Recipe-Book Tag Prediction

A Project

Presented to the

Faculty of

NYU Tandon School of Engineering

In Partial Fulfillment

Of the Requirements for the Coursework of

ECE-GY 6143 Introduction to Machine Learning

Fall 2022

By

Sai Krishna Chaitanya Annavazala

(sa6948)

Sai Tejaswi Veeramreddy

(sv2564)

Quad Chart:

Purpose	Results
→ Given a group of recipes and relevant categories, we can associate new tags based on the existing categories. → In this project we are specifically trying to predict if recipe is vegan or non-vegan by using different classifiers.	→ I have been able to achieve the following accuracy on predicting the 'vegan' tag for recipes. KNN - 94% Random Forest Classifier - 98% Logistic Regression - 99%
Procedures	Analysis
→ The procedure for the project is as follows: 1. Cleaning of data by replacing null values with appropriate values. 2. Consider few targets which have tolerable imbalance in class data. And, we apply pipeline for logistic regression and one vs rest classifier. 3. We then apply KNN, random forest and decision tree and compare accuracy of the recipe tags. (By plotting their AUC graphs).	→ It can be concluded that we have used four classifiers namely KNN, random forestsand logistic regression to compare the accuracy of predicting recipe tags.

Motivation:

Recipe books have been around for a while. Every beginner, who starts to learn cooking goes through the recipe book to learn how to cook. But, there are various cuisines and in a world and by associating 'tags' to them, we can easily choose from the existing categories. So by using ML model, we have associated tags to food which help us in choosing our relevant food.

Why ML?

We have used Machine learning using Python as manually associating tags is a lengthy process and time consuming. By using an ML model, tags accuracy can be established for future reference and use.

Our Contribution:

→ We have taken this dataset from Kaggle. While going through the different ML solutions provided by users of the platform we have observed that there were not many analyses showing the comparison between different classifiers and how each of their prediction accuracy differs.

→ Through code, we have compared against the Logistic Regression, Random Forest Classifier and K-Nearest Neighbors Classifier to check which of these would yield a better prediction.

→ From the features to the end result the process we followed to get the results is as follows.

Dataset Description:

→ The dataset is huge with 20052 rows of recipe titles and 679 of recipe categories.

→ The column names are as follows:

```
[27] recipes.columns

    Index(['rating', 'calories', 'protein', 'fat', 'sodium', '#cakeweek',
           '#wasteless', '22-minute meals', '3-ingredient recipes',
           '30 days of groceries',
           ...
           'yellow squash', 'yogurt', 'yonkers', 'yuca', 'zucchini', 'cookbooks',
           'leftovers', 'snack', 'snack week', 'turkey'],
          dtype='object', length=679)
```

→ The structure of data in the dataset is as follows:

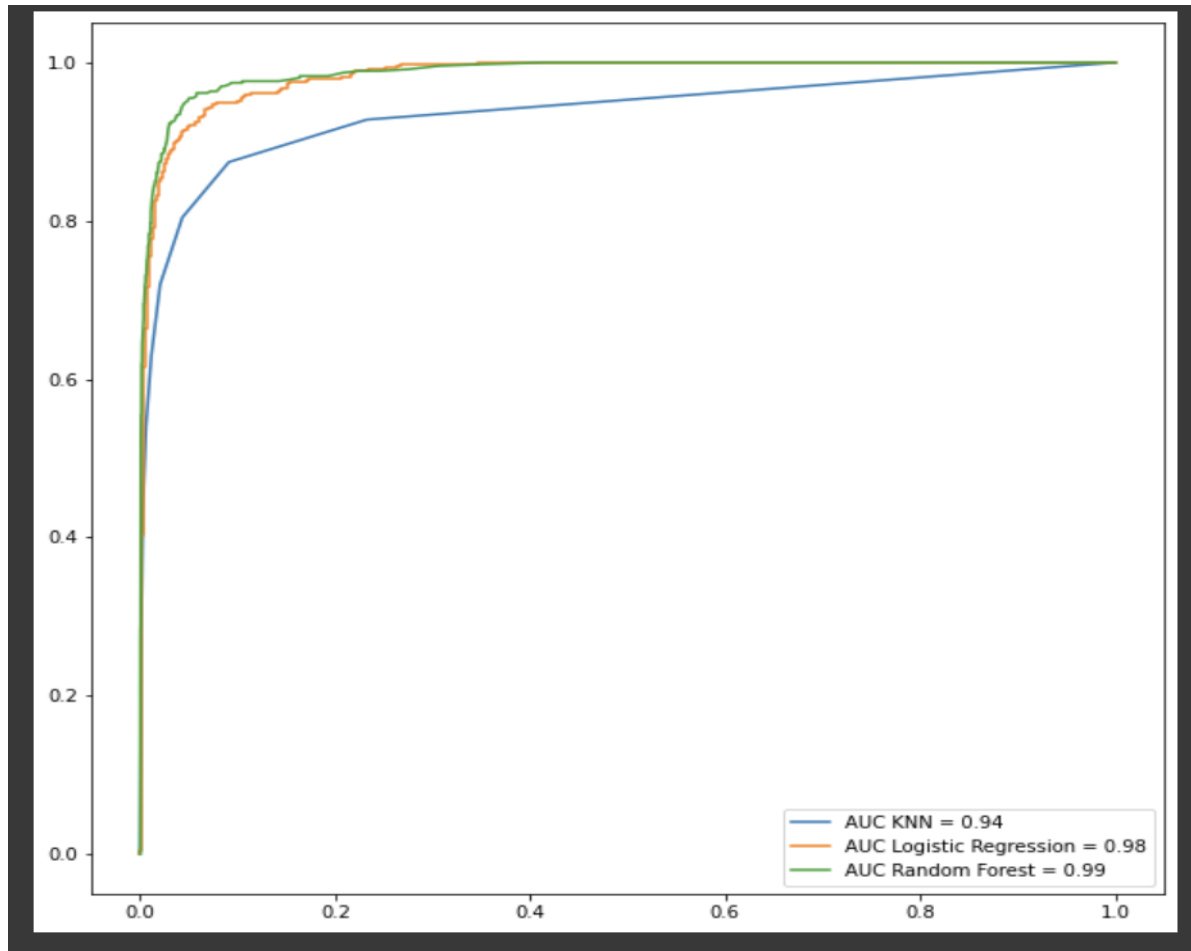| title	rating	calories	protein	fat	sodium	#cakeweek	#wasteless	22-minute meals	3-ingredient recipes	30 days of groceries	...	yellow squash	yogurt	yonkers	yuca
Lentil, Apple, and Turkey Wrap	2.500	426.0	30.0	7.0	559.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0
Boudin Blanc Terrine with Red Onion Confit	4.375	403.0	18.0	23.0	1439.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0
Potato and Fennel Soup Hodge	3.750	165.0	6.0	7.0	165.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0
Mahi-Mahi in Tomato Olive Sauce	5.000	NaN	NaN	NaN	NaN	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0
Spinach Noodle Casserole	3.125	547.0	20.0	32.0	452.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0
The Best Blts	4.375	948.0	19.0	79.0	1042.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0

6 rows × 679 columns

→ The data recipes have been categorised into following categories:
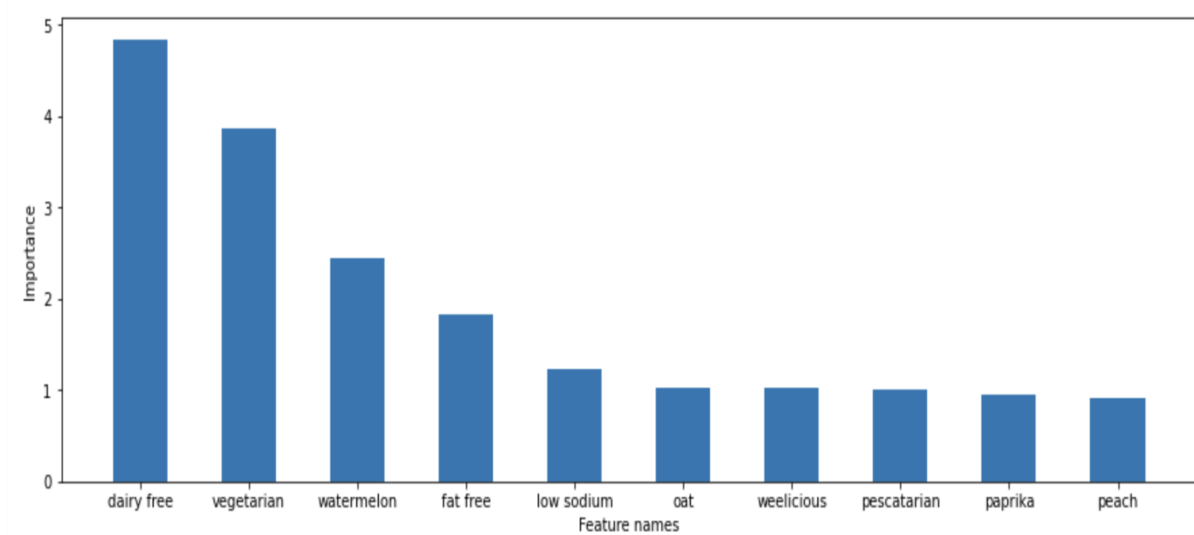
→ We have also plotted a heatmap for the correlation matrix depicting the categories of recipes with high frequency of occurrence.
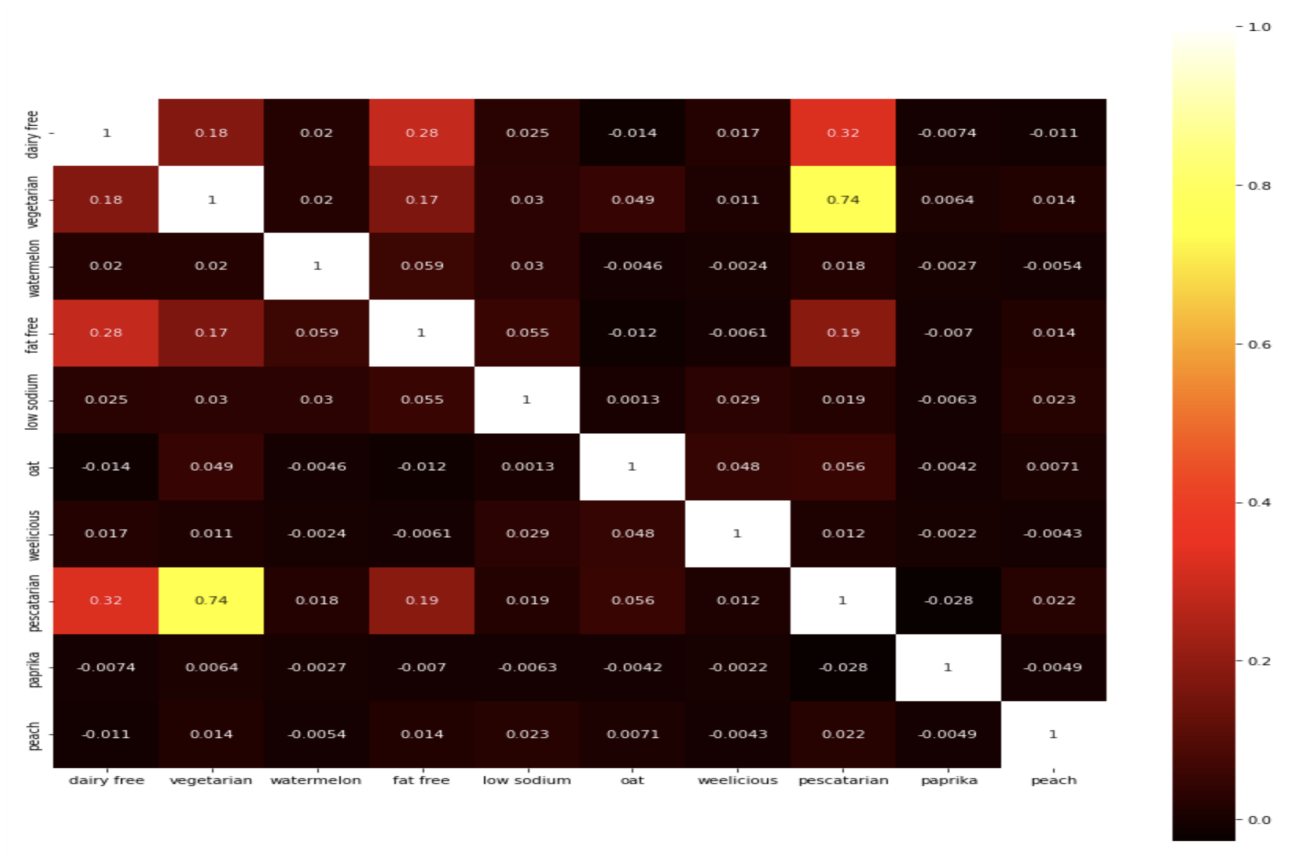
→ This graph has been plotted to represent the efficiencies of the different algorithms used to classify the recipes to be either vegan or non-vegan.

References:

1) Link to data source - https://www.kaggle.com/hugodarwood/epirecipes

2) Recipe information lifted from - http://www.epicurious.com/recipes-menus

3) Code references from the notebooks at this link:
   https://www.kaggle.com/datasets/hugodarwood/epirecipes/code

Recipebook Tag Prediction

December 23, 2022

First we will read the data using *pandas*

```python
[1]: import pandas as pd
     import numpy as np

     import matplotlib.pyplot as plt
     import seaborn as sns
     %matplotlib inline
```

```python
[2]: recipes = pd.read_csv('epi_r.csv', sep=",", index_col=0)
     recipes.head(6)
```

```
[2]:                                             rating  calories  protein   fat  \
     title
     Lentil, Apple, and Turkey Wrap              2.500     426.0     30.0   7.0
     Boudin Blanc Terrine with Red Onion Confit  4.375     403.0     18.0  23.0
     Potato and Fennel Soup Hodge                3.750     165.0      6.0   7.0
     Mahi-Mahi in Tomato Olive Sauce             5.000       NaN      NaN   NaN
     Spinach Noodle Casserole                    3.125     547.0     20.0  32.0
     The Best Blts                               4.375     948.0     19.0  79.0

                                                 sodium  #cakeweek  #wasteless  \
     title
     Lentil, Apple, and Turkey Wrap               559.0        0.0         0.0
     Boudin Blanc Terrine with Red Onion Confit  1439.0        0.0         0.0
     Potato and Fennel Soup Hodge                 165.0        0.0         0.0
     Mahi-Mahi in Tomato Olive Sauce                NaN        0.0         0.0
     Spinach Noodle Casserole                     452.0        0.0         0.0
     The Best Blts                               1042.0        0.0         0.0

                                                 22-minute meals  \
     title
     Lentil, Apple, and Turkey Wrap                          0.0
     Boudin Blanc Terrine with Red Onion Confit              0.0
     Potato and Fennel Soup Hodge                            0.0
     Mahi-Mahi in Tomato Olive Sauce                         0.0
     Spinach Noodle Casserole                                0.0
     The Best Blts                                           0.0
```

```
                                            3-ingredient recipes  \
title
Lentil, Apple, and Turkey Wrap                               0.0
Boudin Blanc Terrine with Red Onion Confit                   0.0
Potato and Fennel Soup Hodge                                 0.0
Mahi-Mahi in Tomato Olive Sauce                              0.0
Spinach Noodle Casserole                                     0.0
The Best Blts                                                0.0

                                            30 days of groceries  …  \
title                                                            …
Lentil, Apple, and Turkey Wrap                               0.0  …
Boudin Blanc Terrine with Red Onion Confit                   0.0  …
Potato and Fennel Soup Hodge                                 0.0  …
Mahi-Mahi in Tomato Olive Sauce                              0.0  …
Spinach Noodle Casserole                                     0.0  …
The Best Blts                                                0.0  …

                                            yellow squash  yogurt  yonkers  \
title
Lentil, Apple, and Turkey Wrap                        0.0     0.0      0.0
Boudin Blanc Terrine with Red Onion Confit            0.0     0.0      0.0
Potato and Fennel Soup Hodge                          0.0     0.0      0.0
Mahi-Mahi in Tomato Olive Sauce                       0.0     0.0      0.0
Spinach Noodle Casserole                              0.0     0.0      0.0
The Best Blts                                         0.0     0.0      0.0

                                            yuca  zucchini  cookbooks  \
title
Lentil, Apple, and Turkey Wrap               0.0       0.0        0.0
Boudin Blanc Terrine with Red Onion Confit   0.0       0.0        0.0
Potato and Fennel Soup Hodge                 0.0       0.0        0.0
Mahi-Mahi in Tomato Olive Sauce              0.0       0.0        0.0
Spinach Noodle Casserole                     0.0       0.0        0.0
The Best Blts                                0.0       0.0        0.0

                                            leftovers  snack  snack week  \
title
Lentil, Apple, and Turkey Wrap                    0.0    0.0         0.0
Boudin Blanc Terrine with Red Onion Confit        0.0    0.0         0.0
Potato and Fennel Soup Hodge                      0.0    0.0         0.0
Mahi-Mahi in Tomato Olive Sauce                   0.0    0.0         0.0
Spinach Noodle Casserole                          0.0    0.0         0.0
The Best Blts                                     0.0    0.0         0.0

                                            turkey
```

```
title
Lentil, Apple, and Turkey Wrap                  1.0
Boudin Blanc Terrine with Red Onion Confit      0.0
Potato and Fennel Soup Hodge                    0.0
Mahi-Mahi in Tomato Olive Sauce                 0.0
Spinach Noodle Casserole                        0.0
The Best Blts                                   0.0

[6 rows x 679 columns]
```

We see a around 20052 samples but quite a lot of features. We need to bring down the number of features to classify results better. But first let's check the summary.

[3]: `recipes.columns`

```
[3]: Index(['rating', 'calories', 'protein', 'fat', 'sodium', '#cakeweek',
            '#wasteless', '22-minute meals', '3-ingredient recipes',
            '30 days of groceries',
            …
            'yellow squash', 'yogurt', 'yonkers', 'yuca', 'zucchini', 'cookbooks',
            'leftovers', 'snack', 'snack week', 'turkey'],
          dtype='object', length=679)
```

[4]: `recipes.shape`

[4]: `(20052, 679)`

Data Summary

[5]:
```python
def getDfSummary(input_data):
    # Place your code here
    output_data = input_data.describe().transpose()
    output_data["number_distinct"] = input_data.nunique()
    output_data["number_nan"] = input_data.isna().sum()
    return output_data
```

[6]:
```python
recipes_stat = getDfSummary(recipes)
recipes_nan_stats = recipes_stat.loc[recipes_stat['number_nan'] != 0]
recipes_full_stats = recipes_stat.loc[recipes_stat['number_nan'] == 0]

print(recipes_nan_stats['number_nan'])
recipes_full_stats
```

```
calories    4117
protein     4162
fat         4183
sodium      4119
Name: number_nan, dtype: int64
```

```
[6]:                       count      mean       std  min   25%    50%    75%  \
     rating              20052.0  3.714467  1.340829  0.0  3.75  4.375  4.375
     #cakeweek           20052.0  0.000299  0.017296  0.0  0.00  0.000  0.000
     #wasteless          20052.0  0.000050  0.007062  0.0  0.00  0.000  0.000
     22-minute meals     20052.0  0.000848  0.029105  0.0  0.00  0.000  0.000
     3-ingredient recipes 20052.0  0.001346  0.036671  0.0  0.00  0.000  0.000
     ...                      ...       ...       ...  ...   ...    ...    ...
     cookbooks           20052.0  0.000150  0.012231  0.0  0.00  0.000  0.000
     leftovers           20052.0  0.000349  0.018681  0.0  0.00  0.000  0.000
     snack               20052.0  0.001396  0.037343  0.0  0.00  0.000  0.000
     snack week          20052.0  0.000948  0.030768  0.0  0.00  0.000  0.000
     turkey              20052.0  0.022741  0.149080  0.0  0.00  0.000  0.000

                          max  number_distinct  number_nan
     rating               5.0                8           0
     #cakeweek            1.0                2           0
     #wasteless           1.0                2           0
     22-minute meals      1.0                2           0
     3-ingredient recipes 1.0                2           0
     ...                  ...              ...         ...
     cookbooks            1.0                2           0
     leftovers            1.0                2           0
     snack                1.0                2           0
     snack week           1.0                2           0
     turkey               1.0                2           0

     [675 rows x 10 columns]
```

We see that some entries above have nan values. It may not be useful for analysis. Hence we remove them.

