Approach

Extracting significant Features(words) using Ensemble Topic modeling

- Tokenize , Lemmatize, Stem , customWord
- Domain and corpus specific Stopwords generation
- Ensemble Topic modeling
- Selecting first 15 words from top 30 topics
- Calculating the weight for the Unique 288 words from the previous step

Example of identified topics

Topics discriptor	List of 15 features within each topic
Pomon	italian, beef-sausage, european, olive, parmesan, mozzarella, pork-sausage, spaghetti, chorizo, pepperoni, lasagna,
Roman	german, marinara, eggplant, romantic
Papapa based descert and drinks	banana, tropical-fruit, frozen-desserts, quick-breads, smoothie, yogurt, muffins, thermomix, walnut, blueberry,
ballalla baseu desselt allu ulliks	pudding, krispies, wheat, bran, batter
Easy and quick spacks	beginner-cook, easy, inexpensive, kid-friendly, appetizer, eggs-dairy, holiday-food, pretzel, toddler-friendly, few-
Easy and quick shacks	ingredients-recipe, served-cold, potluck, egg-free, finger-food, served-hot
Easy fish mains	salmon, fish, saltwater-fish, seafood, main-dish, high-protein, spinach, low-carb, vinaigrette, cream-cheese, halibut,
Easy fish finallis	dijon, asparagus, chickpeas, beginner-cook
Corn based maxisan	tortilla, cornstarch, cornmeal, casserole, salsa, chowder, cornbread, cilantro, jalapeno, chilies, cob, sweet-corn,
Com based maxican	popcorn, soups-stews, roma-tomato
Pico dishos	rice, side-dishes, white-rice, soy, casserole, brown-rice, chinese, one-dish-meal, basmati, long-grain, pea, risotto, pilaf,
Rice distles	stir-fry, stove-top
Tropical juice and descerts	pineapple, tropical-fruit, cherry, hawaiian, mango, maraschino, carrot, low-protein, cornstarch, dessert, soy, pudding,
Topical juice and desserts	marshmallow, easy, caribbean
Roof based mains	beef, meat, ground-beef, main-dish, one-dish-meal, roast-beef, sirloin, crock-pot-slow-cooker, meatloaf, casserole,
beer based mains	meatball, scallion, noodle, breadcrumb, easy
Vagatable dishes	vegetable, carrot, saucepan, zucchini, spinach, soups-stews, tofu, cabbage, sunflower-oil, asparagus, week-night,
vegetable disiles	chickpeas, cauliflower, hard-boiled-egg, mushroom
Seafood mains	seafood, soy-sauce, shellfish, tuna, appetizer, main-dish, wok, low-carb, leek, crabmeat, olive, cod, high-protein, easy,
Searood mains	cream-style
Turkey mains	turkey, meat, cooking-wine, main-dish, couscous, thanksgiving, turkey-breasts, tortilla, sausage, lettuce, cranberry, one-
	dish-meal, burger, gravy, casserole
Health conscious	low-cholesterol, low-protein, low-calorie, ow-carb, easy, side-dishes, soups-stews, appetizer, smoothie, few-steps-
	recipe, cucumber, egg-free, savoryappliance, salsa

The final 288 Food Features

Which we later considered the basis of new data representation format and thereby recommendation algorithms