Data Cleaning and Pre-Processing

```
[7]: recipes_nan_records = recipes[recipes.isna().any(axis=1)]
     recipes_full_records = recipes[~recipes.isna().any(axis=1)]

     recipes_stat = getDfSummary(recipes_full_records)
     recipes_stat
```

```
[7]:                count          mean            std  min     25%      50%  \
     rating       15864.0      3.760952       1.285518  0.0    3.75    4.375
     calories     15864.0   6350.682993  359848.417868  0.0  199.00  333.000
     protein      15864.0    100.324571    3843.462312  0.0    3.00    8.000
     fat          15864.0    346.986826   20459.329549  0.0    7.00   17.000
     sodium       15864.0   6252.742310  334042.078448  0.0   82.00  296.000
     ...              ...           ...            ...  ...     ...      ...
     cookbooks    15864.0      0.000126       0.011228  0.0    0.00    0.000
     leftovers    15864.0      0.000315       0.017751  0.0    0.00    0.000
```

4

```
snack          15864.0      0.001324          0.036360  0.0     0.00     0.000
snack week     15864.0      0.000819          0.028615  0.0     0.00     0.000
turkey         15864.0      0.021243          0.144198  0.0     0.00     0.000

                  75%          max  number_distinct  number_nan
rating           4.375          5.0                8           0
calories       587.250   30111218.0             1858           0
protein         27.000     236489.0              282           0
fat             33.000    1722763.0              326           0
sodium         713.000   27675110.0             2433           0
...                ...          ...              ...         ...
cookbooks        0.000          1.0                2           0
leftovers        0.000          1.0                2           0
snack            0.000          1.0                2           0
snack week       0.000          1.0                2           0
turkey           0.000          1.0                2           0

[679 rows x 10 columns]
```

We then attempt to take only the binary valued entries for target since it will be hard to predict ratings or calories with just binary entries.

```python
[8]: categories = []
     recipes_bin_cols = recipes.isin([0,1]).all()
     for col in recipes.columns:
         if recipes_bin_cols[col] == True:
             categories.append(col)

     categories
```

```
[8]: ['#cakeweek',
      '#wasteless',
      '22-minute meals',
      '3-ingredient recipes',
      '30 days of groceries',
      'advance prep required',
      'alabama',
      'alaska',
      'alcoholic',
      'almond',
      'amaretto',
      'anchovy',
      'anise',
      'anniversary',
      'anthony bourdain',
      'aperitif',
      'appetizer',
```

```
'apple',
'apple juice',
'apricot',
'arizona',
'artichoke',
'arugula',
'asian pear',
'asparagus',
'aspen',
'atlanta',
'australia',
'avocado',
'back to school',
'backyard bbq',
'bacon',
'bake',
'banana',
'barley',
'basil',
'bass',
'bastille day',
'bean',
'beef',
'beef rib',
'beef shank',
'beef tenderloin',
'beer',
'beet',
'bell pepper',
'berry',
'beverly hills',
'birthday',
'biscuit',
'bitters',
'blackberry',
'blender',
'blue cheese',
'blueberry',
'boil',
'bok choy',
'bon appétit',
'bon app tit',
'boston',
'bourbon',
'braise',
'bran',
'brandy',
```

```
'bread',
'breadcrumbs',
'breakfast',
'brie',
'brine',
'brisket',
'broccoli',
'broccoli rabe',
'broil',
'brooklyn',
'brown rice',
'brownie',
'brunch',
'brussel sprout',
'buffalo',
'buffet',
'bulgaria',
'bulgur',
'burrito',
'butter',
'buttermilk',
'butternut squash',
'butterscotch/caramel',
'cabbage',
'cake',
'california',
'calvados',
'cambridge',
'campari',
'camping',
'canada',
'candy',
'candy thermometer',
'cantaloupe',
'capers',
'caraway',
'cardamom',
'carrot',
'cashew',
'casserole/gratin',
'cauliflower',
'caviar',
'celery',
'chambord',
'champagne',
'chard',
'chartreuse',
```

```
'cheddar',
'cheese',
'cherry',
'chestnut',
'chicago',
'chicken',
'chickpea',
'chile',
'chile pepper',
'chili',
'chill',
'chive',
'chocolate',
'christmas',
'christmas eve',
'cilantro',
'cinco de mayo',
'cinnamon',
'citrus',
'clam',
'clove',
'cobbler/crumble',
'cocktail',
'cocktail party',
'coconut',
'cod',
'coffee',
'coffee grinder',
'cognac/armagnac',
'collard greens',
'colorado',
'columbus',
'condiment',
'condiment/spread',
'connecticut',
'cook like a diner',
'cookbook critic',
'cookie',
'cookies',
'coriander',
'corn',
'cornmeal',
'costa mesa',
'cottage cheese',
'couscous',
'crab',
'cranberry',
```

```
'cranberry sauce',
'cream cheese',
'créme de cacao',
'crêpe',
'cr me de cacao',
'cuba',
'cucumber',
'cumin',
'cupcake',
'currant',
'curry',
'custard',
'dairy',
'dairy free',
'dallas',
'date',
'deep-fry',
'denver',
'dessert',
'digestif',
'dill',
'dinner',
'dip',
'diwali',
'dominican republic',
'dorie greenspan',
'double boiler',
'dried fruit',
'drink',
'drinks',
'duck',
'easter',
'eau de vie',
'edible gift',
'egg',
'egg nog',
'eggplant',
'egypt',
'emeril lagasse',
'endive',
'engagement party',
'england',
'entertaining',
'epi + ushg',
'epi loves the microwave',
'escarole',
'fall',
```

```
'family reunion',
'fat free',
"father's day",
'fennel',
'feta',
'fig',
'fish',
'flaming hot summer',
'flat bread',
'florida',
'fontina',
'food processor',
'fortified wine',
'fourth of july',
'france',
'frangelico',
'frankenrecipe',
'freeze/chill',
'freezer food',
'friendsgiving',
'frittata',
'fritter',
'frozen dessert',
'fruit',
'fruit juice',
'fry',
'game',
'garlic',
'georgia',
'germany',
'gin',
'ginger',
'goat cheese',
'goose',
'gouda',
'gourmet',
'graduation',
'grains',
'grand marnier',
'granola',
'grape',
'grapefruit',
'grappa',
'green bean',
'green onion/scallion',
'grill',
'grill/barbecue',
```

```
'ground beef',
'ground lamb',
'guam',
'guava',
'haiti',
'halibut',
'halloween',
'ham',
'hamburger',
'hanukkah',
'harpercollins',
'hawaii',
'hazelnut',
'healdsburg',
'healthy',
'herb',
'high fiber',
'hollywood',
'hominy/cornmeal/masa',
'honey',
'honeydew',
"hors d'oeuvre",
'horseradish',
'hot drink',
'hot pepper',
'house & garden',
'house cocktail',
'houston',
'hummus',
'ice cream',
'ice cream machine',
'iced coffee',
'iced tea',
'idaho',
'illinois',
'indiana',
'iowa',
'ireland',
'israel',
'italy',
'jalapeño',
'jam or jelly',
'jamaica',
'japan',
'jerusalem artichoke',
'juicer',
'jícama',
```

'kahlúa',
'kale',
'kansas',
'kansas city',
'kentucky',
'kentucky derby',
'kid-friendly',
'kidney friendly',
'kirsch',
'kitchen olympics',
'kiwi',
'kosher',
'kosher for passover',
'kumquat',
'kwanzaa',
'labor day',
'lamb',
'lamb chop',
'lamb shank',
'lancaster',
'las vegas',
'lasagna',
'leafy green',
'leek',
'legume',
'lemon',
'lemon juice',
'lemongrass',
'lentil',
'lettuce',
'lima bean',
'lime',
'lime juice',
'lingonberry',
'liqueur',
'lobster',
'london',
'long beach',
'los angeles',
'louisiana',
'louisville',
'low cal',
'low carb',
'low cholesterol',
'low fat',
'low sodium',
'low sugar',

```
'low/no sugar',
'lunar new year',
'lunch',
'lychee',
'macadamia nut',
'macaroni and cheese',
'maine',
'mandoline',
'mango',
'maple syrup',
'mardi gras',
'margarita',
'marinade',
'marinate',
'marsala',
'marscarpone',
'marshmallow',
'martini',
'maryland',
'massachusetts',
'mayonnaise',
'meat',
'meatball',
'meatloaf',
'melon',
'mexico',
'mezcal',
'miami',
'michigan',
'microwave',
'midori',
'milk/cream',
'minneapolis',
'minnesota',
'mint',
'mississippi',
'missouri',
'mixer',
'molasses',
'monterey jack',
'mortar and pestle',
"mother's day",
'mozzarella',
'muffin',
'mushroom',
'mussel',
'mustard',
```

```
'mustard greens',
'nancy silverton',
'nebraska',
'nectarine',
'new hampshire',
'new jersey',
'new mexico',
'new orleans',
"new year's day",
"new year's eve",
'new york',
'no meat, no problem',
'no sugar added',
'no-cook',
'non-alcoholic',
'noodle',
'north carolina',
'nut',
'nutmeg',
'oat',
'oatmeal',
'octopus',
'ohio',
'oklahoma',
'okra',
'oktoberfest',
'olive',
'omelet',
'one-pot meal',
'onion',
'orange',
'orange juice',
'oregano',
'oregon',
'organic',
'orzo',
'oscars',
'oyster',
'pacific palisades',
'paleo',
'pan-fry',
'pancake',
'papaya',
'paprika',
'parade',
'paris',
'parmesan',
```

'parsley',
'parsnip',
'party',
'pasadena',
'passion fruit',
'passover',
'pasta',
'pasta maker',
'pastry',
'pea',
'peach',
'peanut',
'peanut butter',
'peanut free',
'pear',
'pecan',
'pennsylvania',
'pepper',
'pernod',
'persian new year',
'persimmon',
'peru',
'pescatarian',
'philippines',
'phyllo/puff pastry dough',
'pickles',
'picnic',
'pie',
'pine nut',
'pineapple',
'pistachio',
'pittsburgh',
'pizza',
'plantain',
'plum',
'poach',
'poblano',
'poker/game night',
'pomegranate',
'pomegranate juice',
'poppy',
'pork',
'pork chop',
'pork rib',
'pork tenderloin',
'port',
'portland',

```
'pot pie',
'potato',
'potato salad',
'potluck',
'poultry',
'poultry sausage',
'pressure cooker',
'prosciutto',
'providence',
'prune',
'pumpkin',
'punch',
'purim',
'quail',
'quiche',
'quick & easy',
'quick and healthy',
'quince',
'quinoa',
'rabbit',
'rack of lamb',
'radicchio',
'radish',
'raisin',
'ramadan',
'ramekin',
'raspberry',
'raw',
'red wine',
'rhode island',
'rhubarb',
'rice',
'ricotta',
'roast',
'root vegetable',
'rosemary',
'rosh hashanah/yom kippur',
'rosé',
'rub',
'rum',
'rutabaga',
'rye',
'saffron',
'sage',
'sake',
'salad',
'salad dressing',
```

```
'salmon',
'salsa',
'san francisco',
'sandwich',
'sandwich theory',
'sangria',
'santa monica',
'sardine',
'sauce',
'sausage',
'sauté',
'scallop',
'scotch',
'seafood',
'seattle',
'seed',
'self',
'semolina',
'sesame',
'sesame oil',
'shallot',
'shavuot',
'shellfish',
'sherry',
'shower',
'shrimp',
'side',
'simmer',
'skewer',
'slow cooker',
'smoker',
'smoothie',
'snapper',
'sorbet',
'soufflé/meringue',
'soup/stew',
'sour cream',
'sourdough',
'south carolina',
'soy',
'soy free',
'soy sauce',
'spain',
'sparkling wine',
'spice',
'spinach',
'spirit',
```

```
'spring',
'spritzer',
'squash',
'squid',
'st. louis',
"st. patrick's day",
'steak',
'steam',
'stew',
'stir-fry',
'stock',
'strawberry',
'stuffing/dressing',
'sugar conscious',
'sugar snap pea',
'sukkot',
'summer',
'super bowl',
'suzanne goin',
'sweet potato/yam',
'swiss cheese',
'switzerland',
'swordfish',
'taco',
'tailgating',
'tamarind',
'tangerine',
'tapioca',
'tarragon',
'tart',
'tea',
'tennessee',
'tequila',
'tested & improved',
'texas',
'thanksgiving',
'thyme',
'tilapia',
'tofu',
'tomatillo',
'tomato',
'tortillas',
'tree nut',
'tree nut free',
'triple sec',
'tropical fruit',
'trout',
```

```
    'tuna',
    'turnip',
    'utah',
    "valentine's day",
    'vanilla',
    'veal',
    'vegan',
    'vegetable',
    'vegetarian',
    'venison',
    'vermont',
    'vermouth',
    'vinegar',
    'virginia',
    'vodka',
    'waffle',
    'walnut',
    'wasabi',
    'washington',
    'washington, d.c.',
    'watercress',
    'watermelon',
    'wedding',
    'weelicious',
    'west virginia',
    'westwood',
    'wheat/gluten-free',
    'whiskey',
    'white wine',
    'whole wheat',
    'wild rice',
    'windsor',
    'wine',
    'winter',
    'wisconsin',
    'wok',
    'yellow squash',
    'yogurt',
    'yonkers',
    'yuca',
    'zucchini',
    'cookbooks',
    'leftovers',
    'snack',
    'snack week',
    'turkey']
```

We consider only few targets which has tolerable inbalance in class data.

```
[9]: diffdict = {}
     for i, category in enumerate(categories):
         freqarr = recipes_full_records.loc[:, recipes.columns== category].to_numpy()
         freqarr0 = freqarr[freqarr == 0]
         freqarr1 = freqarr[freqarr == 1]
         count0 = len(freqarr0)
         count1 = len(freqarr1)
         diff = count0 - count1
         if diff < 0:
             diff = -1* diff
         diffdict[category] = diff, count0, count1
     diffdict = dict(sorted(diffdict.items(), key=lambda item: item[1]))
     # print(diffdict)
     imbalanced_dict = {k:v for (k,v) in diffdict.items() if v[0] >= 13000}
     targets = list(imbalanced_dict.keys())
     # print(targets)

     values = np.array([list(x) for x in imbalanced_dict.values()])
     N = 5
     # print(values[:N, :])

     width = 0.5
     ind = np.arange(N)
     p1 = plt.bar(ind, values[:N, 1], width)
     p2 = plt.bar(ind, values[:N, 2], width,
                  bottom=values[:N, 1])

     plt.ylabel('Count')
     plt.title('Data Imbalanced')
     plt.xticks(ind, targets[:N])
     plt.legend((p1[0], p2[0]), ('Count 0', 'Count 1'))

     plt.show()
```

Data Imbalanced
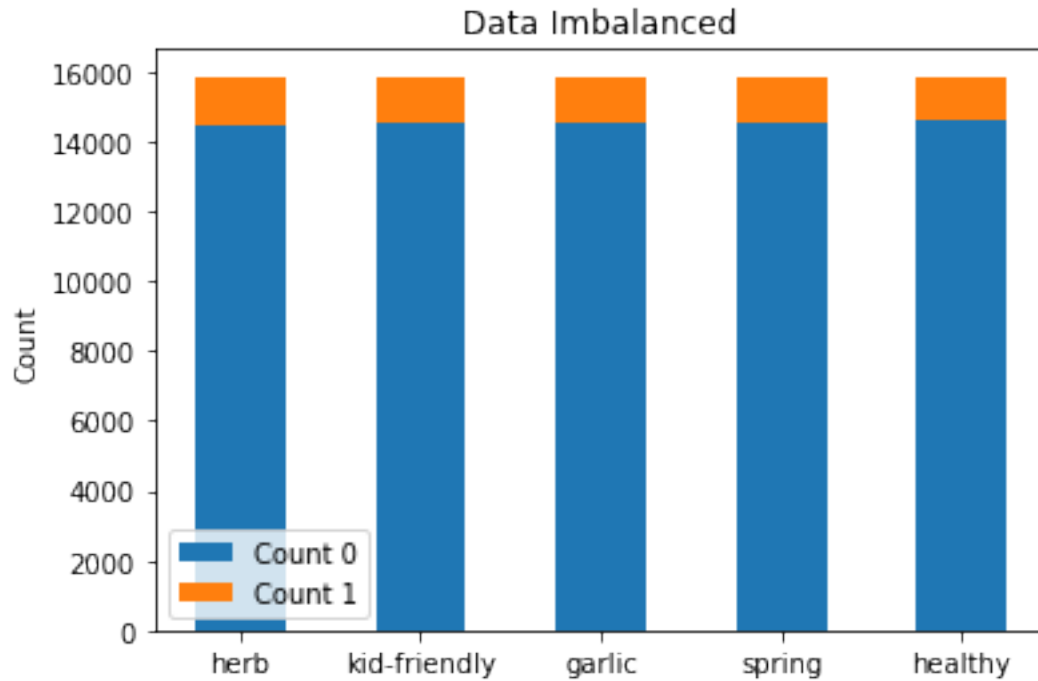
```
[10]: balanced_dict = {k:v for (k,v) in diffdict.items() if v[0] <= 13000}
      targets = list(balanced_dict.keys())
      # print(targets)

      values = np.array([list(x) for x in balanced_dict.values()])
      N = 5
      # print(values[:N, :])

      width = 0.5
      ind = np.arange(N)
      p1 = plt.bar(ind, values[:N, 1], width)
      p2 = plt.bar(ind, values[:N, 2], width,
                   bottom=values[:N, 1])

      plt.ylabel('Count')
      plt.title('Data Balanced')
      plt.xticks(ind, targets[:N])
      plt.legend((p1[0], p2[0]), ('Count 0', 'Count 1'))

      plt.show()
```
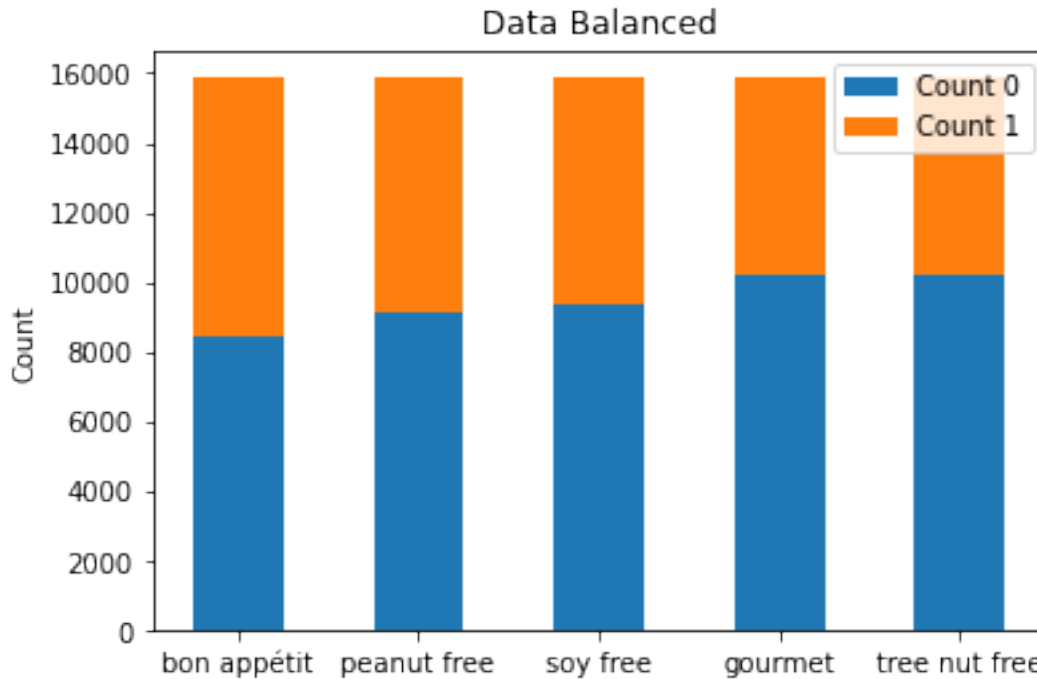
We choose the ones which has correlation to our targets above using decision tree classifier.

```
[11]:  # categories.extend(recipes.columns[:5])
       from sklearn.tree import DecisionTreeClassifier
       from sklearn.model_selection import train_test_split



       def getTestTrainData(X, y):
           xtr, xts, ytr, yts = train_test_split(X, y, test_size=0.33)
           xtr, xts, ytr, yts = xtr.to_numpy(), xts.to_numpy(), ytr.to_numpy().
       ↪ravel(), yts.to_numpy().ravel()
           return xtr, xts, ytr, yts

       def getPredictors(target):
           dt_classifier = DecisionTreeClassifier(criterion="entropy")
           categories.remove(target)
           X = recipes_full_records.loc[:, categories]
           y = recipes_full_records.loc[:, recipes.columns== target]
           xtr, xts, ytr, yts = getTestTrainData(X, y)
           dt_classifier.fit(xtr, ytr)
           predictors = []
           for i, column in enumerate(categories):
               if(dt_classifier.feature_importances_[i] > 0):
                   predictors.append(column)
```

```
        categories.append(target)
        return predictors
```

Since we have multiple targets to predict, we will use a pipeline to apply logistic regression and one vs rest classifier.

Describing targets to predict

```python
[12]: from sklearn.linear_model import LogisticRegression
      from sklearn.pipeline import Pipeline
      from sklearn.metrics import accuracy_score
      from sklearn.multiclass import OneVsRestClassifier# Using pipeline for applying␣
       ↪logistic regression and one vs rest classifier
      from sklearn.metrics import precision_recall_fscore_support

      def getMetrics(yts, yhat):
          accuracy = accuracy_score(yts, yhat)
          precision,recall,f1_score,_  =␣
       ↪precision_recall_fscore_support(yts,yhat,average='binary')
          return accuracy, precision, recall, f1_score




      LogReg_pipeline = Pipeline([
                      ('clf',␣
       ↪OneVsRestClassifier(LogisticRegression(solver='liblinear'), n_jobs=-1)),
                  ])


      for target in targets:

          # Training logistic regression model on train data
          predictors = getPredictors(target)
          X = recipes_full_records.loc[:, predictors]
          y = recipes_full_records.loc[:, recipes.columns== target]

          xtr, xts, ytr, yts = getTestTrainData(X, y)
          LogReg_pipeline.fit(xtr, ytr)

          # calculating test accuracy
          yhat = LogReg_pipeline.predict(xts)
          accuracy, precision, recall, f1_score = getMetrics(yts, yhat)
          print("Category {}".format(target))
          print("predictors "+ str(predictors))
          unique, counts = np.unique(ytr, return_counts=True)
          print(dict(zip(unique, counts)))
          unique, counts = np.unique(yts, return_counts=True)
```

23

```python
    print(dict(zip(unique, counts)))
    print('Test accuracy of {} is {}'.format(target, accuracy))
    print('Test precision of {} is {}'.format(target, precision))
    print('Test recall of {} is {}'.format(target, recall))
    print('Test f1-score of {} is {}'.format(target, f1_score))


    print("\n")
```

Category bon appétit
predictors ['#cakeweek', '22-minute meals', '3-ingredient recipes', '30 days of
groceries', 'advance prep required', 'alcoholic', 'almond', 'amaretto',
'anchovy', 'anise', 'appetizer', 'apple', 'apple juice', 'apricot', 'arizona',
'artichoke', 'arugula', 'asparagus', 'avocado', 'back to school', 'backyard
bbq', 'bacon', 'bake', 'banana', 'basil', 'bean', 'beef', 'beet', 'bell pepper',
'berry', 'birthday', 'biscuit', 'blackberry', 'blender', 'blue cheese',
'blueberry', 'boil', 'bok choy', 'bon app tit', 'braise', 'brandy', 'bread',
'breadcrumbs', 'breakfast', 'broccoli', 'broccoli rabe', 'broil', 'brunch',
'buffet', 'butter', 'buttermilk', 'butternut squash', 'cabbage', 'cake',
'california', 'campari', 'candy', 'candy thermometer', 'capers', 'cardamom',
'carrot', 'cashew', 'cauliflower', 'celery', 'champagne', 'chard', 'chartreuse',
'cheddar', 'cheese', 'cherry', 'chestnut', 'chicken', 'chile pepper', 'chill',
'chive', 'chocolate', 'christmas', 'christmas eve', 'cilantro', 'cinco de mayo',
'cinnamon', 'citrus', 'clam', 'cobbler/crumble', 'cocktail', 'cocktail party',
'coconut', 'cod', 'coffee', 'condiment', 'condiment/spread', 'connecticut',
'cook like a diner', 'cookbook critic', 'cookie', 'cookies', 'coriander',
'corn', 'cornmeal', 'cranberry', 'cucumber', 'cumin', 'currant', 'curry',
'dairy', 'dairy free', 'date', 'deep-fry', 'dessert', 'dill', 'dinner',
'diwali', 'drink', 'duck', 'easter', 'edible gift', 'egg', 'eggplant', 'endive',
'escarole', 'fall', 'family reunion', 'fat free', "father's day", 'fennel',
'feta', 'fig', 'fish', 'flaming hot summer', 'food processor', 'fourth of july',
'freeze/chill', 'freezer food', 'friendsgiving', 'frozen dessert', 'fruit',
'fruit juice', 'fry', 'game', 'garlic', 'gin', 'ginger', 'goat cheese',
'gourmet', 'grapefruit', 'green onion/scallion', 'grill', 'grill/barbecue',
'ground beef', 'guava', 'halibut', 'halloween', 'ham', 'hanukkah',
'harpercollins', 'hazelnut', 'healthy', 'herb', 'high fiber',
'hominy/cornmeal/masa', 'honey', "hors d'oeuvre", 'horseradish', 'hot pepper',
'house & garden', 'house cocktail', 'ice cream machine', 'jalapeño', 'jam or
jelly', 'jerusalem artichoke', 'jícama', 'kale', 'kid-friendly', 'kidney
friendly', 'kosher', 'kosher for passover', 'kumquat', 'lamb', 'leafy green',
'leek', 'lemon', 'lemon juice', 'lemongrass', 'lentil', 'lettuce', 'lime', 'lime
juice', 'liqueur', 'los angeles', 'low cal', 'low carb', 'low cholesterol', 'low
fat', 'low/no sugar', 'lunar new year', 'lunch', 'macadamia nut', 'mandoline',
'mango', 'mardi gras', 'margarita', 'marinate', 'marsala', 'mayonnaise', 'meat',
'melon', 'milk/cream', 'mint', 'mixer', 'monterey jack', "mother's day",
'mozzarella', 'muffin', 'mushroom', 'mussel', 'mustard', 'mustard greens',
'nectarine', "new year's day", "new year's eve", 'no sugar added', 'no-cook',

'non-alcoholic', 'noodle', 'nut', 'nutmeg', 'oat', 'olive', 'onion', 'orange',
'orange juice', 'oregano', 'paleo', 'pan-fry', 'paprika', 'parade', 'parmesan',
'parsley', 'parsnip', 'party', 'passover', 'pasta', 'pastry', 'peach', 'peanut',
'peanut free', 'pear', 'pecan', 'pepper', 'pernod', 'pescatarian', 'phyllo/puff
pastry dough', 'picnic', 'pie', 'pine nut', 'pineapple', 'pistachio', 'plum',
'poach', 'pomegranate', 'pork', 'pork chop', 'pork tenderloin', 'port',
'potato', 'poultry', 'prune', 'pumpkin', 'punch', 'quail', 'quick & easy',
'quick and healthy', 'rack of lamb', 'radicchio', 'radish', 'raisin', 'ramadan',
'raspberry', 'rice', 'ricotta', 'roast', 'root vegetable', 'rosemary', 'rosh
hashanah/yom kippur', 'rum', 'rye', 'sage', 'sake', 'salad', 'salad dressing',
'salmon', 'sandwich', 'sandwich theory', 'sauce', 'sausage', 'sauté', 'scotch',
'seafood', 'seattle', 'seed', 'self', 'semolina', 'sesame', 'sesame oil',
'shallot', 'shellfish', 'sherry', 'shower', 'shrimp', 'side', 'simmer', 'slow
cooker', 'smoker', 'smoothie', 'soufflé/meringue', 'soup/stew', 'sour cream',
'soy', 'soy free', 'soy sauce', 'sparkling wine', 'spice', 'spinach', 'spring',
'squash', "st. patrick's day", 'steak', 'stew', 'stir-fry', 'strawberry',
'stuffing/dressing', 'sugar conscious', 'sugar snap pea', 'summer', 'super
bowl', 'sweet potato/yam', 'swiss cheese', 'taco', 'tailgating', 'tangerine',
'tarragon', 'tart', 'tea', 'texas', 'thanksgiving', 'thyme', 'tomatillo',
'tomato', 'tree nut', 'tree nut free', 'tropical fruit', 'trout', 'tuna',
"valentine's day", 'vanilla', 'veal', 'vegan', 'vegetable', 'vegetarian',
'vermouth', 'vinegar', 'vodka', 'walnut', 'watercress', 'watermelon', 'wedding',
'weelicious', 'wheat/gluten-free', 'whiskey', 'white wine', 'whole wheat',
'wine', 'winter', 'wok', 'yogurt', 'zucchini', 'cookbooks', 'snack', 'snack
week', 'turkey']
{0.0: 5674, 1.0: 4954}
{0.0: 2807, 1.0: 2429}
Test accuracy of bon appétit is 0.8794881588999236
Test precision of bon appétit is 0.8215307582260372
Test recall of bon appétit is 0.9456566488266777
Test f1-score of bon appétit is 0.8792344497607655

Category peanut free
predictors ['advance prep required', 'anniversary', 'appetizer', 'apple',
'artichoke', 'arugula', 'bake', 'bean', 'berry', 'broccoli', 'brown rice',
'brussel sprout', 'butter', 'cabbage', 'cauliflower', 'christmas eve',
'coriander', 'cream cheese', 'cucumber', 'currant', 'curry', 'dairy free',
'dinner', 'duck', 'egg', 'fry', 'gourmet', 'grape', 'healthy', 'herb', 'high
fiber', 'jalapeño', 'kid-friendly', 'kidney friendly', 'kosher', 'leafy green',
'leek', 'legume', 'lemon', 'lemongrass', 'lime juice', 'low fat', 'lunch',
'milk/cream', "new year's eve", 'no sugar added', 'noodle', 'nut', 'nutmeg',
'oat', 'olive', 'parsley', 'peanut', 'peanut butter', 'pescatarian', 'pine nut',
'pomegranate', 'pork', 'potato', 'poultry', 'quick & easy', 'quinoa', 'raisin',
'scallop', 'side', 'soy free', 'soy sauce', 'spring', 'steak', 'stir-fry',
'sugar conscious', 'thanksgiving', 'tree nut free', 'tuna', "valentine's day",
'vegan', 'vegetarian', 'vinegar', 'wheat/gluten-free', 'winter', 'wok',
'yogurt']

```
{0.0: 6141, 1.0: 4487}
{0.0: 3002, 1.0: 2234}
Test accuracy of peanut free is 0.987012987012987
Test precision of peanut free is 0.9821905609973286
Test recall of peanut free is 0.9874664279319606
Test f1-score of peanut free is 0.9848214285714285


Category soy free
predictors ['amaretto', 'anise', 'apple', 'apricot', 'backyard bbq', 'bacon',
'bake', 'blackberry', 'blender', 'bok choy', 'brandy', 'breakfast', 'brunch',
'buffet', 'buttermilk', 'cabbage', 'carrot', 'chicken', 'chile pepper',
'chocolate', 'christmas', 'coconut', 'crab', 'cucumber', 'dairy', 'dairy free',
'dessert', 'dinner', 'egg', 'fall', 'fish', 'fruit', 'ginger', 'green
onion/scallion', 'grill', 'grill/barbecue', 'ground beef', 'ham',
'harpercollins', 'high fiber', 'hot pepper', 'kidney friendly', 'kosher', 'leafy
green', 'lemongrass', 'liqueur', 'low sodium', 'lunar new year', 'lunch',
'marinate', 'marsala', 'meat', 'mushroom', 'mustard greens', "new year's day",
'no sugar added', 'no-cook', 'non-alcoholic', 'noodle', 'nut', 'onion',
'orange', 'orange juice', 'party', 'pea', 'persimmon', 'pescatarian',
'pistachio', 'plum', 'poker/game night', 'portland', 'pumpkin', 'quick & easy',
'rice', 'roast', 'salad', 'sandwich', 'sandwich theory', 'sauté', 'scallop',
'seafood', 'sesame', 'shrimp', 'side', 'soy', 'soy sauce', 'spring', 'squid',
'steam', 'stir-fry', 'sugar conscious', 'tailgating', 'thanksgiving', 'tofu',
'tomato', 'tree nut', 'tree nut free', 'tuna', 'vegetable', 'vegetarian',
'vodka', 'watercress', 'wedding', 'wheat/gluten-free', 'whole wheat', 'wine',
'winter', 'yogurt', 'bon appétit', 'peanut free']
{0.0: 6284, 1.0: 4344}
{0.0: 3078, 1.0: 2158}
Test accuracy of soy free is 0.9845301757066462
Test precision of soy free is 0.976594768242313
Test recall of soy free is 0.9860982391102873
Test f1-score of soy free is 0.9813234955038045


Category gourmet
predictors ['#cakeweek', '22-minute meals', '3-ingredient recipes', '30 days of
groceries', 'advance prep required', 'alcoholic', 'almond', 'amaretto',
'anchovy', 'anise', 'anniversary', 'appetizer', 'apple', 'apple juice',
'apricot', 'artichoke', 'arugula', 'back to school', 'backyard bbq', 'bacon',
'bake', 'banana', 'basil', 'bean', 'beef', 'beef tenderloin', 'beer', 'bell
pepper', 'berry', 'birthday', 'bitters', 'blackberry', 'blender', 'blue cheese',
'blueberry', 'bok choy', 'bon app tit', 'bourbon', 'braise', 'brandy', 'bread',
'breadcrumbs', 'breakfast', 'brisket', 'broccoli', 'broccoli rabe', 'broil',
'brunch', 'buffet', 'butter', 'buttermilk', 'butternut squash', 'cabbage',
'cake', 'calvados', 'candy', 'candy thermometer', 'capers', 'carrot', 'cashew',
'cauliflower', 'celery', 'champagne', 'chard', 'cheddar', 'cheese', 'cherry',
'chestnut', 'chicken', 'chile pepper', 'chill', 'chive', 'chocolate',
```

'christmas', 'christmas eve', 'cilantro', 'cinco de mayo', 'cinnamon', 'citrus', 'clam', 'clove', 'cocktail', 'cocktail party', 'coconut', 'cod', 'coffee', 'cognac/armagnac', 'condiment', 'condiment/spread', 'cookie', 'cookies', 'coriander', 'corn', 'cornmeal', 'cranberry', 'cream cheese', 'crêpe', 'cucumber', 'cumin', 'currant', 'curry', 'dairy', 'dairy free', 'date', 'deep-fry', 'dessert', 'dinner', 'dip', 'diwali', 'dried fruit', 'drink', 'drinks', 'duck', 'easter', 'edible gift', 'egg', 'eggplant', 'endive', 'engagement party', 'fall', 'family reunion', "father's day", 'fennel', 'feta', 'fig', 'fish', 'flaming hot summer', 'fontina', 'food processor', 'fourth of july', 'freeze/chill', 'friendsgiving', 'frozen dessert', 'fruit', 'fruit juice', 'fry', 'game', 'garlic', 'gin', 'ginger', 'goat cheese', 'goose', 'graduation', 'green bean', 'grill', 'grill/barbecue', 'ground beef', 'ham', 'hanukkah', 'harpercollins', 'hazelnut', 'healthy', 'herb', 'high fiber', 'hominy/cornmeal/masa', 'honey', "hors d'oeuvre", 'hot pepper', 'house & garden', 'ice cream machine', 'jalapeño', 'jam or jelly', 'kentucky derby', 'kid-friendly', 'kidney friendly', 'kiwi', 'kosher', 'kosher for passover', 'labor day', 'lamb', 'leafy green', 'leek', 'legume', 'lemon', 'lemon juice', 'lemongrass', 'lentil', 'lime', 'lime juice', 'liqueur', 'lobster', 'low cal', 'low cholesterol', 'low fat', 'low sodium', 'low/no sugar', 'lunar new year', 'lunch', 'macadamia nut', 'mango', 'maple syrup', 'marinate', 'mayonnaise', 'meat', 'milk/cream', 'mint', 'mixer', 'molasses', 'mozzarella', 'mushroom', 'mussel', 'mustard', 'nectarine', "new year's day", "new year's eve", 'no sugar added', 'no-cook', 'non-alcoholic', 'noodle', 'nut', 'nutmeg', 'oat', 'olive', 'one-pot meal', 'onion', 'orange', 'orange juice', 'oregano', 'oscars', 'oyster', 'paleo', 'pan-fry', 'parade', 'parmesan', 'parsley', 'parsnip', 'party', 'passion fruit', 'passover', 'pasta', 'pastry', 'pea', 'pecan', 'pepper', 'pernod', 'pescatarian', 'phyllo/puff pastry dough', 'picnic', 'pie', 'pine nut', 'pineapple', 'pistachio', 'poach', 'poker/game night', 'poppy', 'pork', 'pork tenderloin', 'port', 'potato', 'potluck', 'poultry', 'pressure cooker', 'prune', 'purim', 'quick & easy', 'quinoa', 'radish', 'raisin', 'ramadan', 'raspberry', 'red wine', 'rhubarb', 'rice', 'ricotta', 'roast', 'rosemary', 'rosh hashanah/yom kippur', 'rum', 'saffron', 'salad', 'salad dressing', 'salmon', 'sandwich', 'sardine', 'sauce', 'sausage', 'sauté', 'scallop', 'seafood', 'seed', 'self', 'semolina', 'sesame', 'sesame oil', 'shallot', 'shellfish', 'shower', 'shrimp', 'side', 'simmer', 'slow cooker', 'smoker', 'soufflé/meringue', 'soup/stew', 'sour cream', 'soy', 'soy sauce', 'spice', 'spinach', 'spring', 'squash', 'steak', 'steam', 'stew', 'stir-fry', 'strawberry', 'sugar conscious', 'summer', 'super bowl', 'sweet potato/yam', 'swiss cheese', 'tailgating', 'tarragon', 'tea', 'thanksgiving', 'thyme', 'tomato', 'tortillas', 'tree nut', 'tree nut free', 'triple sec', 'trout', 'tuna', 'vanilla', 'veal', 'vegan', 'vegetable', 'vegetarian', 'venison', 'vinegar', 'vodka', 'wedding', 'weelicious', 'wheat/gluten-free', 'whiskey', 'white wine', 'wine', 'winter', 'wok', 'yellow squash', 'yogurt', 'zucchini', 'snack', 'turkey', 'bon appétit', 'peanut free', 'soy free']
{0.0: 6852, 1.0: 3776}
{0.0: 3329, 1.0: 1907}
Test accuracy of gourmet is 0.8905653170359052
Test precision of gourmet is 0.811391223155929

Test recall of gourmet is 0.9113791295228107
Test f1-score of gourmet is 0.8584835761916524

Category tree nut free
predictors ['almond', 'amaretto', 'anise', 'anniversary', 'appetizer', 'apple',
'apricot', 'artichoke', 'arugula', 'asparagus', 'back to school', 'backyard
bbq', 'bake', 'bean', 'beet', 'bell pepper', 'berry', 'blue cheese',
'blueberry', 'bok choy', 'brandy', 'bread', 'broil', 'brunch', 'brussel sprout',
'buffet', 'butter', 'cabbage', 'cake', 'carrot', 'cashew', 'chard', 'cheese',
'cherry', 'chestnut', 'chicken', 'chickpea', 'chile pepper', 'chill',
'chocolate', 'christmas', 'cinnamon', 'citrus', 'coconut', 'coffee', 'cookies',
'cornmeal', 'couscous', 'cranberry', 'cream cheese', 'cucumber', 'cumin',
'currant', 'curry', 'dairy', 'dairy free', 'date', 'dessert', 'easter', 'edible
gift', 'egg', 'eggplant', 'engagement party', 'escarole', 'fall', 'family
reunion', 'fennel', 'feta', 'food processor', 'frangelico', 'freeze/chill',
'frozen dessert', 'fruit', 'ginger', 'grill', 'hazelnut', 'healthy', 'herb',
'honey', 'hot pepper', 'jerusalem artichoke', 'kid-friendly', 'kidney friendly',
'kiwi', 'kosher', 'lamb', 'leafy green', 'lemon', 'lettuce', 'lime', 'liqueur',
'low cal', 'low cholesterol', 'low fat', 'low sodium', 'macadamia nut',
'milk/cream', 'mint', "mother's day", 'mushroom', 'nectarine', "new year's day",
'no sugar added', 'noodle', 'nut', 'nutmeg', 'oat', 'onion', 'orange', 'paleo',
'pancake', 'pasadena', 'passion fruit', 'pasta', 'pastry', 'peach', 'peanut',
'pear', 'pecan', 'pepper', 'pescatarian', 'picnic', 'pine nut', 'pistachio',
'poach', 'pomegranate', 'pumpkin', 'purim', 'quick & easy', 'radicchio',
'raisin', 'ramekin', 'raspberry', 'red wine', 'rice', 'ricotta', 'roast',
'salad', 'salad dressing', 'sauce', 'sausage', 'sauté', 'seafood', 'semolina',
'shrimp', 'side', 'smoothie', 'snapper', 'soy', 'soy sauce', 'spring', 'squash',
'sugar conscious', 'sugar snap pea', 'sukkot', 'summer', 'sweet potato/yam',
'tapioca', 'thanksgiving', 'tofu', 'tomato', "valentine's day", 'vanilla',
'vegan', 'vegetable', 'vegetarian', 'vinegar', 'walnut', 'watercress',
'wheat/gluten-free', 'winter', 'yogurt', 'zucchini', 'bon appétit', 'peanut
free', 'soy free', 'gourmet']
{0.0: 6878, 1.0: 3750}
{0.0: 3370, 1.0: 1866}
Test accuracy of tree nut free is 0.9723071046600459
Test precision of tree nut free is 0.939703628002044
Test recall of tree nut free is 0.9855305466237942
Test f1-score of tree nut free is 0.9620716714622024

Category vegetarian
predictors ['almond', 'anchovy', 'appetizer', 'apple', 'artichoke', 'asparagus',
'avocado', 'back to school', 'backyard bbq', 'bacon', 'bake', 'basil', 'bass',
'bean', 'beef', 'beer', 'beet', 'bell pepper', 'berry', 'blackberry', 'blender',
'blue cheese', 'blueberry', 'boil', 'bread', 'breakfast', 'broil', 'brunch',
'brussel sprout', 'buffet', 'butter', 'butternut squash', 'cake', 'canada',
'capers', 'caraway', 'carrot', 'cashew', 'cauliflower', 'celery', 'cheddar',

'cheese', 'chicken', 'chill', 'chive', 'christmas', 'christmas eve', 'cilantro', 'cinco de mayo', 'cinnamon', 'citrus', 'cocktail party', 'coconut', 'cod', 'cognac/armagnac', 'collard greens', 'colorado', 'condiment', 'condiment/spread', 'cookie', 'coriander', 'corn', 'cornmeal', 'cream cheese', 'cucumber', 'currant', 'curry', 'dairy', 'dairy free', 'dessert', 'dill', 'dinner', 'dried fruit', 'drink', 'duck', 'egg', 'eggplant', 'escarole', 'fall', 'family reunion', "father's day", 'fennel', 'feta', 'fig', 'fish', 'food processor', 'freeze/chill', 'frozen dessert', 'fruit', 'fry', 'garlic', 'ginger', 'goat cheese', 'graduation', 'granola', 'green bean', 'grill', 'grill/barbecue', 'ground beef', 'halibut', 'halloween', 'hamburger', 'hanukkah', 'harpercollins', 'hazelnut', 'healthy', 'herb', 'high fiber', 'hominy/cornmeal/masa', 'honey', 'honeydew', 'horseradish', 'hot pepper', 'jalapeño', 'jícama', 'kale', 'kentucky', 'kid-friendly', 'kidney friendly', 'kosher', 'kumquat', 'lamb', 'leafy green', 'leek', 'lemon', 'lemon juice', 'lentil', 'lettuce', 'lime', 'los angeles', 'low cal', 'low carb', 'low cholesterol', 'low fat', 'low sodium', 'low/no sugar', 'lunch', 'mango', 'marinate', 'mayonnaise', 'microwave', 'milk/cream', 'minnesota', 'mint', 'missouri', 'mixer', 'molasses', "mother's day", 'mozzarella', 'mushroom', 'mustard', 'no meat, no problem', 'no-cook', 'noodle', 'nut', 'olive', 'onion', 'orange', 'orange juice', 'oregon', 'pan-fry', 'papaya', 'parmesan', 'parsley', 'parsnip', 'party', 'pasta', 'pastry', 'pear', 'pecan', 'pepper', 'pescatarian', 'phyllo/puff pastry dough', 'picnic', 'pine nut', 'poker/game night', 'pomegranate', 'poppy', 'pork', 'potato', 'potluck', 'prosciutto', 'quick & easy', 'radish', 'raisin', 'ramekin', 'raspberry', 'raw', 'rice', 'roast', 'root vegetable', 'rosemary', 'saffron', 'sage', 'salad', 'salad dressing', 'salmon', 'sandwich', 'sauce', 'sausage', 'sauté', 'seafood', 'seed', 'self', 'sesame', 'shallot', 'shellfish', 'shrimp', 'side', 'simmer', 'slow cooker', 'soup/stew', 'sour cream', 'soy', 'soy sauce', 'spice', 'spinach', 'spring', 'squash', 'squid', "st. patrick's day", 'stir-fry', 'stuffing/dressing', 'sugar conscious', 'sugar snap pea', 'summer', 'super bowl', 'sweet potato/yam', 'swiss cheese', 'thanksgiving', 'thyme', 'tilapia', 'tofu', 'tomato', 'tortillas', 'trout', 'tuna', 'vanilla', 'veal', 'vegan', 'vegetable', 'vinegar', 'walnut', 'wasabi', 'watercress', 'weelicious', 'wheat/gluten-free', 'white wine', 'wine', 'winter', 'yellow squash', 'yogurt', 'zucchini', 'snack week', 'turkey', 'bon appétit', 'peanut free', 'gourmet']
{0.0: 6960, 1.0: 3668}
{0.0: 3375, 1.0: 1861}
Test accuracy of vegetarian is 0.9241787624140565
Test precision of vegetarian is 0.927071178529755
Test recall of vegetarian is 0.8538420204191295
Test f1-score of vegetarian is 0.8889510489510489

Category kosher
predictors ['advance prep required', 'anchovy', 'apricot', 'asparagus', 'bacon', 'bake', 'basil', 'bean', 'beef', 'beef rib', 'beef tenderloin', 'beer', 'beet', 'blender', 'braise', 'broccoli', 'broil', 'cabbage', 'cake', 'candy thermometer', 'carrot', 'celery', 'cheddar', 'chicken', 'chill', 'christmas',

'cilantro', 'cinnamon', 'citrus', 'clam', 'cocktail party', 'coriander', 'corn',
'cornmeal', 'crab', 'cucumber', 'currant', 'curry', 'dairy free', 'dinner',
'dried fruit', 'duck', 'egg', 'fall', 'fat free', "father's day", 'feta',
'fish', 'fourth of july', 'fruit', 'fruit juice', 'ginger', 'grill/barbecue',
'halibut', 'ham', 'hanukkah', 'healthy', 'high fiber', 'hot pepper', 'kale',
'kid-friendly', 'kidney friendly', 'leafy green', 'lemon', 'lemongrass',
'lettuce', 'lime', 'lobster', 'low cal', 'low carb', 'low fat', 'lunch',
'marinate', 'milk/cream', 'mint', 'mushroom', 'mussel', 'mustard', "new year's
eve", 'no sugar added', 'noodle', 'nut', 'octopus', 'olive', 'paleo', 'pan-fry',
'party', 'passover', 'pasta', 'peanut', 'pescatarian', 'picnic', 'pistachio',
'pomegranate', 'pork', 'pork chop', 'pork rib', 'pork tenderloin', 'potato',
'potluck', 'poultry', 'prosciutto', 'purim', 'quick & easy', 'raspberry', 'red
wine', 'rice', 'roast', 'root vegetable', 'rosemary', 'salad', 'salad dressing',
'salmon', 'sausage', 'sauté', 'scallop', 'seafood', 'shellfish', 'shrimp',
'simmer', 'soup/stew', 'sour cream', 'soy sauce', 'spice', 'spring', 'squid',
'stir-fry', 'strawberry', 'sugar conscious', 'sukkot', 'summer', 'sweet
potato/yam', 'tequila', 'thanksgiving', 'tomato', 'trout', 'tuna', 'veal',
'vegan', 'vegetable', 'vinegar', 'watermelon', 'wheat/gluten-free', 'white
wine', 'whole wheat', 'winter', 'turkey', 'bon appétit', 'peanut free', 'soy
free', 'gourmet', 'tree nut free', 'vegetarian']
{0.0: 7291, 1.0: 3337}
{0.0: 3588, 1.0: 1648}
Test accuracy of kosher is 0.9734530175706646
Test precision of kosher is 0.9637369391518131
Test recall of kosher is 0.9514563106796117
Test f1-score of kosher is 0.9575572519083969