The final 288 Food Features

maplephyllo mash tilapia^{snack} yukon indian krispies coffee^{lentil}sirloin biscuit pretzel tea blender pancake toothpick hickpeas wing roast-beef pork-chops raspberry combreadbran pine wok mango easter tuna russetoatmeal meatloaf cookies enchilada white-ricehalloween tenderloinvinaigrette semi-sweetgoat-cheese cocoa refrigerate crabmeat brownie citrus cherry meatballlasagna florentine hawaiian maraschino coffee-cakes macadamia ground-beef rolls-biscuits salad-dressings muffins mozzarellagravy sourdough saucepan pea served-cold half-an-hour breadcrumb roma-tomato soups-stewsartichoke cooking-wine pork-sausage mayonnaise beef-sausagezucchiniturkey-breasts frozen-desserts chinese quick-breads casserole toddler-friendly chorizo Spaghetti penne less-than-one-hour orzomarshmallow shallot celery_{sweet-potatoes} european carrot eggs-dairy rice broccoli pilafheirloom-historical talian ^{SOY} Orange crock-pot-slow-cooker banana bisques-cream-soupshalibut noodle cheese cake spicy northeastern-united-states batter dutch-ovenserved-hot green-yellow-beans cheese apple or more set of for-large-groups cheese apple pear beginner-cook cornstarch pasta sun-dried-tomato/capcicumsouthern-united-states chicken-breasts green-yellow-beans chicken strawberry appetizer american shredded-meatchilles pot-roast main-dish low-protein low-cholesterol low-calorienorth-american cream-cheese few-ingredients-recipe oatturkey potatoham vegetable sausage pumpkin bread shellfish few-steps-recipeeasv beef bacon^{patty} dessert olive low-carb pineapple meat ^{pudding inexpensivespringform} spinach tortilla electric-equipmentstex-mexsouthwestern-united-states dijon side-dishes hard-boiled-egg reole pies-and-tarts nacaroon whole-chicken saltwater-fish one-dish-meal leek saltwater-fish electric-equipmentstex-mexsouthwestern-united-statesdijon side-dishes parmesan high-protein almond chive pork coconut elbow-macaroni sunflower-oil midwestern south-west-pacific fettuccine creole pies-and-tarts saltwater-fish poultry pecan cornmeal peanut-butter sets cholled to the small-appliance marmalade caribbean pitted-fruit mushroom macaroon whole-chicken sweet-corn australian asparagus cream-style cauliflower finger-food pasta-shells blueberry lemongrass grapefruit cranberry smoothle marinara velveeta cucumber flour mandarin pork-ribs yolk pepperoni cod long-grain week-night black-bean feta christmas stove-top egg-free smoothie marinara basmati peanut cabbage rhubarb scallion eel bouillon canadian chowder cilantroromantic quiche soy-sauce cheddar caramel walnutyogurt linguine german wheat peach risotto burger scone potluck tidbit cupcake apricot granola_{lettuce} fried cob loaf cajun



The final 288 Food Features



Our proposed n-by-288 data representation format

Recipes	Plaintext			f_1	f_2	•	•	f_{288}
R_1	Document ₁		R_1	0.79	0	•	•	.31
R_2	Document ₂	\xrightarrow{EnsTM}	R_2	0	0		•	0
•			•	•	•	•	•	•
R_n	Document _n		R_n	0.61	1	•	•	.08

Food Features based Recommendation (FFbR)

- u_a is an user
- r_n is a new recipe
- i is each feature in r_n i=0,1,2....m
- $f_{i,ua}$ is u_a 's rating on i

$$P(u_a, r_n) = \left(\sum_{i=0}^{m} f_{i, u_a}\right)^{\prime(0, 5)}$$

Weighted Food Features based Recommendation (WFFbR)

- u_a is an user
- r_n is a new recipe
- i is each ingredient in r_n i=0,1,2....m
- $f_{i,ua}$ is u_a 's rating on i
- w_i is significance score of each feature

$$P(u_a, r_n) = \left(\sum_{i=0}^m f_{i, u_a} \times w_i\right)^{\prime(0, 5)}$$

1/- ->

Food Features based Collaborative Filtering (FFbCF)

 $P(f_{b,u_a}) = \frac{\sum_{i=0}^{n} f_{b,i}}{n}$

- u_a is an user
- i is each neighbor of u_a i=0,1,2....n
- $f_{b,i}$ is neighbor i's rating on a feature f_b
- We calculate $f_{r,ua}$ using FFbR after learning u_a 's rating on all unknown feature f_b using FFbCF