Category pescatarian
predictors ['appetizer', 'backyard bbq', 'bacon', 'bake', 'basil', 'bean',
'beer', 'bell pepper', 'blender', 'breadcrumbs', 'breakfast', 'broccoli rabe',
'broil', 'cabbage', 'cheddar', 'chicken', 'chile pepper', 'chill', 'chive',
'christmas', 'cilantro', 'cocktail party', 'coconut', 'cod', 'coriander',
'couscous', 'crab', 'cucumber', 'curry', 'dairy free', 'dinner', 'egg', 'fall',
'fish', 'fourth of july', 'garlic', 'green bean', 'halibut', 'halloween', 'ham',
'hanukkah', 'healthy', 'herb', 'high fiber', 'horseradish', 'kid-friendly',
'kidney friendly', 'lemon', 'lentil', 'lettuce', 'lime juice', 'lobster', 'low
cal', 'lunch', 'massachusetts', 'meat', 'mixer', 'molasses', 'mushroom',
'mussel', 'mustard', "new year's eve", 'no sugar added', 'no-cook', 'noodle',
'onion', 'oyster', 'passover', 'pea', 'peanut', 'pecan', 'pepper', 'pine nut',
'poach', 'pork', 'port', 'potato', 'potluck', 'quick & easy', 'rice', 'ricotta',
'rosh hashanah/yom kippur', 'sage', 'salad', 'salmon', 'sausage', 'scallop',
'seafood', 'shallot', 'shellfish', 'shower', 'shrimp', 'side', 'simmer',
'soup/stew', 'south carolina', 'spice', 'squid', 'steam', 'stir-fry',
'stuffing/dressing', 'sugar conscious', 'summer', 'tequila', 'tofu', 'tomato',
'trout', 'tuna', "valentine's day", 'vegan', 'vegetable', 'vinegar', 'vodka',
'wheat/gluten-free', 'white wine', 'bon appétit', 'peanut free', 'soy free',
'gourmet', 'tree nut free', 'vegetarian', 'kosher']

```
{0.0: 7344, 1.0: 3284}
{0.0: 3644, 1.0: 1592}
Test accuracy of pescatarian is 0.9812834224598931
Test precision of pescatarian is 0.966875
Test recall of pescatarian is 0.9717336683417085
Test f1-score of pescatarian is 0.9692982456140351
```

Category quick & easy
predictors ['22-minute meals', '3-ingredient recipes', 'almond', 'anchovy',
'anise', 'appetizer', 'apple', 'apricot', 'arizona', 'artichoke', 'arugula',
'asparagus', 'avocado', 'back to school', 'backyard bbq', 'bacon', 'bake',
'banana', 'barley', 'basil', 'bass', 'bastille day', 'bean', 'beef', 'beef
tenderloin', 'beer', 'beet', 'bell pepper', 'berry', 'birthday', 'blackberry',
'blender', 'blue cheese', 'blueberry', 'boil', 'bok choy', 'braise', 'brandy',
'bread', 'breadcrumbs', 'breakfast', 'brine', 'brisket', 'broccoli', 'broccoli
rabe', 'broil', 'brooklyn', 'brown rice', 'brunch', 'brussel sprout', 'buffet',
'bulgur', 'butter', 'cabbage', 'cake', 'california', 'canada', 'candy',
'capers', 'carrot', 'cashew', 'casserole/gratin', 'cauliflower', 'celery',
'chard', 'cheddar', 'cheese', 'cherry', 'chicken', 'chickpea', 'chile pepper',
'chill', 'chive', 'chocolate', 'christmas', 'christmas eve', 'cilantro', 'cinco
de mayo', 'cinnamon', 'citrus', 'clam', 'clove', 'cocktail', 'cocktail party',
'coconut', 'cod', 'coffee', 'coffee grinder', 'cognac/armagnac', 'condiment',
'condiment/spread', 'cookies', 'coriander', 'corn', 'cornmeal', 'cottage
cheese', 'couscous', 'crab', 'cranberry', 'cream cheese', 'cucumber', 'cumin',
'currant', 'curry', 'dairy', 'dairy free', 'date', 'deep-fry', 'dessert',
'dill', 'dinner', 'dip', 'dried fruit', 'drink', 'duck', 'easter', 'edible
gift', 'egg', 'eggplant', 'endive', 'engagement party', 'escarole', 'fall',
'family reunion', 'fat free', "father's day", 'fennel', 'feta', 'fish',
'florida', 'fontina', 'food processor', 'fortified wine', 'fourth of july',
'france', 'freeze/chill', 'frozen dessert', 'fruit', 'fruit juice', 'fry',
'game', 'garlic', 'ginger', 'goat cheese', 'goose', 'graduation', 'grape',
'grapefruit', 'green bean', 'green onion/scallion', 'grill', 'grill/barbecue',
'ground beef', 'guam', 'halibut', 'ham', 'hanukkah', 'healthy', 'herb', 'high
fiber', 'hominy/cornmeal/masa', 'honey', 'horseradish', 'hot drink', 'hot
pepper', 'house & garden', 'ice cream machine', 'illinois', 'indiana',
'ireland', 'italy', 'jalapeño', 'jam or jelly', 'kale', 'kentucky derby', 'kid-
friendly', 'kidney friendly', 'kiwi', 'kosher for passover', 'kwanzaa', 'lamb',
'lamb chop', 'leafy green', 'leek', 'legume', 'lemon', 'lettuce', 'lima bean',
'lime', 'liqueur', 'lobster', 'los angeles', 'low cal', 'low carb', 'low
cholesterol', 'low fat', 'low sodium', 'low/no sugar', 'lunar new year',
'lunch', 'macadamia nut', 'mango', 'maple syrup', 'mardi gras', 'marinate',
'maryland', 'massachusetts', 'mayonnaise', 'meat', 'melon', 'microwave',
'milk/cream', 'mint', 'missouri', 'mixer', 'molasses', 'mozzarella', 'mushroom',
'mussel', 'mustard', 'mustard greens', 'new mexico', "new year's day", 'new
york', 'no sugar added', 'no-cook', 'noodle', 'nut', 'nutmeg', 'oat', 'octopus',
'ohio', 'oktoberfest', 'olive', 'onion', 'orange', 'orange juice', 'oregano',
'oregon', 'organic', 'orzo', 'oscars', 'oyster', 'paleo', 'pan-fry', 'pancake',

31

'papaya', 'paris', 'parmesan', 'parsley', 'party', 'passion fruit', 'passover',
'pasta', 'pasta maker', 'pastry', 'pea', 'peach', 'peanut', 'peanut butter',
'pear', 'pecan', 'pennsylvania', 'pepper', 'pernod', 'phyllo/puff pastry dough',
'pickles', 'picnic', 'pine nut', 'pineapple', 'pistachio', 'plum', 'poach',
'poker/game night', 'pomegranate', 'pork', 'pork chop', 'pork rib', 'pork
tenderloin', 'port', 'potato', 'potluck', 'poultry', 'pressure cooker',
'prosciutto', 'prune', 'pumpkin', 'purim', 'quick and healthy', 'radish',
'raisin', 'raspberry', 'raw', 'red wine', 'rhubarb', 'rice', 'ricotta', 'roast',
'root vegetable', 'rosemary', 'rosh hashanah/yom kippur', 'rum', 'rutabaga',
'saffron', 'sage', 'salad', 'salad dressing', 'salmon', 'sandwich', 'sauce',
'sausage', 'sauté', 'scallop', 'seafood', 'seed', 'self', 'sesame', 'shallot',
'shavuot', 'shellfish', 'sherry', 'shower', 'shrimp', 'side', 'simmer',
'snapper', 'soup/stew', 'sour cream', 'south carolina', 'soy', 'soy sauce',
'sparkling wine', 'spice', 'spinach', 'spring', 'squash', 'squid', 'st. louis',
'steak', 'steam', 'stew', 'stir-fry', 'strawberry', 'sugar conscious', 'sugar
snap pea', 'summer', 'super bowl', 'sweet potato/yam', 'swiss cheese',
'tailgating', 'tamarind', 'tangerine', 'tarragon', 'tea', 'texas',
'thanksgiving', 'thyme', 'tofu', 'tomatillo', 'tomato', 'tortillas', 'triple
sec', 'tuna', 'turnip', "valentine's day", 'vanilla', 'veal', 'vegan',
'vegetable', 'venison', 'vinegar', 'virginia', 'vodka', 'walnut', 'wasabi',
'washington', 'watercress', 'watermelon', 'weelicious', 'wheat/gluten-free',
'white wine', 'wild rice', 'wine', 'winter', 'wok', 'yogurt', 'yonkers',
'zucchini', 'snack week', 'turkey', 'bon appétit', 'peanut free', 'soy free',
'gourmet', 'tree nut free', 'vegetarian', 'kosher', 'pescatarian']
{0.0: 7808, 1.0: 2820}
{0.0: 3845, 1.0: 1391}
Test accuracy of quick & easy is 0.7694805194805194
Test precision of quick & easy is 0.6045454545454545
Test recall of quick & easy is 0.38245866283249463
Test f1-score of quick & easy is 0.46851607221488334

Category wheat/gluten-free
predictors ['#cakeweek', '22-minute meals', '3-ingredient recipes', 'advance
prep required', 'almond', 'amaretto', 'anise', 'anniversary', 'appetizer',
'apricot', 'artichoke', 'asparagus', 'avocado', 'backyard bbq', 'bacon', 'bake',
'banana', 'barley', 'basil', 'bastille day', 'bean', 'beef', 'beef rib', 'beef
tenderloin', 'beer', 'beet', 'bell pepper', 'berry', 'birthday', 'biscuit',
'blender', 'blue cheese', 'blueberry', 'braise', 'bread', 'breakfast',
'brisket', 'broccoli', 'broccoli rabe', 'broil', 'brunch', 'bulgur', 'butter',
'butternut squash', 'cabbage', 'cake', 'candy', 'capers', 'carrot',
'casserole/gratin', 'cauliflower', 'chard', 'cheddar', 'cheese', 'cherry',
'chestnut', 'chicken', 'chickpea', 'chile pepper', 'chill', 'chive',
'chocolate', 'christmas', 'cilantro', 'cinnamon', 'citrus', 'cocktail party',
'coconut', 'coffee', 'cognac/armagnac', 'collard greens', 'condiment/spread',
'cookies', 'corn', 'couscous', 'cranberry', 'cream cheese', 'cucumber', 'curry',
'dairy', 'dairy free', 'date', 'deep-fry', 'dessert', 'dinner', 'double boiler',
'drink', 'duck', 'egg', 'eggplant', 'endive', 'engagement party', 'fall',

'family reunion', 'fat free', 'fennel', 'feta', 'fig', 'fish', 'food processor',
'fourth of july', 'freezer food', 'frozen dessert', 'fruit', 'fruit juice',
'fry', 'garlic', 'georgia', 'gin', 'ginger', 'goat cheese', 'goose',
'grapefruit', 'green onion/scallion', 'grill', 'grill/barbecue', 'halibut',
'hanukkah', 'harpercollins', 'hazelnut', 'healthy', 'herb', 'high fiber',
'hominy/cornmeal/masa', 'honey', 'honeydew', 'hot pepper', 'house & garden',
'ice cream', 'ice cream machine', 'jalapeño', 'jam or jelly', 'jerusalem
artichoke', 'kid-friendly', 'kidney friendly', 'kumquat', 'lamb', 'lamb chop',
'lasagna', 'leafy green', 'leek', 'legume', 'lemon', 'lime', 'lime juice',
'liqueur', 'low cal', 'low carb', 'low cholesterol', 'low fat', 'low sodium',
'low/no sugar', 'lunch', 'macadamia nut', 'maple syrup', 'marinate', 'maryland',
'massachusetts', 'mayonnaise', 'meat', 'michigan', 'microwave', 'milk/cream',
'mint', 'mixer', 'molasses', 'mushroom', 'mussel', 'mustard', 'mustard greens',
"new year's eve", 'no meat, no problem', 'no sugar added', 'no-cook', 'noodle',
'nut', 'oat', 'oktoberfest', 'olive', 'onion', 'orange', 'oregano', 'oregon',
'oscars', 'paleo', 'pan-fry', 'paprika', 'parmesan', 'parsley', 'parsnip',
'party', 'passover', 'pasta', 'pastry', 'pea', 'peach', 'peanut', 'pear',
'pecan', 'pepper', 'pickles', 'pistachio', 'plum', 'poach', 'pomegranate',
'pomegranate juice', 'pork', 'pork chop', 'pork tenderloin', 'port', 'potato',
'potluck', 'poultry', 'prune', 'quick and healthy', 'quinoa', 'rabbit',
'raisin', 'ramekin', 'raspberry', 'raw', 'red wine', 'rice', 'ricotta', 'roast',
'root vegetable', 'rosemary', 'rosh hashanah/yom kippur', 'rum', 'saffron',
'salad', 'salad dressing', 'salmon', 'sandwich', 'sauce', 'sausage', 'sauté',
'scallop', 'seafood', 'seed', 'semolina', 'sesame', 'shallot', 'shavuot',
'shrimp', 'side', 'simmer', 'smoker', 'smoothie', 'snapper', 'soup/stew', 'sour
cream', 'south carolina', 'soy sauce', 'spice', 'spinach', 'spring', 'squash',
"st. patrick's day", 'steak', 'steam', 'stew', 'stir-fry', 'strawberry',
'stuffing/dressing', 'sugar conscious', 'summer', 'super bowl', 'sweet
potato/yam', 'taco', 'tea', 'thanksgiving', 'thyme', 'tofu', 'tomato',
'tortillas', 'trout', 'tuna', 'turnip', "valentine's day", 'vanilla', 'vegan',
'vegetable', 'vinegar', 'walnut', 'washington', 'watercress', 'weelicious',
'whiskey', 'white wine', 'winter', 'yellow squash', 'yogurt', 'zucchini',
'snack', 'turkey', 'bon appétit', 'peanut free', 'soy free', 'gourmet', 'tree
nut free', 'vegetarian', 'kosher', 'pescatarian', 'quick & easy']
{0.0: 7985, 1.0: 2643}
{0.0: 3978, 1.0: 1258}
Test accuracy of wheat/gluten-free is 0.922077922077922
Test precision of wheat/gluten-free is 0.8489326765188834
Test recall of wheat/gluten-free is 0.821939586645469
Test f1-score of wheat/gluten-free is 0.8352180936995154

Category bake
predictors ['3-ingredient recipes', 'advance prep required', 'alcoholic',
'almond', 'amaretto', 'anchovy', 'anise', 'appetizer', 'apple', 'apricot',
'arizona', 'artichoke', 'arugula', 'asparagus', 'back to school', 'bacon',
'banana', 'barley', 'basil', 'bass', 'bastille day', 'bean', 'beef', 'beef rib',
'beer', 'beet', 'berry', 'birthday', 'biscuit', 'blackberry', 'blender', 'blue

cheese', 'blueberry', 'bourbon', 'braise', 'brandy', 'bread', 'breadcrumbs',
'breakfast', 'brisket', 'broil', 'brownie', 'brunch', 'buffet', 'butter',
'buttermilk', 'butternut squash', 'cabbage', 'cake', 'california', 'candy',
'candy thermometer', 'carrot', 'cashew', 'casserole/gratin', 'celery',
'cheddar', 'cheese', 'cherry', 'chestnut', 'chicken', 'chill', 'chive',
'chocolate', 'christmas', 'christmas eve', 'cilantro', 'cinco de mayo',
'cinnamon', 'citrus', 'clove', 'cocktail party', 'coconut', 'cod', 'coffee',
'colorado', 'condiment', 'condiment/spread', 'cookie', 'cookies', 'coriander',
'corn', 'cornmeal', 'couscous', 'crab', 'cranberry', 'cream cheese', 'cucumber',
'currant', 'curry', 'dairy', 'dairy free', 'date', 'deep-fry', 'dessert',
'dill', 'dinner', 'dried fruit', 'drink', 'duck', 'easter', 'edible gift',
'egg', 'eggplant', 'engagement party', 'fall', 'family reunion', "father's day",
'fennel', 'feta', 'fig', 'fish', 'fontina', 'food processor', 'fourth of july',
'france', 'freeze/chill', 'friendsgiving', 'frozen dessert', 'fruit', 'fruit
juice', 'fry', 'garlic', 'georgia', 'ginger', 'goat cheese', 'gouda', 'grappa',
'green onion/scallion', 'grill/barbecue', 'ground beef', 'halibut', 'ham',
'hanukkah', 'harpercollins', 'hazelnut', 'healthy', 'herb', 'high fiber',
'hominy/cornmeal/masa', 'honey', 'horseradish', 'hot pepper', 'house & garden',
'ice cream machine', 'illinois', 'jalapeño', 'jam or jelly', 'jerusalem
artichoke', 'kahlúa', 'kale', 'kentucky derby', 'kid-friendly', 'kidney
friendly', 'kiwi', 'kosher for passover', 'lamb', 'leafy green', 'leek',
'legume', 'lemon', 'lentil', 'lettuce', 'lime', 'lime juice', 'liqueur', 'low
cal', 'low carb', 'low cholesterol', 'low fat', 'low sodium', 'low/no sugar',
'lunch', 'macaroni and cheese', 'maple syrup', 'mardi gras', 'marinate',
'mayonnaise', 'meat', 'michigan', 'microwave', 'milk/cream', 'mint', 'mixer',
'molasses', 'monterey jack', "mother's day", 'mozzarella', 'mushroom',
'mustard', 'mustard greens', 'nebraska', 'nectarine', 'new jersey', 'new
mexico', "new year's eve", 'new york', 'no sugar added', 'no-cook', 'nut',
'nutmeg', 'oat', 'oktoberfest', 'olive', 'onion', 'orange', 'oscars', 'oyster',
'pan-fry', 'parmesan', 'parsley', 'parsnip', 'party', 'pasadena', 'passover',
'pasta', 'pastry', 'pea', 'peach', 'peanut butter', 'pear', 'pecan',
'pennsylvania', 'pepper', 'pernod', 'phyllo/puff pastry dough', 'picnic', 'pie',
'pine nut', 'pineapple', 'pistachio', 'pizza', 'plum', 'poach', 'poker/game
night', 'pomegranate', 'pork', 'pork rib', 'pork tenderloin', 'potato',
'potluck', 'poultry', 'poultry sausage', 'pressure cooker', 'prosciutto',
'prune', 'pumpkin', 'quince', 'quinoa', 'radish', 'raisin', 'ramadan',
'raspberry', 'red wine', 'rhubarb', 'rice', 'ricotta', 'roast', 'root
vegetable', 'rosemary', 'rosh hashanah/yom kippur', 'rum', 'sage', 'salad',
'salmon', 'sandwich', 'sauce', 'sausage', 'sauté', 'scallop', 'seafood',
'semolina', 'shallot', 'shellfish', 'sherry', 'shower', 'shrimp', 'side',
'simmer', 'soup/stew', 'sour cream', 'soy sauce', 'spice', 'spinach', 'spring',
'squash', 'squid', "st. patrick's day", 'steak', 'steam', 'stew', 'stir-fry',
'strawberry', 'stuffing/dressing', 'sugar conscious', 'summer', 'super bowl',
'sweet potato/yam', 'swiss cheese', 'tapioca', 'tarragon', 'tea',
'thanksgiving', 'thyme', 'tomato', 'tortillas', 'tree nut', 'triple sec',
'trout', "valentine's day", 'vanilla', 'veal', 'vegan', 'vegetable', 'vinegar',
'virginia', 'walnut', 'white wine', 'winter', 'yogurt', 'turkey', 'bon appétit',
'peanut free', 'soy free', 'gourmet', 'tree nut free', 'vegetarian', 'kosher',

'pescatarian', 'quick & easy', 'wheat/gluten-free']
{0.0: 8171, 1.0: 2457}
{0.0: 3968, 1.0: 1268}
Test accuracy of bake is 0.8919022154316272
Test precision of bake is 0.8150807899461401
Test recall of bake is 0.7160883280757098
Test f1-score of bake is 0.762384550797649

Category summer
predictors ['#cakeweek', '22-minute meals', 'advance prep required',
'alcoholic', 'almond', 'amaretto', 'appetizer', 'apple', 'apricot', 'arizona',
'artichoke', 'arugula', 'asparagus', 'avocado', 'back to school', 'backyard
bbq', 'bacon', 'banana', 'basil', 'bass', 'bean', 'beef', 'beef rib', 'beef
tenderloin', 'beer', 'beet', 'bell pepper', 'berry', 'birthday', 'bitters',
'blackberry', 'blender', 'blueberry', 'boil', 'bok choy', 'brandy', 'bread',
'breakfast', 'brine', 'broccoli', 'broccoli rabe', 'broil', 'brunch', 'buffet',
'bulgur', 'butter', 'buttermilk', 'cabbage', 'cake', 'california', 'campari',
'cantaloupe', 'capers', 'carrot', 'cauliflower', 'celery', 'champagne', 'chard',
'cheddar', 'cheese', 'cherry', 'chestnut', 'chicken', 'chickpea', 'chile
pepper', 'chill', 'chive', 'chocolate', 'christmas', 'cilantro', 'cinco de
mayo', 'cinnamon', 'citrus', 'clam', 'cocktail', 'cocktail party', 'coconut',
'coffee', 'cognac/armagnac', 'condiment/spread', 'cookies', 'corn', 'cornmeal',
'cottage cheese', 'couscous', 'crab', 'cranberry', 'cream cheese', 'cucumber',
'cumin', 'currant', 'curry', 'dairy', 'dairy free', 'date', 'dessert',
'digestif', 'dill', 'dinner', 'dried fruit', 'drink', 'duck', 'easter', 'egg',
'eggplant', 'endive', 'engagement party', 'fall', 'family reunion', 'fat free',
"father's day", 'fennel', 'feta', 'fig', 'fish', 'food processor', 'fortified
wine', 'fourth of july', 'freeze/chill', 'frozen dessert', 'fruit', 'fruit
juice', 'fry', 'garlic', 'georgia', 'gin', 'ginger', 'goat cheese',
'graduation', 'grape', 'grapefruit', 'green bean', 'green onion/scallion',
'grill', 'grill/barbecue', 'guava', 'halibut', 'ham', 'hanukkah', 'hazelnut',
'healthy', 'herb', 'high fiber', 'honey', "hors d'oeuvre", 'horseradish', 'hot
pepper', 'house & garden', 'ice cream', 'ice cream machine', 'iced tea',
'jalapeño', 'jícama', 'kansas', 'kentucky derby', 'kid-friendly', 'kidney
friendly', 'kitchen olympics', 'kumquat', 'kwanzaa', 'lamb', 'leafy green',
'leek', 'lemon', 'lemon juice', 'lemongrass', 'lentil', 'lettuce', 'lime', 'lime
juice', 'liqueur', 'lobster', 'low cal', 'low carb', 'low cholesterol', 'low
fat', 'low sodium', 'low/no sugar', 'lunch', 'macadamia nut', 'mango', 'mardi
gras', 'margarita', 'marinate', 'marsala', 'maryland', 'mayonnaise', 'meat',
'melon', 'milk/cream', 'mint', 'missouri', 'mixer', 'molasses', 'monterey jack',
"mother's day", 'muffin', 'mushroom', 'mussel', 'mustard', 'nectarine', 'new
mexico', "new year's eve", 'no sugar added', 'no-cook', 'non-alcoholic',
'noodle', 'nut', 'octopus', 'oktoberfest', 'olive', 'onion', 'orange',
'oregano', 'oyster', 'paleo', 'pan-fry', 'papaya', 'parmesan', 'parsley',
'party', 'pasadena', 'passion fruit', 'passover', 'pasta', 'pasta maker', 'pea',
'peach', 'peanut', 'pear', 'pecan', 'pepper', 'pernod', 'phyllo/puff pastry
dough', 'pickles', 'picnic', 'pine nut', 'pineapple', 'pistachio', 'plum',