Content Based(CB)

$$P(u_a, r_n) = \frac{\sum_{i=0}^{m} Ing_{i, u_a}}{m}$$

- u_a is an user
- r_n is a new recipe
- i is each ingredient in r_n i=0,1,2....m
- Ing_{i,ua} is u_a's rating on i

User modeling and testing the recommendation algorithms



Users' spread over gender category

Users' spread over activity category



Users' spread over age range



localhost:4003/signUP#/~ × +				
→ C ③ localhost:4003/signUP#/concent/personalInfo/create_prf_profile/like/dislike/select_time/CP_collect	¢	2 \$	ß	٩
Please Type in the ingreadients you like. Put a comma after each ingredients's name. You can type in as many	as you			
want. Total number of ingredients have to be atleast 20	-			
Click next to save and continue				
rice, coconut, chicken, onion, lamb, ca				
			_/	
carrots carrot				
cabbage				
macaroni avocado				
	Su	bmit		













Demo on our developed system could be found on

https://www.youtube.com/watch?v=ujaB0FiqRwk

Results

Coverage



User preference



Results

Identifying user's food requirement and restriction



Recommending Healthy Recipe ??

Distributions of items(recipes)in my data-set over WHO and FSA Health Scales.

	Total (percentage)	_	Total (percentage)
WHO	Recipes	FSA	Recipes $n = 230872$
0	88(.04)	4	40877(17.71)
1	2732(1.18)	5	32227(13.96)
2	15603(6.76)	6	44807(19.41)
3	63969(27.71)	7	40647(17.61)
4	123453(53.47)	8	30754(13.32)
5	22395(9.7)	9	24249(10.5)
6	2533(1.1)	10	13002(5.63)
7	99(0.04)	11	4009(1.74)
8	0(0)	12	300(0.13)

Recommending Healthy Recipe ??

	%	Features
contout	4.2	holiday-food, beginner-cook, week-night, inexpensive,
context	4.2	6-people-or-more, potluck
cuisine	6.7	italian, hawaiian, tex-mex, chinese, cajun
equipment	2.4	saucepan, thermomix, wok, dutch-oven
cooking process	2 1 2	few-steps-recipe, less-than-one-hour, fried, slow-cooked,
cooking process	3.12	marinated, 4-hours-or-more
ingredient	61.15	poultry, feta, spaghetti, shredded-mea
category	15.2	risotto, lasagna, stew, appetizer, pot-roast
nutrition	3.01	high-calcium, low-cholesterol, egg-free

Looking for user group with healthier life style??

Scale	Guideline	User Group
Activity Level	FAO: activity level	<pre>sedentary, lightly_active, moderately_active, extra_active</pre>
BMI	WHO:BMI	underweight, normal_weight , pre_obesity, obesityclass_1
Food Choices	FSA: nutrient intake guideline	less_healthy, moderately_healthy, very_healthy

Looking for user group with healthier life style??



Users' spread over different Healthiness scale

Positively correlated features to healthier user groups

Average Food He	ealthScore	Activity Level			
Feature	r	Feature	r		
peanut-butter	0.447989	wing	0.441152		
granola	0.365171	tuna	0.430467		
lentil	0.360767	tilapia	0.363502		
indian	0.356347	salmon	0.359852		
cauliflower	0.352353	hawaiian	0.346401		
low-cholesterol	0.350818	canadian	0.322470		
maple	0.321131	smoothy	0.314174		
vegetable	0.307459	chicken-thighs-legs	0.314059		
wheat	0.303326	halibut	0.310990		
carrot	0.303052	main-dish	0.303345		

Negatively correlated features to healthier user groups

Average Food HealthScore		Activity Level	
	r		r
roast-beef	-0.564273476	hard-boiled	-0.345550067
lasagna	-0.464811126	chicken-breasts	-0.329366275
sausage	-0.389612531	white-rice	-0.32198579
wing	-0.364559511	rice	-0.32198579
thanksgiving	-0.350745293	ham	-0.289471946
chicken-breast	-0.329366275	pear	-0.292700772

Limitation

Feature work

Centre for Data Analytics



Thank You