'poker/game night', 'pomegranate', 'poppy', 'pork', 'pork chop', 'pork rib',
'port', 'potato', 'potluck', 'poultry', 'quick and healthy', 'radish', 'raisin',
'raspberry', 'raw', 'red wine', 'rhubarb', 'rice', 'ricotta', 'roast', 'root
vegetable', 'rosemary', 'rum', 'rye', 'sake', 'salad', 'salad dressing',
'salmon', 'san francisco', 'sandwich', 'sauce', 'sausage', 'sauté', 'scallop',
'seafood', 'self', 'sesame', 'shellfish', 'sherry', 'shower', 'shrimp', 'side',
'simmer', 'smoothie', 'snapper', 'soup/stew', 'sour cream', 'south carolina',
'soy', 'soy sauce', 'sparkling wine', 'spice', 'spinach', 'spirit', 'spring',
'squash', 'steak', 'steam', 'stew', 'stir-fry', 'strawberry', 'sugar conscious',
'sugar snap pea', 'super bowl', 'swiss cheese', 'tailgating', 'tamarind',
'tapioca', 'tarragon', 'tequila', 'thanksgiving', 'thyme', 'tilapia', 'tofu',
'tomato', 'tortillas', 'tree nut', 'triple sec', 'tropical fruit', 'tuna',
'turnip', "valentine's day", 'vanilla', 'veal', 'vegan', 'vegetable', 'vermont',
'vinegar', 'walnut', 'washington', 'watercress', 'watermelon', 'wedding',
'weelicious', 'white wine', 'wine', 'winter', 'yogurt', 'zucchini', 'snack',
'turkey', 'bon appétit', 'peanut free', 'soy free', 'gourmet', 'tree nut free',
'vegetarian', 'kosher', 'pescatarian', 'quick & easy', 'wheat/gluten-free',
'bake']
{0.0: 8437, 1.0: 2191}
{0.0: 4177, 1.0: 1059}
Test accuracy of summer is 0.8474025974025974
Test precision of summer is 0.6825842696629213
Test recall of summer is 0.45892351274787535
Test f1-score of summer is 0.5488424618859401

Category dessert
predictors ['3-ingredient recipes', 'advance prep required', 'alcoholic',
'almond', 'anise', 'appetizer', 'apple', 'apricot', 'avocado', 'backyard bbq',
'banana', 'barley', 'bastille day', 'beef', 'berry', 'blackberry', 'blender',
'blue cheese', 'blueberry', 'boil', 'bourbon', 'brandy', 'bread', 'breakfast',
'broil', 'brownie', 'brunch', 'cake', 'candy', 'cantaloupe', 'cashew', 'cheese',
'cherry', 'chicken', 'chill', 'chocolate', 'christmas', 'cinnamon', 'citrus',
'cobbler/crumble', 'cocktail party', 'coconut', 'coffee', 'cognac/armagnac',
'condiment/spread', 'cookies', 'cornmeal', 'cream cheese', 'dairy', 'dairy
free', 'date', 'dinner', 'double boiler', 'dried fruit', 'drink', 'duck',
'edible gift', 'egg', 'fall', 'family reunion', 'fat free', 'fennel', 'fig',
'fish', 'food processor', 'freeze/chill', 'friendsgiving', 'frozen dessert',
'fruit', 'garlic', 'georgia', 'ginger', 'grill', 'grill/barbecue', 'ham',
'hazelnut', 'healthy', 'honey', 'hot pepper', 'ice cream machine', 'jalapeño',
'jam or jelly', 'kid-friendly', 'kidney friendly', 'lamb', 'leafy green',
'leek', 'lemon', 'lime', 'liqueur', 'lobster', 'low carb', 'low cholesterol',
'low fat', 'low sodium', 'lunch', 'maple syrup', 'marsala', 'marshmallow',
'meat', 'milk/cream', 'mixer', 'molasses', 'mushroom', 'mustard', 'nebraska',
'new mexico', 'no sugar added', 'no-cook', 'non-alcoholic', 'nut', 'oat',
'oatmeal', 'ohio', 'onion', 'orange', 'organic', 'oscars', 'party', 'pasta',
'pastry', 'peach', 'pear', 'pecan', 'pennsylvania', 'pepper', 'phyllo/puff
pastry dough', 'picnic', 'pie', 'pineapple', 'pistachio', 'poach', 'pork', 'pork

rib', 'prune', 'pumpkin', 'ramadan', 'raspberry', 'red wine', 'rhubarb', 'rice',
'ricotta', 'roast', 'root vegetable', 'rosemary', 'rum', 'rye', 'salad',
'salmon', 'sauce', 'sauté', 'seafood', 'sherry', 'shrimp', 'side', 'simmer',
'soup/stew', 'sour cream', 'soy sauce', 'spice', 'spring', 'squash', "st.
patrick's day", 'steam', 'strawberry', 'sugar conscious', 'sweet potato/yam',
'tapioca', 'thanksgiving', 'thyme', 'tomato', 'tree nut', 'tropical fruit',
"valentine's day", 'vanilla', 'vegan', 'vegetable', 'vodka', 'walnut',
'wedding', 'white wine', 'winter', 'yogurt', 'snack week', 'turkey', 'bon
appétit', 'peanut free', 'soy free', 'gourmet', 'tree nut free', 'vegetarian',
'kosher', 'pescatarian', 'quick & easy', 'wheat/gluten-free', 'bake', 'summer']
{0.0: 8669, 1.0: 1959}
{0.0: 4243, 1.0: 993}
Test accuracy of dessert is 0.974025974025974
Test precision of dessert is 0.9544008483563097
Test recall of dessert is 0.9063444108761329
Test f1-score of dessert is 0.9297520661157024

Category winter
predictors ['advance prep required', 'alcoholic', 'almond', 'anise',
'anniversary', 'appetizer', 'apple', 'apricot', 'arugula', 'asian pear',
'backyard bbq', 'bacon', 'barley', 'basil', 'bass', 'bastille day', 'bean',
'beef', 'beef rib', 'beef tenderloin', 'beet', 'bell pepper', 'berry',
'birthday', 'bitters', 'blender', 'blue cheese', 'blueberry', 'boil', 'bok
choy', 'bourbon', 'braise', 'brandy', 'bread', 'breadcrumbs', 'breakfast',
'broccoli', 'broil', 'brunch', 'buffet', 'butternut squash', 'cabbage', 'cake',
'california', 'capers', 'cardamom', 'carrot', 'casserole/gratin', 'cauliflower',
'celery', 'champagne', 'chard', 'cheddar', 'cheese', 'cherry', 'chestnut',
'chicken', 'chickpea', 'chile pepper', 'chill', 'chive', 'chocolate',
'christmas', 'christmas eve', 'cilantro', 'cinco de mayo', 'citrus', 'clove',
'cobbler/crumble', 'cocktail', 'cocktail party', 'coconut', 'cod', 'coffee',
'coffee grinder', 'cognac/armagnac', 'collard greens', 'condiment',
'condiment/spread', 'cookies', 'coriander', 'corn', 'cornmeal', 'crab',
'cranberry', 'cream cheese', 'cucumber', 'currant', 'curry', 'dairy', 'dairy
free', 'digestif', 'dill', 'dinner', 'double boiler', 'dried fruit', 'drink',
'duck', 'easter', 'eau de vie', 'edible gift', 'egg', 'eggplant', 'endive',
'engagement party', 'escarole', 'fall', 'family reunion', 'fat free', "father's
day", 'fennel', 'feta', 'fish', 'florida', 'fontina', 'food processor', 'fourth
of july', 'frangelico', 'freeze/chill', 'frozen dessert', 'fruit', 'fruit
juice', 'fry', 'game', 'garlic', 'ginger', 'goat cheese', 'goose', 'graduation',
'grape', 'grapefruit', 'grill/barbecue', 'ground beef', 'ground lamb',
'halibut', 'halloween', 'ham', 'hanukkah', 'harpercollins', 'hazelnut',
'healthy', 'herb', 'high fiber', 'hominy/cornmeal/masa', 'honey', 'horseradish',
'hot drink', 'hot pepper', 'house & garden', 'ice cream machine', 'jalapeño',
'jam or jelly', 'jerusalem artichoke', 'kahlúa', 'kale', 'kid-friendly', 'kidney
friendly', 'kosher for passover', 'kumquat', 'lamb', 'lamb shank', 'leafy
green', 'leek', 'lemon', 'lentil', 'lima bean', 'lime', 'liqueur', 'long beach',
'louisiana', 'low cal', 'low carb', 'low fat', 'low sodium', 'low/no sugar',

'lunar new year', 'lunch', 'maine', 'mango', 'maple syrup', 'mardi gras',
'marinate', 'marsala', 'massachusetts', 'mayonnaise', 'meat', 'milk/cream',
'mint', 'missouri', 'mixer', 'molasses', 'monterey jack', "mother's day",
'mushroom', 'mussel', 'mustard', 'mustard greens', 'nectarine', "new year's
eve", 'no sugar added', 'no-cook', 'non-alcoholic', 'noodle', 'north carolina',
'nut', 'nutmeg', 'oat', 'oktoberfest', 'olive', 'onion', 'orange', 'oregano',
'oscars', 'oyster', 'paleo', 'pan-fry', 'parmesan', 'parsley', 'parsnip',
'party', 'passover', 'pasta', 'pastry', 'pea', 'peach', 'peanut', 'pear',
'pecan', 'pennsylvania', 'pepper', 'pernod', 'picnic', 'pine nut', 'pistachio',
'poker/game night', 'pomegranate', 'poppy', 'pork', 'pork chop', 'pork
tenderloin', 'port', 'potato', 'potluck', 'poultry', 'prune', 'pumpkin',
'punch', 'purim', 'quail', 'quince', 'rack of lamb', 'radish', 'raisin',
'ramekin', 'raspberry', 'red wine', 'rice', 'roast', 'root vegetable',
'rosemary', 'rum', 'rye', 'saffron', 'sage', 'sake', 'salad', 'salad dressing',
'salmon', 'sandwich', 'santa monica', 'sauce', 'sausage', 'sauté', 'scallop',
'seafood', 'seed', 'sesame', 'shallot', 'shellfish', 'sherry', 'shower',
'shrimp', 'side', 'simmer', 'snapper', 'soup/stew', 'sour cream', 'soy sauce',
'spice', 'spinach', 'spring', 'squash', "st. patrick's day", 'steak', 'stew',
'stir-fry', 'strawberry', 'sugar conscious', 'super bowl', 'sweet potato/yam',
'swiss cheese', 'tangerine', 'tapioca', 'tarragon', 'tea', 'thanksgiving',
'thyme', 'tofu', 'tomatillo', 'tomato', 'tortillas', 'tree nut', 'tropical
fruit', 'trout', 'tuna', 'turnip', "valentine's day", 'vanilla', 'veal',
'vegan', 'vegetable', 'vinegar', 'vodka', 'walnut', 'washington', 'watercress',
'wedding', 'white wine', 'wine', 'wok', 'yogurt', 'zucchini', 'turkey', 'bon
appétit', 'peanut free', 'soy free', 'gourmet', 'tree nut free', 'vegetarian',
'kosher', 'pescatarian', 'quick & easy', 'wheat/gluten-free', 'bake', 'summer',
'dessert']
{0.0: 8889, 1.0: 1739}
{0.0: 4394, 1.0: 842}
Test accuracy of winter is 0.855423987776929
Test precision of winter is 0.606516290726817
Test recall of winter is 0.28741092636579574
Test f1-score of winter is 0.39000805801772764

Category side
predictors ['alcoholic', 'almond', 'appetizer', 'apple', 'artichoke',
'asparagus', 'back to school', 'backyard bbq', 'bacon', 'barley', 'basil',
'bean', 'beef', 'beet', 'bell pepper', 'blender', 'boil', 'bok choy', 'bread',
'breakfast', 'broccoli rabe', 'broil', 'brunch', 'brussel sprout', 'buffet',
'bulgur', 'butter', 'buttermilk', 'butternut squash', 'cabbage', 'california',
'carrot', 'casserole/gratin', 'cauliflower', 'cheddar', 'cheese', 'chicken',
'chile pepper', 'chill', 'chive', 'christmas', 'christmas eve', 'cilantro',
'cinco de mayo', 'cinnamon', 'citrus', 'cocktail party', 'coconut', 'condiment',
'condiment/spread', 'cook like a diner', 'corn', 'cornmeal', 'couscous',
'cucumber', 'currant', 'curry', 'dairy', 'dairy free', 'dill', 'dinner',
'drink', 'easter', 'edible gift', 'egg', 'eggplant', 'escarole', 'fall', 'family
reunion', 'fennel', 'fig', 'fish', 'florida', 'food processor', 'fruit', 'fry',

'game', 'garlic', 'ginger', 'goat cheese', 'green bean', 'green onion/scallion', 'grill', 'grill/barbecue', 'ground beef', 'ham', 'hanukkah', 'harpercollins', 'healthy', 'herb', 'high fiber', 'hominy/cornmeal/masa', 'honey', 'hot drink', 'house & garden', 'jalapeño', 'jícama', 'kale', 'kid-friendly', 'kidney friendly', 'kwanzaa', 'lamb', 'leafy green', 'leek', 'lemon', 'lima bean', 'lime', 'low cal', 'low cholesterol', 'low fat', 'lunar new year', 'lunch', 'mayonnaise', 'meat', 'melon', 'microwave', 'milk/cream', 'minnesota', 'mint', 'molasses', 'monterey jack', 'mozzarella', 'mushroom', 'mustard', 'mustard greens', "new year's day", "new year's eve", 'no sugar added', 'no-cook', 'noodle', 'nut', 'nutmeg', 'olive', 'onion', 'oscars', 'paleo', 'pan-fry', 'papaya', 'paprika', 'parade', 'parmesan', 'parsley', 'passover', 'pasta', 'pecan', 'pepper', 'picnic', 'pine nut', 'poach', 'pork', 'potato', 'potluck', 'poultry', 'radish', 'raisin', 'ramadan', 'raspberry', 'red wine', 'rice', 'roast', 'rosemary', 'rutabaga', 'saffron', 'sage', 'salad', 'santa monica', 'sauté', 'seafood', 'seed', 'shallot', 'shrimp', 'simmer', 'soup/stew', 'sour cream', 'soy sauce', 'spinach', 'spring', 'squash', "st. patrick's day", 'steam', 'stir-fry', 'stuffing/dressing', 'sugar conscious', 'sugar snap pea', 'sweet potato/yam', 'tailgating', 'tarragon', 'texas', 'thanksgiving', 'thyme', 'tomato', 'tree nut', 'vanilla', 'vegan', 'vegetable', 'vinegar', 'walnut', 'washington', 'white wine', 'wine', 'yellow squash', 'zucchini', 'turkey', 'bon appétit', 'soy free', 'gourmet', 'tree nut free', 'vegetarian', 'kosher', 'pescatarian', 'quick & easy', 'wheat/gluten-free', 'bake', 'summer', 'dessert', 'winter']
{0.0: 8940, 1.0: 1688}
{0.0: 4362, 1.0: 874}
Test accuracy of side is 0.9459511077158136
Test precision of side is 0.8832684824902723
Test recall of side is 0.7791762013729977
Test f1-score of side is 0.8279635258358663

Category fall
predictors ['almond', 'anise', 'anniversary', 'appetizer', 'apple', 'apricot', 'arugula', 'asian pear', 'asparagus', 'avocado', 'back to school', 'backyard bbq', 'bacon', 'banana', 'barley', 'basil', 'bass', 'bean', 'beef', 'beef tenderloin', 'beer', 'beet', 'bell pepper', 'berry', 'birthday', 'blackberry', 'blender', 'blue cheese', 'boil', 'bourbon', 'braise', 'brandy', 'bread', 'breadcrumbs', 'breakfast', 'brine', 'brisket', 'broccoli', 'broccoli rabe', 'broil', 'brunch', 'brussel sprout', 'bulgur', 'butternut squash', 'cabbage', 'cake', 'california', 'candy', 'capers', 'carrot', 'casserole/gratin', 'cauliflower', 'celery', 'cheddar', 'cheese', 'cherry', 'chestnut', 'chicken', 'chickpea', 'chill', 'chive', 'chocolate', 'christmas', 'christmas eve', 'cilantro', 'cinnamon', 'citrus', 'clam', 'clove', 'cobbler/crumble', 'cocktail', 'cocktail party', 'coconut', 'cod', 'coffee', 'cognac/armagnac', 'condiment/spread', 'connecticut', 'cookie', 'cookies', 'coriander', 'corn', 'cornmeal', 'couscous', 'crab', 'cranberry', 'cream cheese', 'cucumber', 'currant', 'curry', 'custard', 'dairy', 'dairy free', 'date', 'deep-fry', 'dill', 'dinner', 'dried fruit', 'drink', 'duck', 'edible gift', 'egg', 'egg

nog', 'eggplant', 'endive', 'engagement party', 'escarole', 'family reunion',
'fat free', "father's day", 'fennel', 'fig', 'fish', 'florida', 'food
processor', 'freeze/chill', 'friendsgiving', 'frozen dessert', 'fruit', 'fruit
juice', 'fry', 'game', 'garlic', 'ginger', 'goat cheese', 'gouda', 'graduation',
'grape', 'grapefruit', 'green onion/scallion', 'grill', 'grill/barbecue',
'ground beef', 'guam', 'halloween', 'ham', 'hanukkah', 'harpercollins',
'hazelnut', 'healthy', 'herb', 'high fiber', 'hominy/cornmeal/masa', 'honey',
'horseradish', 'hot pepper', 'house & garden', 'ice cream machine', 'jalapeño',
'jam or jelly', 'jerusalem artichoke', 'kale', 'kid-friendly', 'kidney
friendly', 'kiwi', 'kumquat', 'labor day', 'lamb', 'leafy green', 'leek',
'legume', 'lemon', 'lentil', 'lima bean', 'lime', 'liqueur', 'lobster', 'low
cal', 'low carb', 'low cholesterol', 'low sodium', 'low sugar', 'low/no sugar',
'lunch', 'macadamia nut', 'maine', 'mandoline', 'mango', 'maple syrup', 'mardi
gras', 'marinate', 'maryland', 'mayonnaise', 'meat', 'milk/cream', 'minnesota',
'missouri', 'mixer', 'molasses', "mother's day", 'mozzarella', 'mushroom',
'mussel', 'mustard', 'mustard greens', "new year's day", "new year's eve", 'new
york', 'no sugar added', 'no-cook', 'noodle', 'north carolina', 'nut', 'nutmeg',
'oklahoma', 'oktoberfest', 'olive', 'one-pot meal', 'onion', 'orange',
'oregano', 'oscars', 'oyster', 'pacific palisades', 'paleo', 'pan-fry',
'parmesan', 'parsley', 'parsnip', 'party', 'pasadena', 'passion fruit',
'passover', 'pasta', 'pastry', 'peach', 'peanut', 'pear', 'pecan', 'pepper',
'pernod', 'phyllo/puff pastry dough', 'picnic', 'pine nut', 'pineapple',
'pistachio', 'plum', 'pomegranate', 'pork', 'pork chop', 'pork tenderloin',
'port', 'potato', 'potluck', 'poultry', 'poultry sausage', 'pressure cooker',
'prosciutto', 'prune', 'pumpkin', 'quail', 'quince', 'radish', 'raisin',
'raspberry', 'red wine', 'rice', 'ricotta', 'roast', 'root vegetable',
'rosemary', 'rum', 'sage', 'salad', 'salad dressing', 'sandwich', 'sauce',
'sausage', 'sauté', 'scallop', 'seafood', 'seed', 'sesame', 'shallot',
'shavuot', 'shellfish', 'sherry', 'shower', 'shrimp', 'simmer', 'soup/stew',
'sour cream', 'soy sauce', 'spice', 'spinach', 'spring', 'squash', 'st. louis',
"st. patrick's day", 'steak', 'stew', 'stir-fry', 'stuffing/dressing', 'sugar
conscious', 'sukkot', 'super bowl', 'sweet potato/yam', 'swiss cheese',
'swordfish', 'taco', 'tailgating', 'tangerine', 'tarragon', 'tea', 'tequila',
'thanksgiving', 'thyme', 'tofu', 'tomato', 'tree nut', 'tuna', "valentine's
day", 'vanilla', 'veal', 'vegan', 'vegetable', 'vinegar', 'vodka', 'walnut',
'washington', 'watercress', 'whiskey', 'white wine', 'whole wheat', 'wine',
'yogurt', 'zucchini', 'turkey', 'bon appétit', 'peanut free', 'soy free',
'gourmet', 'tree nut free', 'vegetarian', 'kosher', 'pescatarian', 'quick &
easy', 'wheat/gluten-free', 'bake', 'summer', 'dessert', 'winter', 'side']
{0.0: 8949, 1.0: 1679}
{0.0: 4415, 1.0: 821}
Test accuracy of fall is 0.8754774637127578
Test precision of fall is 0.7046004842615012
Test recall of fall is 0.35444579780755175
Test f1-score of fall is 0.47163695299837927

Category dairy free

predictors ['22-minute meals', 'advance prep required', 'almond', 'amaretto', 'anchovy', 'anise', 'appetizer', 'apple', 'asparagus', 'back to school', 'backyard bbq', 'bacon', 'basil', 'bean', 'beef', 'beef tenderloin', 'beet', 'bell pepper', 'birthday', 'blackberry', 'blender', 'blue cheese', 'boil', 'bread', 'breakfast', 'broccoli', 'broccoli rabe', 'broil', 'brunch', 'butter', 'cabbage', 'cake', 'capers', 'cardamom', 'carrot', 'casserole/gratin', 'cauliflower', 'cheese', 'chicken', 'chickpea', 'chill', 'chive', 'chocolate', 'christmas', 'christmas eve', 'cilantro', 'citrus', 'clam', 'cobbler/crumble', 'cod', 'coffee', 'cookies', 'cornmeal', 'cucumber', 'curry', 'dairy', 'deep-fry', 'dinner', 'dried fruit', 'easter', 'egg', 'fat free', "father's day", 'fish', 'fortified wine', 'freeze/chill', 'frozen dessert', 'fruit', 'fruit juice', 'fry', 'garlic', 'goat cheese', 'gouda', 'grapefruit', 'green bean', 'grill', 'grill/barbecue', 'ground beef', 'ham', 'healthy', 'herb', 'high fiber', 'honey', 'hot pepper', 'jerusalem artichoke', 'jícama', 'kale', 'kid-friendly', 'kidney friendly', 'lamb', 'leafy green', 'leek', 'lemon', 'lemongrass', 'lettuce', 'lobster', 'low cal', 'low cholesterol', 'low fat', 'low sodium', 'lunch', 'maple syrup', 'marinate', 'mayonnaise', 'meat', 'microwave', 'milk/cream', 'mint', 'mixer', "mother's day", 'mushroom', 'mussel', 'mustard', "new year's day", 'no sugar added', 'no-cook', 'noodle', 'nut', 'nutmeg', 'oat', 'onion', 'orange', 'oregano', 'oyster', 'paleo', 'pan-fry', 'paprika', 'parmesan', 'parsnip', 'passover', 'pasta', 'pea', 'peach', 'peanut butter', 'pecan', 'pickles', 'picnic', 'pineapple', 'poker/game night', 'poppy', 'pork', 'pork chop', 'pork rib', 'pork tenderloin', 'port', 'potato', 'potluck', 'prune', 'quinoa', 'radish', 'raisin', 'ramadan', 'ramekin', 'rice', 'roast', 'rosemary', 'rosh hashanah/yom kippur', 'salad', 'salad dressing', 'salmon', 'sandwich', 'sauce', 'sausage', 'sauté', 'seafood', 'sesame', 'shallot', 'shrimp', 'simmer', 'soup/stew', 'sour cream', 'soy sauce', 'spice', 'spring', "st. patrick's day", 'steak', 'steam', 'stir-fry', 'stuffing/dressing', 'sugar conscious', 'sugar snap pea', 'super bowl', 'tea', 'thanksgiving', 'tomato', 'tortillas', 'trout', 'tuna', "valentine's day", 'vanilla', 'vegan', 'vegetable', 'vinegar', 'walnut', 'white wine', 'wok', 'yogurt', 'zucchini', 'snack', 'bon appétit', 'peanut free', 'soy free', 'gourmet', 'tree nut free', 'vegetarian', 'kosher', 'pescatarian', 'quick & easy', 'wheat/gluten-free', 'bake', 'summer', 'dessert', 'winter', 'side', 'fall']
{0.0: 8969, 1.0: 1659}
{0.0: 4405, 1.0: 831}
Test accuracy of dairy free is 0.9560733384262796
Test precision of dairy free is 0.9022757697456493
Test recall of dairy free is 0.8110709987966306
Test f1-score of dairy free is 0.8542458808618504

Category no sugar added
predictors ['22-minute meals', 'almond', 'appetizer', 'apple', 'artichoke', 'arugula', 'asparagus', 'back to school', 'backyard bbq', 'bacon', 'barley', 'bastille day', 'bean', 'beef', 'beef rib', 'beer', 'beet', 'bell pepper', 'birthday', 'blackberry', 'blender', 'blue cheese', 'boil', 'braise', 'brandy',

'bread', 'breakfast', 'broccoli', 'broil', 'brown rice', 'brunch', 'brussel
sprout', 'cabbage', 'carrot', 'casserole/gratin', 'celery', 'cheddar', 'cheese',
'chestnut', 'chicken', 'chickpea', 'chile pepper', 'chive', 'christmas',
'christmas eve', 'cilantro', 'citrus', 'cocktail party', 'coconut', 'cod',
'cognac/armagnac', 'collard greens', 'condiment', 'coriander', 'corn',
'cornmeal', 'couscous', 'crab', 'cranberry', 'cucumber', 'curry', 'dairy',
'deep-fry', 'dill', 'dinner', 'egg', 'endive', 'entertaining', 'family reunion',
'fat free', "father's day", 'fennel', 'feta', 'fig', 'fish', 'food processor',
'fourth of july', 'fruit', 'garlic', 'ginger', 'goat cheese', 'grapefruit',
'grill', 'grill/barbecue', 'ham', 'hanukkah', 'harpercollins', 'healthy',
'herb', 'high fiber', 'hominy/cornmeal/masa', 'honey', 'horseradish', 'hot
pepper', 'house & garden', 'jícama', 'kid-friendly', 'kidney friendly', 'lamb',
'leafy green', 'legume', 'lemon', 'lemon juice', 'lemongrass', 'lentil',
'lettuce', 'lime', 'lime juice', 'low cal', 'low carb', 'low cholesterol', 'low
fat', 'low sodium', 'low sugar', 'low/no sugar', 'lunch', 'maple syrup', 'mardi
gras', 'marinate', 'massachusetts', 'mayonnaise', 'meat', 'milk/cream', 'mint',
'molasses', 'mozzarella', 'mushroom', 'mussel', 'mustard', 'no-cook', 'olive',
'onion', 'orange', 'orange juice', 'oscars', 'paleo', 'pancake', 'paprika',
'parmesan', 'parsley', 'parsnip', 'party', 'passover', 'pasta', 'pastry', 'pea',
'pear', 'pepper', 'picnic', 'pineapple', 'poppy', 'pork', 'pork chop', 'pork
rib', 'pork tenderloin', 'potato', 'potluck', 'poultry', 'pressure cooker',
'quick and healthy', 'radish', 'red wine', 'rice', 'roast', 'root vegetable',
'rosemary', 'sage', 'salad', 'salmon', 'sandwich', 'sauce', 'sausage', 'sauté',
'scallop', 'seafood', 'seed', 'semolina', 'sesame', 'shallot', 'shellfish',
'shrimp', 'simmer', 'soup/stew', 'sour cream', 'soy', 'spice', 'spinach',
'spring', 'squash', 'squid', 'steam', 'stir-fry', 'stuffing/dressing', 'sugar
conscious', 'sugar snap pea', 'super bowl', 'sweet potato/yam', 'tailgating',
'tangerine', 'tarragon', 'tequila', 'thanksgiving', 'thyme', 'tomato', 'tuna',
'turnip', "valentine's day", 'veal', 'vegan', 'vegetable', 'vinegar', 'walnut',
'watermelon', 'white wine', 'wine', 'wok', 'yogurt', 'zucchini', 'bon appétit',
'peanut free', 'soy free', 'gourmet', 'tree nut free', 'vegetarian', 'kosher',
'pescatarian', 'quick & easy', 'wheat/gluten-free', 'bake', 'summer', 'dessert',
'winter', 'side', 'fall', 'dairy free']
{0.0: 8990, 1.0: 1638}
{0.0: 4456, 1.0: 780}
Test accuracy of no sugar added is 0.9276165011459129
Test precision of no sugar added is 0.7587096774193548
Test recall of no sugar added is 0.7538461538461538
Test f1-score of no sugar added is 0.7562700964630225

Category kidney friendly
predictors ['advance prep required', 'almond', 'amaretto', 'anchovy', 'anise',
'appetizer', 'apple', 'apricot', 'artichoke', 'arugula', 'asparagus', 'avocado',
'back to school', 'backyard bbq', 'bacon', 'banana', 'basil', 'bean', 'beef',
'bell pepper', 'berry', 'blackberry', 'blender', 'blue cheese', 'blueberry',
'boil', 'boston', 'bourbon', 'brandy', 'bread', 'breakfast', 'broccoli rabe',
'broil', 'brunch', 'butter', 'buttermilk', 'cabbage', 'cake', 'candy', 'candy

thermometer', 'carrot', 'celery', 'chartreuse', 'cheddar', 'cheese', 'cherry', 'chicken', 'chile pepper', 'chill', 'chive', 'chocolate', 'christmas', 'christmas eve', 'cinco de mayo', 'cinnamon', 'citrus', 'cocktail party', 'coconut', 'coffee', 'cognac/armagnac', 'condiment/spread', 'cookies', 'coriander', 'corn', 'cornmeal', 'crab', 'cranberry', 'cream cheese', 'cucumber', 'dairy', 'date', 'dill', 'dinner', 'dried fruit', 'drink', 'easter', 'edible gift', 'egg', 'eggplant', 'endive', 'engagement party', 'fat free', "father's day", 'fennel', 'feta', 'food processor', 'freeze/chill', 'frozen dessert', 'fruit', 'fruit juice', 'fry', 'game', 'garlic', 'gin', 'ginger', 'goat cheese', 'green onion/scallion', 'grill', 'grill/barbecue', 'ham', 'harpercollins', 'healthy', 'herb', 'high fiber', 'hominy/cornmeal/masa', 'honey', 'honeydew', 'horseradish', 'ice cream machine', 'jalapeño', 'jam or jelly', 'kid-friendly', 'kosher for passover', 'lamb shank', 'leafy green', 'leek', 'lemon', 'lime', 'liqueur', 'low cal', 'low fat', 'low sodium', 'low/no sugar', 'lunar new year', 'lunch', 'macadamia nut', 'maple syrup', 'massachusetts', 'mayonnaise', 'milk/cream', 'mint', 'mixer', "mother's day", 'mushroom', 'mustard', "new year's eve", 'no-cook', 'nut', 'oat', 'olive', 'omelet', 'onion', 'orange', 'orange juice', 'oscars', 'oyster', 'paleo', 'pan-fry', 'pancake', 'papaya', 'parmesan', 'parsley', 'parsnip', 'party', 'passion fruit', 'passover', 'pasta', 'pastry', 'pea', 'peach', 'peanut', 'pear', 'pecan', 'pepper', 'persimmon', 'phyllo/puff pastry dough', 'picnic', 'pine nut', 'pineapple', 'pistachio', 'plum', 'poach', 'pomegranate', 'poppy', 'pork rib', 'potato', 'potluck', 'poultry', 'prune', 'pumpkin', 'radish', 'raisin', 'ramadan', 'raspberry', 'raw', 'red wine', 'rhubarb', 'rice', 'roast', 'rosemary', 'rosh hashanah/yom kippur', 'rum', 'sage', 'salad', 'sauce', 'sauté', 'seafood', 'shallot', 'shower', 'shrimp', 'simmer', 'smoothie', 'soup/stew', 'sour cream', 'soy', 'spice', 'spinach', 'spirit', 'spring', "st. patrick's day", 'stir-fry', 'strawberry', 'sugar conscious', 'sweet potato/yam', 'swiss cheese', 'tea', 'thanksgiving', 'thyme', 'tofu', 'tomato', 'triple sec', "valentine's day", 'vanilla', 'vegan', 'vegetable', 'vinegar', 'vodka', 'walnut', 'wasabi', 'weelicious', 'whiskey', 'whole wheat', 'yogurt', 'zucchini', 'bon appétit', 'peanut free', 'soy free', 'gourmet', 'tree nut free', 'vegetarian', 'kosher', 'pescatarian', 'quick & easy', 'wheat/gluten-free', 'bake', 'summer', 'dessert', 'winter', 'side', 'fall', 'dairy free', 'no sugar added']
{0.0: 9335, 1.0: 1293}
{0.0: 4606, 1.0: 630}
Test accuracy of kidney friendly is 0.9155844155844156
Test precision of kidney friendly is 0.6721611721611722
Test recall of kidney friendly is 0.5825396825396826
Test f1-score of kidney friendly is 0.6241496598639457

Category sugar conscious
predictors ['advance prep required', 'anise', 'anniversary', 'appetizer', 'apple', 'artichoke', 'arugula', 'asparagus', 'avocado', 'backyard bbq', 'bacon', 'basil', 'bass', 'bean', 'beef', 'beef tenderloin', 'beet', 'bell pepper', 'biscuit', 'blue cheese', 'blueberry', 'boil', 'braise', 'brandy',

'bread', 'breakfast', 'brie', 'broccoli', 'broil', 'brown rice', 'brunch',
'brussel sprout', 'buffet', 'bulgur', 'butter', 'butternut squash', 'cabbage',
'cake', 'california', 'capers', 'carrot', 'cashew', 'casserole/gratin',
'cauliflower', 'celery', 'chard', 'cheddar', 'cheese', 'chestnut', 'chicken',
'chickpea', 'chile pepper', 'chill', 'chive', 'chocolate', 'christmas',
'christmas eve', 'cilantro', 'cinco de mayo', 'cinnamon', 'citrus', 'clam',
'cocktail party', 'coffee', 'colorado', 'condiment/spread', 'cookies', 'corn',
'cornmeal', 'crab', 'cream cheese', 'cucumber', 'cumin', 'curry', 'dairy',
'date', 'deep-fry', 'dill', 'dinner', 'diwali', 'duck', 'easter', 'egg',
'eggplant', 'endive', 'family reunion', 'fat free', 'fennel', 'feta', 'fish',
'food processor', 'fruit', 'garlic', 'ginger', 'goat cheese', 'grapefruit',
'green bean', 'green onion/scallion', 'grill/barbecue', 'halibut', 'ham',
'hanukkah', 'harpercollins', 'healthy', 'herb', 'high fiber',
'hominy/cornmeal/masa', 'honey', 'horseradish', 'hot pepper', 'ireland',
'jalapeño', 'jícama', 'kale', 'kid-friendly', 'kwanzaa', 'leafy green', 'leek',
'legume', 'lemon', 'lemon juice', 'lemongrass', 'lentil', 'lettuce', 'lime',
'low cal', 'low cholesterol', 'low/no sugar', 'lunch', 'mango', 'mardi gras',
'marinate', 'massachusetts', 'mayonnaise', 'meat', 'microwave', 'milk/cream',
'mint', 'mixer', 'molasses', 'mozzarella', 'mushroom', 'mussel', 'mustard',
'mustard greens', "new year's day", "new year's eve", 'no-cook', 'noodle',
'nut', 'oat', 'oktoberfest', 'olive', 'onion', 'orange', 'orzo', 'oscars',
'paleo', 'pan-fry', 'parmesan', 'parsley', 'parsnip', 'party', 'pasta',
'pastry', 'pea', 'peanut', 'pear', 'pepper', 'picnic', 'pine nut', 'pork', 'pork
chop', 'potato', 'potluck', 'poultry', 'prosciutto', 'purim', 'quinoa',
'radish', 'raisin', 'red wine', 'rice', 'ricotta', 'roast', 'saffron', 'sage',
'salad', 'salmon', 'sauce', 'sausage', 'sauté', 'scallop', 'seafood', 'sesame
oil', 'shallot', 'shellfish', 'shrimp', 'simmer', 'snapper', 'soup/stew', 'soy
sauce', 'spice', 'spinach', 'spring', 'squid', 'steam', 'stew',
'stuffing/dressing', 'sugar snap pea', 'super bowl', 'sweet potato/yam', 'swiss
cheese', 'tailgating', 'tarragon', 'thanksgiving', 'thyme', 'tofu', 'tomato',
'tortillas', 'tuna', 'turnip', "valentine's day", 'vegan', 'vegetable',
'vinegar', 'walnut', 'watercress', 'watermelon', 'weelicious', 'white wine',
'wok', 'yellow squash', 'yogurt', 'zucchini', 'bon appétit', 'peanut free', 'soy
free', 'gourmet', 'tree nut free', 'vegetarian', 'kosher', 'pescatarian', 'quick
& easy', 'wheat/gluten-free', 'bake', 'summer', 'dessert', 'winter', 'side',
'fall', 'dairy free', 'no sugar added', 'kidney friendly']
{0.0: 9341, 1.0: 1287}
{0.0: 4607, 1.0: 629}
Test accuracy of sugar conscious is 0.9146294881588999
Test precision of sugar conscious is 0.6685185185185185
Test recall of sugar conscious is 0.5739268680445151
Test f1-score of sugar conscious is 0.6176218990590249

Category dinner
predictors ['#wasteless', '3-ingredient recipes', 'advance prep required',
'anise', 'anniversary', 'appetizer', 'apple', 'artichoke', 'arugula',
'asparagus', 'back to school', 'backyard bbq', 'bacon', 'barley', 'basil',

'bastille day', 'bean', 'beef', 'beef tenderloin', 'bell pepper', 'blender',
'boil', 'bok choy', 'bourbon', 'braise', 'bread', 'breadcrumbs', 'breakfast',
'broccoli', 'broccoli rabe', 'broil', 'brunch', 'bulgur', 'buttermilk',
'butternut squash', 'cabbage', 'california', 'cantaloupe', 'carrot', 'cashew',
'casserole/gratin', 'cauliflower', 'caviar', 'chard', 'cheddar', 'cheese',
'chicken', 'chile pepper', 'chili', 'chill', 'chive', 'chocolate', 'christmas',
'christmas eve', 'cinco de mayo', 'cinnamon', 'citrus', 'cocktail party',
'coconut', 'cognac/armagnac', 'colorado', 'condiment', 'condiment/spread',
'cookie', 'coriander', 'corn', 'cumin', 'dairy', 'date', 'deep-fry', 'dill',
'drink', 'egg', 'family reunion', 'fat free', "father's day", 'fennel', 'feta',
'fish', 'fortified wine', 'fry', 'game', 'garlic', 'georgia', 'ginger', 'goat
cheese', 'green bean', 'green onion/scallion', 'grill', 'grill/barbecue', 'ham',
'hanukkah', 'harpercollins', 'hazelnut', 'healthy', 'herb', 'high fiber',
'hominy/cornmeal/masa', 'honey', 'horseradish', 'house & garden', 'jalapeño',
'kale', 'kid-friendly', 'kosher for passover', 'kumquat', 'lamb', 'lancaster',
'lasagna', 'leafy green', 'leek', 'legume', 'lemon', 'lentil', 'lettuce', 'low
cal', 'lunch', 'maple syrup', 'mardi gras', 'marinate', 'massachusetts',
'mayonnaise', 'meat', 'meatloaf', 'milk/cream', 'mixer', 'molasses', "mother's
day", 'mozzarella', 'mushroom', 'mussel', 'mustard', 'mustard greens', 'new
jersey', 'new orleans', 'no meat, no problem', 'no-cook', 'noodle', 'nut',
'olive', 'onion', 'orange', 'oscars', 'pan-fry', 'paprika', 'parmesan', 'party',
'pasta', 'pea', 'peanut', 'peanut butter', 'pepper', 'peru', 'picnic',
'pineapple', 'pistachio', 'poach', 'poker/game night', 'pomegranate', 'pork',
'port', 'potato', 'potluck', 'poultry', 'purim', 'quick and healthy', 'radish',
'raisin', 'rhubarb', 'rice', 'ricotta', 'roast', 'root vegetable', 'rosemary',
'rutabaga', 'saffron', 'sage', 'salad', 'sandwich', 'sardine', 'sauce',
'sausage', 'sauté', 'seafood', 'seed', 'shallot', 'shellfish', 'sherry',
'shrimp', 'simmer', 'slow cooker', 'soup/stew', 'soy', 'soy sauce', 'spice',
'spinach', 'spring', 'steak', 'stew', 'stir-fry', 'sweet potato/yam', 'swiss
cheese', 'tailgating', 'tamarind', 'tarragon', 'thanksgiving', 'thyme',
'tomatillo', 'tomato', 'tortillas', 'tree nut', 'turnip', "valentine's day",
'vegan', 'vegetable', 'vinegar', 'wedding', 'white wine', 'wine', 'yogurt',
'turkey', 'bon appétit', 'peanut free', 'soy free', 'gourmet', 'tree nut free',
'vegetarian', 'kosher', 'pescatarian', 'quick & easy', 'wheat/gluten-free',
'bake', 'summer', 'dessert', 'winter', 'side', 'fall', 'dairy free', 'no sugar
added', 'kidney friendly', 'sugar conscious']
{0.0: 9382, 1.0: 1246}
{0.0: 4584, 1.0: 652}
Test accuracy of dinner is 0.9428953399541635
Test precision of dinner is 0.8551307847082495
Test recall of dinner is 0.651840490797546
Test f1-score of dinner is 0.7397737162750218

Category onion
predictors ['3-ingredient recipes', 'advance prep required', 'almond',
'anchovy', 'anise', 'appetizer', 'apple', 'arizona', 'artichoke', 'arugula',
'asparagus', 'australia', 'avocado', 'backyard bbq', 'bacon', 'barley', 'basil',

'bass', 'bean', 'beef', 'beef rib', 'beef shank', 'beef tenderloin', 'beet',
'bell pepper', 'bitters', 'blender', 'blue cheese', 'boil', 'bok choy',
'braise', 'bread', 'breadcrumbs', 'breakfast', 'broccoli', 'broil', 'brunch',
'brussel sprout', 'butter', 'butternut squash', 'cake', 'california', 'canada',
'cantaloupe', 'capers', 'caraway', 'carrot', 'cashew', 'casserole/gratin',
'cauliflower', 'celery', 'chard', 'cheddar', 'cheese', 'cherry', 'chicken',
'chile pepper', 'chill', 'chive', 'christmas', 'cilantro', 'cinco de mayo',
'cinnamon', 'citrus', 'clam', 'clove', 'cocktail party', 'condiment/spread',
'coriander', 'corn', 'cornmeal', 'cottage cheese', 'couscous', 'crab',
'cranberry', 'cream cheese', 'cucumber', 'cumin', 'currant', 'curry', 'dairy',
'date', 'dill', 'dip', 'dried fruit', 'drink', 'egg', 'eggplant', 'endive',
'engagement party', 'escarole', 'family reunion', 'fat free', "father's day",
'fennel', 'feta', 'fish', 'florida', 'fontina', 'food processor', 'fortified
wine', 'fourth of july', 'freeze/chill', 'fruit', 'fruit juice', 'fry',
'garlic', 'georgia', 'gin', 'ginger', 'goat cheese', 'goose', 'graduation',
'grapefruit', 'green bean', 'grill', 'grill/barbecue', 'ground lamb', 'ham',
'hanukkah', 'harpercollins', 'healthy', 'herb', 'high fiber', 'honey',
'horseradish', 'hot pepper', 'house & garden', 'ireland', 'italy', 'jalapeño',
'jam or jelly', 'jerusalem artichoke', 'jícama', 'kale', 'kansas city',
'kentucky derby', 'kid-friendly', 'kiwi', 'lamb', 'lamb chop', 'lamb shank',
'leafy green', 'leek', 'legume', 'lemon', 'lentil', 'lettuce', 'lime', 'low
cal', 'low carb', 'low cholesterol', 'low fat', 'low sodium', 'low/no sugar',
'lunar new year', 'lunch', 'mango', 'maple syrup', 'mardi gras', 'marinate',
'maryland', 'massachusetts', 'mayonnaise', 'meat', 'milk/cream', 'mint',
'missouri', 'mixer', 'molasses', "mother's day", 'mozzarella', 'mushroom',
'mussel', 'mustard', "new year's day", "new year's eve", 'no-cook', 'noodle',
'north carolina', 'nut', 'okra', 'olive', 'orange', 'paleo', 'papaya',
'parmesan', 'parsley', 'parsnip', 'party', 'passover', 'pasta', 'pastry', 'pea',
'pecan', 'pepper', 'pernod', 'picnic', 'pine nut', 'pineapple', 'plantain',
'poach', 'poker/game night', 'pomegranate', 'pork', 'pork chop', 'pork
tenderloin', 'port', 'potato', 'potluck', 'poultry', 'prosciutto', 'prune',
'pumpkin', 'purim', 'rack of lamb', 'radish', 'raisin', 'ramadan', 'raw', 'red
wine', 'rice', 'ricotta', 'roast', 'root vegetable', 'rosemary', 'rutabaga',
'saffron', 'sage', 'salad', 'salad dressing', 'salmon', 'sandwich', 'sauce',
'sausage', 'sauté', 'scallop', 'seafood', 'seed', 'sesame oil', 'shallot',
'shellfish', 'sherry', 'shower', 'shrimp', 'simmer', 'soup/stew', 'sour cream',
'soy', 'soy sauce', 'spice', 'spinach', 'spirit', 'spring', 'squash', "st.
patrick's day", 'steak', 'steam', 'stew', 'stir-fry', 'stock', 'sugar snap pea',
'super bowl', 'sweet potato/yam', 'swiss cheese', 'taco', 'tailgating',
'tangerine', 'tarragon', 'tequila', 'texas', 'thanksgiving', 'thyme', 'tofu',
'tomatillo', 'tomato', 'tortillas', 'tree nut', 'trout', 'tuna', 'turnip',
"valentine's day", 'veal', 'vegan', 'vegetable', 'vinegar', 'walnut', 'wasabi',
'watermelon', 'weelicious', 'white wine', 'wild rice', 'wok', 'yellow squash',
'yogurt', 'zucchini', 'turkey', 'bon appétit', 'peanut free', 'soy free',
'gourmet', 'tree nut free', 'vegetarian', 'kosher', 'pescatarian', 'quick &
easy', 'wheat/gluten-free', 'bake', 'summer', 'dessert', 'winter', 'side',
'fall', 'dairy free', 'no sugar added', 'kidney friendly', 'sugar conscious',
'dinner']

```
{0.0: 9365, 1.0: 1263}
{0.0: 4651, 1.0: 585}
Test accuracy of onion is 0.8875095492742552
Test precision of onion is 0.4807692307692308
Test recall of onion is 0.08547008547008547
Test f1-score of onion is 0.1451378809869376


Category milk/cream
predictors ['#cakeweek', '3-ingredient recipes', 'advance prep required',
'alcoholic', 'almond', 'amaretto', 'anise', 'anniversary', 'appetizer', 'apple',
'apricot', 'artichoke', 'asparagus', 'avocado', 'backyard bbq', 'bacon',
'banana', 'basil', 'bastille day', 'bean', 'beef', 'beef tenderloin', 'bell
pepper', 'berry', 'birthday', 'bitters', 'blackberry', 'blender', 'blue cheese',
'blueberry', 'boil', 'boston', 'bourbon', 'braise', 'brandy', 'bread',
'breadcrumbs', 'breakfast', 'brine', 'broccoli', 'broil', 'brunch', 'brussel
sprout', 'butter', 'buttermilk', 'butternut squash', 'butterscotch/caramel',
'cabbage', 'cake', 'california', 'calvados', 'candy thermometer', 'capers',
'carrot', 'cashew', 'casserole/gratin', 'cauliflower', 'celery', 'champagne',
'chard', 'cheddar', 'cheese', 'cherry', 'chestnut', 'chicken', 'chill', 'chive',
'chocolate', 'christmas', 'christmas eve', 'cilantro', 'cinco de mayo',
'cinnamon', 'citrus', 'clam', 'cocktail', 'cocktail party', 'coconut', 'coffee',
'cognac/armagnac', 'condiment/spread', 'cookie', 'cookies', 'coriander', 'corn',
'cornmeal', 'couscous', 'crab', 'cranberry', 'cream cheese', 'créme de cacao',
'cucumber', 'curry', 'dairy', 'date', 'deep-fry', 'digestif', 'dill', 'double
boiler', 'dried fruit', 'drink', 'easter', 'eau de vie', 'egg', 'endive',
'engagement party', 'family reunion', 'fennel', 'feta', 'fig', 'fish', 'food
processor', 'fortified wine', 'fourth of july', 'france', 'freeze/chill',
'freezer food', 'frozen dessert', 'fruit', 'fruit juice', 'fry', 'garlic',
'gin', 'ginger', 'goat cheese', 'grapefruit', 'grappa', 'green bean', 'grill',
'grill/barbecue', 'ground beef', 'ham', 'hanukkah', 'hawaii', 'hazelnut',
'healthy', 'herb', 'high fiber', 'hominy/cornmeal/masa', 'honey', 'horseradish',
'hot drink', 'house & garden', 'ice cream machine', 'italy', 'jam or jelly',
'kahlúa', 'kid-friendly', 'lamb', 'leafy green', 'leek', 'lemon', 'lemon juice',
'lemongrass', 'lima bean', 'lime', 'liqueur', 'louisiana', 'low cal', 'low
carb', 'low fat', 'low sodium', 'lunch', 'mango', 'maple syrup', 'marinate',
'marsala', 'mayonnaise', 'microwave', 'mint', 'mixer', 'molasses', 'monterey
jack', 'mushroom', 'mussel', 'mustard', "new year's day", "new year's eve", 'no-
cook', 'north carolina', 'nut', 'nutmeg', 'oktoberfest', 'olive', 'one-pot
meal', 'orange', 'organic', 'oscars', 'oyster', 'pan-fry', 'parmesan',
'parsley', 'parsnip', 'party', 'pasadena', 'pasta', 'pea', 'peach', 'peanut',
'pear', 'pecan', 'pennsylvania', 'pepper', 'pernod', 'phyllo/puff pastry dough',
'picnic', 'pineapple', 'pork', 'pork tenderloin', 'potato', 'potluck',
'poultry', 'poultry sausage', 'prune', 'pumpkin', 'quinoa', 'radish', 'raisin',
'raspberry', 'rice', 'ricotta', 'roast', 'root vegetable', 'rosemary', 'rum',
'rutabaga', 'saffron', 'sage', 'salad', 'salad dressing', 'salmon', 'san
francisco', 'sandwich', 'sauce', 'sausage', 'sauté', 'scallop', 'seafood',
'seed', 'semolina', 'shallot', 'shavuot', 'shellfish', 'sherry', 'shower',
```

'shrimp', 'simmer', 'soup/stew', 'sour cream', 'sparkling wine', 'spice',
'spinach', 'spirit', 'spring', 'squash', "st. patrick's day", 'steam', 'stew',
'strawberry', 'super bowl', 'sweet potato/yam', 'swiss cheese', 'tangerine',
'tapioca', 'tarragon', 'tea', 'thanksgiving', 'thyme', 'tomatillo', 'tomato',
'tortillas', 'tropical fruit', 'trout', 'tuna', 'turnip', 'utah', "valentine's
day", 'vanilla', 'veal', 'vegan', 'vegetable', 'vinegar', 'virginia', 'vodka',
'walnut', 'watercress', 'wedding', 'whiskey', 'white wine', 'wine', 'yogurt',
'zucchini', 'snack', 'turkey', 'bon appétit', 'peanut free', 'soy free',
'gourmet', 'tree nut free', 'vegetarian', 'kosher', 'pescatarian', 'quick &
easy', 'wheat/gluten-free', 'bake', 'summer', 'dessert', 'winter', 'side',
'fall', 'dairy free', 'no sugar added', 'kidney friendly', 'sugar conscious',
'dinner', 'onion']
{0.0: 9482, 1.0: 1146}
{0.0: 4669, 1.0: 567}
Test accuracy of milk/cream is 0.8997326203208557
Test precision of milk/cream is 0.584
Test recall of milk/cream is 0.25749559082892415
Test f1-score of milk/cream is 0.35740514075887386

Category sauté
predictors ['almond', 'anise', 'anniversary', 'appetizer', 'apple', 'artichoke',
'arugula', 'asparagus', 'backyard bbq', 'bacon', 'banana', 'barley', 'basil',
'bass', 'bean', 'beef', 'beef shank', 'bell pepper', 'berry', 'birthday',
'blender', 'boil', 'bon app tit', 'braise', 'brandy', 'bread', 'breadcrumbs',
'breakfast', 'broccoli', 'broccoli rabe', 'broil', 'brown rice', 'brunch',
'brussel sprout', 'bulgur', 'cabbage', 'cake', 'california', 'capers',
'caraway', 'carrot', 'cashew', 'casserole/gratin', 'cauliflower', 'celery',
'chard', 'cheese', 'cherry', 'chestnut', 'chicken', 'chickpea', 'chile pepper',
'chill', 'christmas', 'cilantro', 'cinco de mayo', 'cinnamon', 'citrus', 'clam',
'clove', 'cocktail party', 'coconut', 'cod', 'cognac/armagnac', 'collard
greens', 'condiment', 'condiment/spread', 'connecticut', 'cookie', 'corn',
'cornmeal', 'cottage cheese', 'crab', 'cranberry', 'cream cheese', 'cucumber',
'currant', 'curry', 'dairy', 'date', 'dill', 'double boiler', 'dried fruit',
'drink', 'duck', 'easter', 'egg', 'eggplant', 'endive', 'escarole', 'family
reunion', "father's day", 'fennel', 'feta', 'fish', 'florida', 'fontina', 'food
processor', 'fortified wine', 'france', 'fruit', 'fry', 'game', 'garlic',
'ginger', 'goat cheese', 'green bean', 'grill/barbecue', 'ground beef',
'halibut', 'ham', 'hazelnut', 'healthy', 'herb', 'high fiber',
'hominy/cornmeal/masa', 'honey', 'horseradish', 'hot pepper', 'house & garden',
'jalapeño', 'jerusalem artichoke', 'kale', 'kentucky derby', 'kid-friendly',
'lamb', 'lamb chop', 'lancaster', 'leafy green', 'leek', 'lemon', 'lemongrass',
'lentil', 'lettuce', 'lima bean', 'lime', 'lingonberry', 'los angeles', 'low
cal', 'low carb', 'low cholesterol', 'low fat', 'low sodium', 'low/no sugar',
'lunch', 'maine', 'mango', 'mardi gras', 'marinate', 'marsala', 'maryland',
'massachusetts', 'mayonnaise', 'meat', 'minnesota', 'mint', 'mixer', "mother's
day", 'mushroom', 'mussel', 'mustard', 'mustard greens', 'new jersey', "new
year's day", "new year's eve", 'no-cook', 'noodle', 'nut', 'oklahoma', 'olive',

'orange', 'oregano', 'paleo', 'pan-fry', 'parmesan', 'parsley', 'parsnip',
'passover', 'pasta', 'pea', 'peanut', 'pear', 'pecan', 'pennsylvania', 'pepper',
'picnic', 'pine nut', 'pistachio', 'poach', 'pork', 'pork chop', 'pork
tenderloin', 'port', 'potato', 'poultry', 'prosciutto', 'prune', 'quick and
healthy', 'rabbit', 'raisin', 'red wine', 'rice', 'ricotta', 'roast', 'root
vegetable', 'rosemary', 'saffron', 'sage', 'salad', 'salmon', 'san francisco',
'santa monica', 'sauce', 'sausage', 'scallop', 'seafood', 'seed', 'shallot',
'shavuot', 'shellfish', 'sherry', 'shrimp', 'simmer', 'soup/stew', 'sour cream',
'south carolina', 'soy', 'soy sauce', 'spice', 'spinach', 'spring', 'squash',
'squid', "st. patrick's day", 'steak', 'steam', 'stew', 'stir-fry',
'stuffing/dressing', 'sugar snap pea', 'super bowl', 'sweet potato/yam',
'tarragon', 'thanksgiving', 'thyme', 'tofu', 'tomato', 'tortillas', 'triple
sec', 'trout', 'tuna', 'turnip', "valentine's day", 'veal', 'vegan',
'vegetable', 'vinegar', 'vodka', 'walnut', 'wasabi', 'washington', 'watercress',
'white wine', 'wine', 'yellow squash', 'yogurt', 'zucchini', 'turkey', 'bon
appétit', 'peanut free', 'soy free', 'gourmet', 'tree nut free', 'vegetarian',
'kosher', 'pescatarian', 'quick & easy', 'wheat/gluten-free', 'bake', 'summer',
'dessert', 'winter', 'side', 'fall', 'dairy free', 'no sugar added', 'kidney
friendly', 'sugar conscious', 'dinner', 'onion', 'milk/cream']
{0.0: 9490, 1.0: 1138}
{0.0: 4685, 1.0: 551}
Test accuracy of sauté is 0.9033613445378151
Test precision of sauté is 0.5991189427312775
Test recall of sauté is 0.24682395644283123
Test f1-score of sauté is 0.3496143958868895

Category tomato
predictors ['22-minute meals', '3-ingredient recipes', '30 days of groceries',
'advance prep required', 'alcoholic', 'almond', 'anise', 'appetizer',
'artichoke', 'arugula', 'avocado', 'backyard bbq', 'bacon', 'basil', 'bass',
'bean', 'beef', 'beef rib', 'beef tenderloin', 'beer', 'beet', 'bell pepper',
'bitters', 'blender', 'blue cheese', 'boil', 'bourbon', 'braise', 'bread',
'breadcrumbs', 'breakfast', 'brine', 'brisket', 'broccoli rabe', 'broil',
'brunch', 'buffet', 'bulgur', 'butter', 'butternut squash', 'cabbage',
'california', 'capers', 'carrot', 'cashew', 'casserole/gratin', 'cauliflower',
'celery', 'chard', 'cheddar', 'cheese', 'chicken', 'chickpea', 'chile pepper',
'chill', 'chocolate', 'christmas', 'christmas eve', 'cilantro', 'cinco de mayo',
'citrus', 'clam', 'clove', 'cocktail', 'cocktail party', 'coconut', 'coffee',
'cognac/armagnac', 'collard greens', 'condiment', 'condiment/spread', 'cookie',
'cookies', 'corn', 'cornmeal', 'couscous', 'crab', 'cucumber', 'currant',
'curry', 'dairy', 'date', 'deep-fry', 'dill', 'diwali', 'drink', 'duck',
'easter', 'egg', 'eggplant', 'endive', 'escarole', 'family reunion', 'fat free',
"father's day", 'fennel', 'feta', 'fish', 'florida', 'food processor',
'freeze/chill', 'frozen dessert', 'fruit', 'fruit juice', 'fry', 'garlic',
'ginger', 'goat cheese', 'goose', 'gouda', 'graduation', 'green bean', 'green
onion/scallion', 'grill', 'grill/barbecue', 'ground beef', 'halibut', 'ham',
'hamburger', 'hazelnut', 'healthy', 'herb', 'high fiber',

'hominy/cornmeal/masa', 'horseradish', 'hot pepper', 'house & garden', 'italy',
'jalapeño', 'jícama', 'kentucky derby', 'kid-friendly', 'lamb', 'lamb chop',
'leafy green', 'leek', 'lemon', 'lemon juice', 'lentil', 'lettuce', 'lime',
'lime juice', 'liqueur', 'lobster', 'los angeles', 'low cal', 'low carb', 'low
cholesterol', 'low fat', 'low/no sugar', 'lunar new year', 'lunch', 'maine',
'mango', 'maple syrup', 'marinate', 'marsala', 'massachusetts', 'mayonnaise',
'meat', 'mexico', 'mint', 'mixer', 'molasses', 'monterey jack', "mother's day",
'mozzarella', 'mushroom', 'mussel', 'mustard', 'nectarine', 'new jersey', "new
year's eve", 'new york', 'no-cook', 'noodle', 'nut', 'nutmeg', 'ohio',
'oklahoma', 'okra', 'olive', 'orange', 'oregano', 'oyster', 'paleo', 'pan-fry',
'parmesan', 'parsley', 'party', 'pasta', 'pasta maker', 'pea', 'peanut',
'pecan', 'pepper', 'pernod', 'phyllo/puff pastry dough', 'picnic', 'pine nut',
'pittsburgh', 'plantain', 'poach', 'pomegranate', 'pork', 'pork rib', 'pork
tenderloin', 'port', 'portland', 'potato', 'potluck', 'poultry', 'prosciutto',
'quick and healthy', 'rack of lamb', 'raisin', 'ramadan', 'red wine', 'rhode
island', 'rice', 'ricotta', 'roast', 'root vegetable', 'rosemary', 'rutabaga',
'saffron', 'sage', 'salad', 'salmon', 'salsa', 'san francisco', 'sandwich',
'sauce', 'sausage', 'scallop', 'seafood', 'shellfish', 'sherry', 'shrimp',
'simmer', 'snapper', 'soup/stew', 'sour cream', 'soy', 'soy sauce', 'spain',
'spice', 'spinach', 'spirit', 'spring', 'squash', 'steak', 'stew', 'stir-fry',
'strawberry', 'super bowl', 'swiss cheese', 'switzerland', 'swordfish',
'tailgating', 'tarragon', 'tequila', 'thanksgiving', 'thyme', 'tofu',
'tomatillo', 'tortillas', 'tuna', "valentine's day", 'veal', 'vegan',
'vegetable', 'vermont', 'vinegar', 'virginia', 'vodka', 'washington, d.c.',
'white wine', 'wine', 'yellow squash', 'yogurt', 'zucchini', 'turkey', 'bon
appétit', 'peanut free', 'soy free', 'gourmet', 'tree nut free', 'vegetarian',
'kosher', 'pescatarian', 'quick & easy', 'wheat/gluten-free', 'bake', 'summer',
'dessert', 'winter', 'side', 'fall', 'dairy free', 'no sugar added', 'kidney
friendly', 'sugar conscious', 'dinner', 'onion', 'milk/cream', 'sauté']
{0.0: 9499, 1.0: 1129}
{0.0: 4683, 1.0: 553}
Test accuracy of tomato is 0.8949579831932774
Test precision of tomato is 0.5081081081081081
Test recall of tomato is 0.16998191681735986
Test f1-score of tomato is 0.25474254742547425

Category vegetable
predictors ['advance prep required', 'almond', 'appetizer', 'apple',
'artichoke', 'arugula', 'asparagus', 'avocado', 'backyard bbq', 'bacon',
'barley', 'basil', 'bean', 'beef', 'beef shank', 'beet', 'bell pepper',
'blender', 'blue cheese', 'blueberry', 'boil', 'bourbon', 'braise', 'brandy',
'bread', 'breakfast', 'brine', 'broccoli', 'broccoli rabe', 'broil', 'brunch',
'brussel sprout', 'buffet', 'bulgur', 'buttermilk', 'butternut squash',
'cabbage', 'california', 'capers', 'caraway', 'carrot', 'cashew',
'casserole/gratin', 'cauliflower', 'celery', 'chard', 'cheese', 'cherry',
'chicken', 'chickpea', 'chill', 'chocolate', 'christmas', 'christmas eve',
'cilantro', 'cinco de mayo', 'cinnamon', 'citrus', 'clam', 'clove', 'cocktail

party', 'coconut', 'cod', 'condiment', 'condiment/spread', 'cookie', 'corn',
'couscous', 'crab', 'cranberry', 'cream cheese', 'cucumber', 'curry', 'dairy',
'dill', 'dried fruit', 'drink', 'easter', 'egg', 'eggplant', 'fennel', 'feta',
'fig', 'fish', 'florida', 'food processor', 'fourth of july', 'france',
'freeze/chill', 'fruit', 'fry', 'garlic', 'ginger', 'goat cheese', 'gouda',
'graduation', 'green bean', 'grill/barbecue', 'harpercollins', 'hazelnut',
'healthy', 'herb', 'high fiber', 'honey', 'horseradish', 'hot pepper', 'house &
garden', 'ireland', 'jalapeño', 'jerusalem artichoke', 'kentucky derby', 'kid-
friendly', 'lamb', 'lamb shank', 'leafy green', 'leek', 'legume', 'lemon',
'lentil', 'lettuce', 'lima bean', 'lime', 'lobster', 'louisiana', 'low cal',
'low carb', 'low fat', 'low/no sugar', 'lunch', 'mandoline', 'mango', 'maple
syrup', 'marinate', 'mayonnaise', 'meat', 'microwave', 'missouri', 'mixer',
"mother's day", 'mozzarella', 'mushroom', 'mustard', "new year's day", "new
year's eve", 'no-cook', 'noodle', 'nut', 'nutmeg', 'okra', 'oktoberfest',
'olive', 'one-pot meal', 'orange', 'paleo', 'pan-fry', 'parmesan', 'parsley',
'parsnip', 'party', 'pasadena', 'passover', 'pasta', 'pastry', 'pea', 'peanut',
'pennsylvania', 'pepper', 'phyllo/puff pastry dough', 'picnic', 'pine nut',
'poach', 'poker/game night', 'pomegranate', 'pork', 'potato', 'potluck',
'poultry', 'prosciutto', 'prune', 'pumpkin', 'purim', 'radish', 'raisin',
'ramadan', 'raw', 'rhubarb', 'rice', 'ricotta', 'roast', 'root vegetable',
'rosemary', 'rum', 'saffron', 'sage', 'salad', 'salad dressing', 'salmon',
'sandwich', 'santa monica', 'sauce', 'sausage', 'seafood', 'seed', 'self',
'sesame', 'shallot', 'shellfish', 'shrimp', 'simmer', 'snapper', 'soup/stew',
'sour cream', 'soy', 'soy sauce', 'spice', 'spinach', 'spring', 'squash',
'steak', 'steam', 'stew', 'stir-fry', 'stuffing/dressing', 'sugar snap pea',
'sweet potato/yam', 'tailgating', 'thanksgiving', 'thyme', 'tomatillo',
'tortillas', 'tree nut', 'turnip', "valentine's day", 'veal', 'vegan',
'vinegar', 'walnut', 'watercress', 'watermelon', 'white wine', 'wine', 'wok',
'yogurt', 'zucchini', 'turkey', 'bon appétit', 'peanut free', 'soy free',
'gourmet', 'tree nut free', 'vegetarian', 'kosher', 'pescatarian', 'quick &
easy', 'wheat/gluten-free', 'bake', 'summer', 'dessert', 'winter', 'side',
'fall', 'dairy free', 'no sugar added', 'kidney friendly', 'sugar conscious',
'dinner', 'onion', 'milk/cream', 'sauté', 'tomato']
{0.0: 9506, 1.0: 1122}
{0.0: 4683, 1.0: 553}
Test accuracy of vegetable is 0.8972498090145149
Test precision of vegetable is 0.5405405405405406
Test recall of vegetable is 0.18083182640144665
Test f1-score of vegetable is 0.27100271002710025

Category fruit
predictors ['advance prep required', 'alcoholic', 'almond', 'amaretto', 'anise',
'anniversary', 'appetizer', 'apple', 'apple juice', 'apricot', 'asian pear',
'avocado', 'back to school', 'backyard bbq', 'bacon', 'banana', 'beet', 'bell
pepper', 'berry', 'blackberry', 'blender', 'blue cheese', 'blueberry',
'bourbon', 'brandy', 'bread', 'breakfast', 'broil', 'brunch', 'butter',
'butterscotch/caramel', 'cabbage', 'cake', 'california', 'cantaloupe', 'cashew',

'casserole/gratin', 'champagne', 'cheddar', 'cheese', 'cherry', 'chestnut',
'chill', 'chocolate', 'christmas', 'christmas eve', 'cinnamon', 'citrus',
'clove', 'cocktail', 'cocktail party', 'coconut', 'coffee', 'cognac/armagnac',
'condiment/spread', 'cookies', 'cottage cheese', 'cranberry', 'cream cheese',
'créme de cacao', 'cucumber', 'currant', 'dairy', 'date', 'deep-fry', 'diwali',
'dried fruit', 'drink', 'duck', 'easter', 'eau de vie', 'edible gift', 'egg',
'engagement party', 'escarole', 'family reunion', 'fat free', 'fennel', 'fig',
'fish', 'food processor', 'freeze/chill', 'fruit juice', 'garlic', 'gin',
'ginger', 'grape', 'grapefruit', 'green onion/scallion', 'grill',
'grill/barbecue', 'hanukkah', 'harpercollins', 'hazelnut', 'healthy', 'herb',
'high fiber', 'honey', 'honeydew', 'hot pepper', 'ice cream machine',
'jalapeño', 'jam or jelly', 'kale', 'kid-friendly', 'kirsch', 'kiwi', 'kumquat',
'leafy green', 'lemon', 'lime', 'lingonberry', 'liqueur', 'low cal', 'low
cholesterol', 'low fat', 'low sodium', 'low/no sugar', 'lunch', 'macadamia nut',
'mango', 'marinate', 'marsala', 'meat', 'melon', 'mint', 'mixer', 'molasses',
"mother's day", 'mushroom', 'mustard', 'nectarine', "new year's day", 'no-cook',
'nut', 'nutmeg', 'oat', 'octopus', 'orange', 'papaya', 'passion fruit', 'pasta',
'pastry', 'peach', 'peanut', 'peanut butter', 'pear', 'pecan', 'pepper',
'persimmon', 'phyllo/puff pastry dough', 'picnic', 'pineapple', 'plum', 'poach',
'pork', 'pork tenderloin', 'port', 'potluck', 'poultry', 'prosciutto', 'prune',
'pumpkin', 'quail', 'quince', 'raisin', 'ramadan', 'raspberry', 'red wine',
'rhode island', 'rhubarb', 'rice', 'roast', 'rosemary', 'rum', 'saffron',
'salad', 'sauce', 'seafood', 'self', 'sesame', 'shavuot', 'sherry', 'shower',
'simmer', 'soup/stew', 'sour cream', 'south carolina', 'spice', 'spinach',
'spirit', 'spring', "st. patrick's day", 'stew', 'strawberry', 'sweet
potato/yam', 'tapioca', 'tea', 'tequila', 'thanksgiving', 'tropical fruit',
"valentine's day", 'vanilla', 'vegan', 'vodka', 'walnut', 'watercress',
'watermelon', 'white wine', 'wine', 'yogurt', 'yuca', 'turkey', 'bon appétit',
'peanut free', 'soy free', 'gourmet', 'tree nut free', 'vegetarian', 'kosher',
'pescatarian', 'quick & easy', 'wheat/gluten-free', 'bake', 'summer', 'dessert',
'winter', 'side', 'fall', 'dairy free', 'no sugar added', 'kidney friendly',
'sugar conscious', 'dinner', 'onion', 'milk/cream', 'sauté', 'tomato',
'vegetable']
{0.0: 9546, 1.0: 1082}
{0.0: 4724, 1.0: 512}
Test accuracy of fruit is 0.921886936592819
Test precision of fruit is 0.6484149855907781
Test recall of fruit is 0.439453125
Test f1-score of fruit is 0.5238649592549476

Category vegan
predictors ['appetizer', 'artichoke', 'asparagus', 'banana', 'barley', 'basil',
'bean', 'beer', 'bell pepper', 'berry', 'birthday', 'blender', 'blueberry',
'bread', 'butter', 'butternut squash', 'cabbage', 'cake', 'candy', 'cantaloupe',
'capers', 'caraway', 'cardamom', 'carrot', 'cashew', 'casserole/gratin',
'cauliflower', 'celery', 'champagne', 'cheese', 'cherry', 'chickpea', 'chile
pepper', 'chill', 'chive', 'christmas eve', 'cilantro', 'cinnamon', 'cocktail

party', 'colorado', 'condiment/spread', 'cookies', 'corn', 'cucumber', 'dairy', 'date', 'dip', 'egg', 'eggplant', 'escarole', 'fat free', 'fennel', 'feta', 'fry', 'garlic', 'ginger', 'goat cheese', 'graduation', 'grapefruit', 'green bean', 'grill', 'hanukkah', 'hazelnut', 'healthy', 'herb', 'high fiber', 'honey', 'honeydew', 'jam or jelly', 'kid-friendly', 'kosher for passover', 'leafy green', 'legume', 'lemon', 'lettuce', 'lime juice', 'low cal', 'low cholesterol', 'low fat', 'low sodium', 'low/no sugar', 'lunch', 'mayonnaise', 'melon', 'mint', 'mixer', 'mushroom', 'mustard', 'nectarine', 'no-cook', 'noodle', 'nut', 'oat', 'olive', 'orange', 'paleo', 'pan-fry', 'paprika', 'parmesan', 'parsley', 'parsnip', 'party', 'passover', 'pasta', 'pasta maker', 'pea', 'peach', 'picnic', 'pine nut', 'pistachio', 'poach', 'pomegranate', 'potato', 'potluck', 'pumpkin', 'quinoa', 'radish', 'ramadan', 'raw', 'red wine', 'rhubarb', 'rice', 'salad', 'sauce', 'seed', 'sesame', 'simmer', 'soup/stew', 'south carolina', 'soy', 'soy sauce', 'spice', 'spinach', 'spring', 'stuffing/dressing', 'tamarind', 'tangerine', 'tarragon', 'tomatillo', 'vermont', 'vinegar', 'walnut', 'watermelon', 'white wine', 'wok', 'yogurt', 'bon appétit', 'peanut free', 'soy free', 'gourmet', 'tree nut free', 'vegetarian', 'pescatarian', 'quick & easy', 'wheat/gluten-free', 'bake', 'summer', 'dessert', 'winter', 'side', 'fall', 'dairy free', 'no sugar added', 'kidney friendly', 'sugar conscious', 'dinner', 'onion', 'milk/cream', 'sauté', 'tomato', 'vegetable', 'fruit']
{0.0: 9626, 1.0: 1002}
{0.0: 4783, 1.0: 453}
Test accuracy of vegan is 0.9684873949579832
Test precision of vegan is 0.8582089552238806
Test recall of vegan is 0.7615894039735099
Test f1-score of vegan is 0.8070175438596492

Category egg
predictors ['22-minute meals', 'advance prep required', 'alcoholic', 'almond', 'amaretto', 'anise', 'appetizer', 'apple', 'apricot', 'artichoke', 'asian pear', 'asparagus', 'avocado', 'backyard bbq', 'bacon', 'banana', 'basil', 'bastille day', 'bean', 'beef', 'beef tenderloin', 'beer', 'beet', 'bell pepper', 'berry', 'birthday', 'bitters', 'blender', 'blue cheese', 'blueberry', 'boil', 'bourbon', 'braise', 'brandy', 'bread', 'breadcrumbs', 'breakfast', 'brie', 'broil', 'brunch', 'buffet', 'butter', 'cake', 'california', 'candy', 'candy thermometer', 'capers', 'cardamom', 'carrot', 'casserole/gratin', 'celery', 'chard', 'chartreuse', 'cheddar', 'cheese', 'cherry', 'chestnut', 'chicken', 'chill', 'chive', 'chocolate', 'christmas', 'christmas eve', 'cinco de mayo', 'cinnamon', 'citrus', 'cocktail', 'cocktail party', 'coconut', 'coffee', 'colorado', 'condiment', 'condiment/spread', 'cookie', 'cookies', 'coriander', 'corn', 'cornmeal', 'cottage cheese', 'crab', 'cranberry', 'cream cheese', 'cucumber', 'curry', 'dairy', 'dallas', 'date', 'deep-fry', 'denver', 'dill', 'dried fruit', 'drink', 'easter', 'edible gift', 'eggplant', 'endive', 'engagement party', 'fat free', "father's day", 'fennel', 'feta', 'fig', 'fish', 'flaming hot summer', 'fontina', 'food processor', 'fortified wine', 'france', 'frangelico', 'freeze/chill', 'frozen dessert', 'fry', 'game', 'garlic', 'gin',

'ginger', 'goat cheese', 'graduation', 'grains', 'grapefruit', 'green onion/scallion', 'grill/barbecue', 'halloween', 'ham', 'hanukkah', 'hazelnut', 'healthy', 'herb', 'high fiber', 'hominy/cornmeal/masa', 'honey', 'hot pepper', 'house & garden', 'ice cream machine', 'jalapeño', 'jam or jelly', 'kahlúa', 'kid-friendly', 'kirsch', 'kosher for passover', 'leafy green', 'leek', 'lemon', 'lemon juice', 'lime', 'lime juice', 'liqueur', 'low cal', 'low carb', 'low cholesterol', 'lunar new year', 'lunch', 'macadamia nut', 'mango', 'maple syrup', 'marinate', 'marsala', 'mayonnaise', 'mint', 'mixer', 'molasses', 'monterey jack', "mother's day", 'mozzarella', 'mushroom', 'mustard', "new year's day", "new year's eve", 'no-cook', 'nut', 'nutmeg', 'olive', 'orange', 'orzo', 'oscars', 'paleo', 'pan-fry', 'parade', 'parmesan', 'parsley', 'party', 'passover', 'pasta', 'pastry', 'peach', 'peanut', 'pear', 'pecan', 'pepper', 'picnic', 'pie', 'pineapple', 'pistachio', 'poach', 'pork', 'pork tenderloin', 'potato', 'prosciutto', 'radicchio', 'radish', 'raisin', 'ramadan', 'ramekin', 'raspberry', 'red wine', 'rhode island', 'rice', 'ricotta', 'roast', 'root vegetable', 'rosemary', 'rum', 'rutabaga', 'saffron', 'salad', 'salad dressing', 'salmon', 'sandwich', 'sandwich theory', 'sauce', 'sausage', 'seafood', 'seattle', 'seed', 'semolina', 'sesame', 'shallot', 'sherry', 'shower', 'shrimp', 'simmer', 'soufflé/meringue', 'soup/stew', 'sour cream', 'soy', 'soy sauce', 'spain', 'spice', 'spinach', 'spirit', 'spring', 'stir-fry', 'strawberry', 'stuffing/dressing', 'sugar snap pea', 'super bowl', 'sweet potato/yam', 'swiss cheese', 'tarragon', 'tea', 'thanksgiving', 'tomatillo', 'tortillas', 'tuna', "valentine's day", 'vanilla', 'veal', 'vinegar', 'walnut', 'washington', 'wedding', 'whiskey', 'white wine', 'yogurt', 'zucchini', 'turkey', 'bon appétit', 'peanut free', 'soy free', 'gourmet', 'tree nut free', 'vegetarian', 'kosher', 'pescatarian', 'quick & easy', 'wheat/gluten-free', 'bake', 'summer', 'dessert', 'winter', 'side', 'fall', 'dairy free', 'no sugar added', 'kidney friendly', 'sugar conscious', 'dinner', 'onion', 'milk/cream', 'sauté', 'tomato', 'vegetable', 'fruit', 'vegan']
{0.0: 9678, 1.0: 950}
{0.0: 4752, 1.0: 484}
Test accuracy of egg is 0.9134835752482812
Test precision of egg is 0.5786802030456852
Test recall of egg is 0.23553719008264462
Test f1-score of egg is 0.3348017621145374
```

We see that while some (dessert, kosher) have good results, some have poor results. Let's take the use case of vegan which has fairly decent metrics. This time we will refine the number of features to be considered by choosing a lasso regulariser and considering a number of options for C. Since it takes time to run analysis, we run it with 10 values of C for now!

**Classification**

```
[13]:  target = 'vegan'

       categories.remove(target)
       X = recipes_full_records.loc[:, categories]
```

```
categories.append(target)
y = recipes_full_records.loc[:, recipes.columns== target]
xtr, xts, ytr, yts = getTestTrainData(X, y)
```

[14]:
```python
from sklearn.metrics import precision_recall_fscore_support
from sklearn.model_selection import StratifiedKFold
from sklearn.preprocessing import StandardScaler
from sklearn import linear_model
from sklearn.metrics import confusion_matrix
from sklearn.model_selection import KFold


# example of grid searching key hyperparametres for logistic regression
from sklearn.model_selection import RepeatedStratifiedKFold
from sklearn.model_selection import GridSearchCV
from sklearn.linear_model import LogisticRegression
# define dataset

model = LogisticRegression()

solvers = ['liblinear']
penalty = ['l1']
c_values = np.logspace(0.0001, 100, num= 20)
scoring = ['accuracy','precision', 'recall', 'f1']
# define grid search
grid = dict(solver=solvers,penalty=penalty,C=c_values)
cv = RepeatedStratifiedKFold(n_splits=10, n_repeats=3, random_state=1)
grid_search = GridSearchCV(estimator=model, param_grid=grid, n_jobs=-1, cv=cv,␣
 ↪scoring=scoring,error_score=0, refit='accuracy')
grid_result = grid_search.fit(xtr,ytr)
# summarize results
print("Best: %f using %s" % (grid_result.best_score_, grid_result.best_params_))
```

```
Best: 0.966221 using {'C': 1.0002302850208247, 'penalty': 'l1', 'solver':
'liblinear'}
```

[14]:

We use the best C we got from above to do our logistic regression now. We should see a fairly improved recall and f1 score for target 'vegan' now!

[15]:
```python
import sklearn.metrics as metrics
from sklearn.metrics import auc
logreg = linear_model.LogisticRegression(penalty= 'l1', C=grid_result.
 ↪best_params_['C'],solver='liblinear')
logreg.fit(xtr,ytr)
yhat = logreg.predict(xts)
```

```
accuracy,precision,recall,f1_score = getMetrics(yts, yhat)
print('Logistic Regression metrics')
print('Test accuracy of {} is {}'.format(target, accuracy))
print('Test precision of {} is {}'.format(target, precision))
print('Test recall of {} is {}'.format(target, recall))
print('Test f1-score of {} is {}'.format(target, f1_score))
```

```
Logistic Regression metrics
Test accuracy of vegan is 0.9698242933537051
Test precision of vegan is 0.8677494199535963
Test recall of vegan is 0.7873684210526316
Test f1-score of vegan is 0.82560706401766
```

[16]:
```
W = logreg.coef_
W = W.ravel()

predictors = []
wnonzero = []
for i, wval in enumerate(W):
    if(wval > 0):
        predictors.append(categories[i])
        wnonzero.append(wval)
#         print(categories[i] + " " + str(wval))

# print(set(predictors))
# print(set(wnonzero))
plt.figure(figsize = (15, 5))
plt.stem(W)

# print(len(categories))
```

```
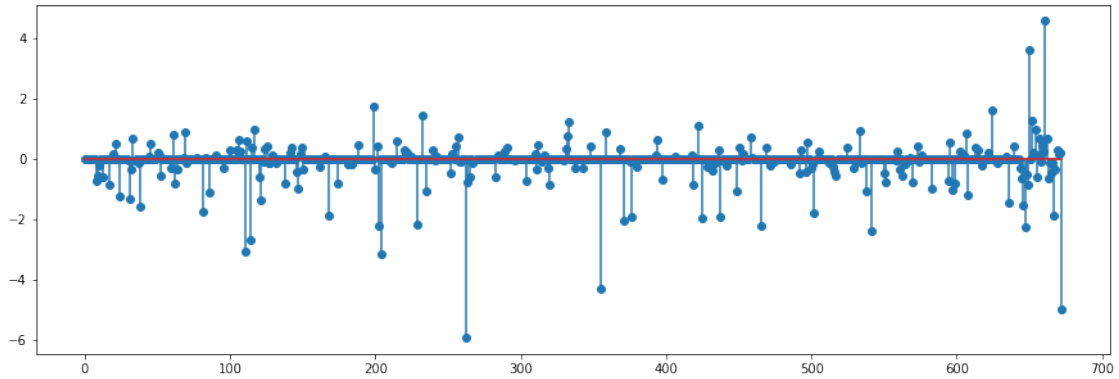<ipython-input-16-704a2cbfacb6>:15: UserWarning: In Matplotlib 3.3 individual
lines on a stem plot will be added as a LineCollection instead of individual
lines. This significantly improves the performance of a stem plot. To remove
this warning and switch to the new behaviour, set the "use_line_collection"
keyword argument to True.
  plt.stem(W)
```

[16]: `<StemContainer object of 3 artists>`

```
[17]: index = np.argsort(np.abs(wnonzero))

top_predictors = []
top_wval = []
for i in range(10):
    top_predictors.append(predictors[index[-i-1]])
    top_wval.append(wnonzero[index[-i-1]])

#Place your code here
import matplotlib.pyplot as plt

fig = plt.figure(figsize = (15, 5))
plt.bar(top_predictors, top_wval,width = 0.5)
plt.xlabel("Feature names")
plt.ylabel("Importance")
plt.show()

yproba_lr= logreg.predict_proba(xts)
fpr_lr, tpr_lr, _ = metrics.roc_curve(yts, yproba_lr[:, 1])
```
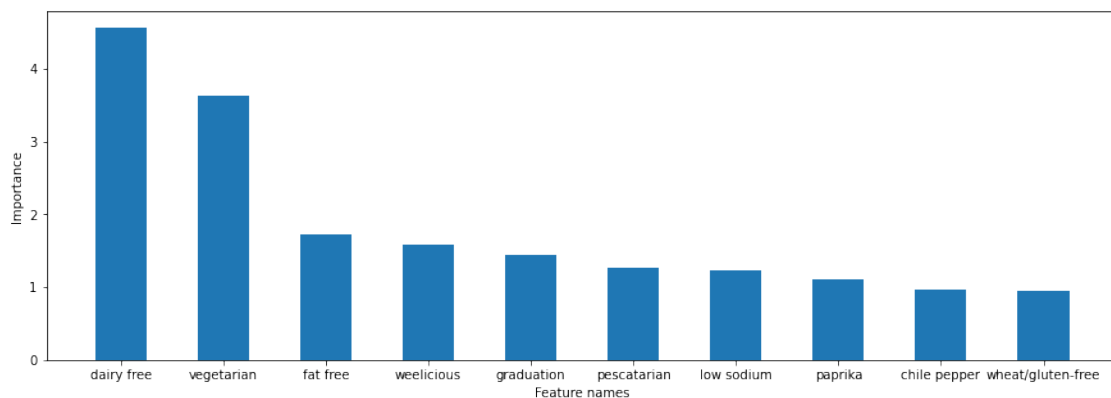
```
[18]: import seaborn as sns
      corr = recipes_full_records.loc[:, top_predictors].corr(method ='pearson')
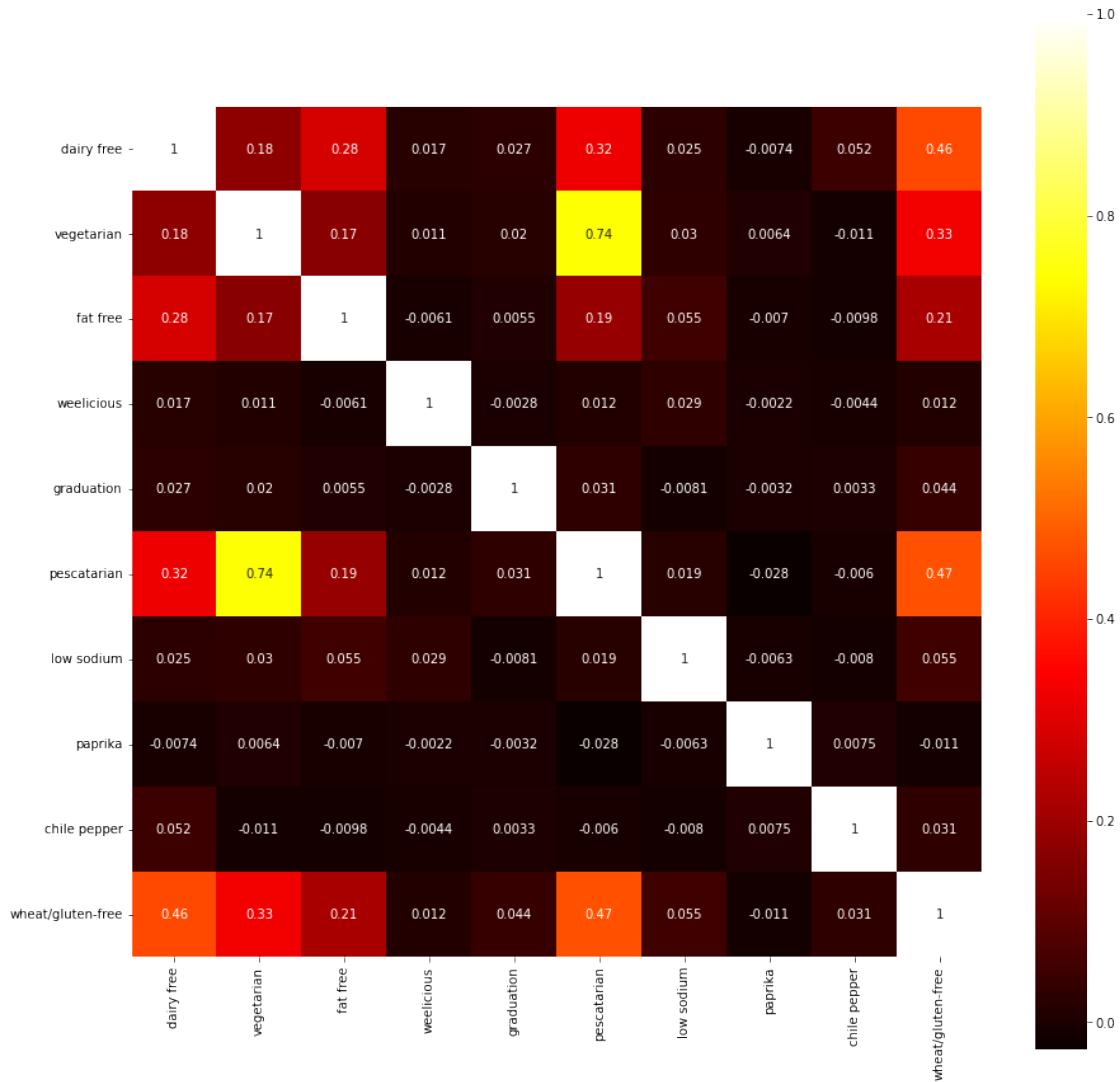      plt.figure(figsize=(15, 15))
      sns.heatmap(corr, vmax=1, square=True,annot=True,cmap='hot')
```

[18]: <matplotlib.axes._subplots.AxesSubplot at 0x7f22858141c0>

	dairy free	vegetarian	fat free	weelicious	graduation	pescatarian	low sodium	paprika	chile pepper	wheat/gluten-free
dairy free	1	0.18	0.28	0.017	0.027	0.32	0.025	-0.0074	0.052	0.46
vegetarian	0.18	1	0.17	0.011	0.02	0.74	0.03	0.0064	-0.011	0.33
fat free	0.28	0.17	1	-0.0061	0.0055	0.19	0.055	-0.007	-0.0098	0.21
weelicious	0.017	0.011	-0.0061	1	-0.0028	0.012	0.029	-0.0022	-0.0044	0.012
graduation	0.027	0.02	0.0055	-0.0028	1	0.031	-0.0081	-0.0032	0.0033	0.044
pescatarian	0.32	0.74	0.19	0.012	0.031	1	0.019	-0.028	-0.006	0.47
low sodium	0.025	0.03	0.055	0.029	-0.0081	0.019	1	-0.0063	-0.008	0.055
paprika	-0.0074	0.0064	-0.007	-0.0022	-0.0032	-0.028	-0.0063	1	0.0075	-0.011
chile pepper	0.052	-0.011	-0.0098	-0.0044	0.0033	-0.006	-0.008	0.0075	1	0.031
wheat/gluten-free	0.46	0.33	0.21	0.012	0.044	0.47	0.055	-0.011	0.031	1

```
[19]: from sklearn.ensemble import RandomForestClassifier

      model = RandomForestClassifier()
      target = 'vegan'
      categories.remove(target)
      X = recipes_full_records.loc[:, categories]
      categories.append(target)
```

```
y = recipes_full_records.loc[:, recipes.columns== target]
xtr, xts, ytr, yts = getTestTrainData(X, y)
n_estimators =  [50, 200]
max_features = ['auto', 'sqrt', 'log2']
scoring = ['accuracy','precision', 'recall', 'f1']
# define grid search
grid = dict(n_estimators=n_estimators,max_features=max_features)
cv = RepeatedStratifiedKFold(n_splits=10, n_repeats=3, random_state=1)
grid_search = GridSearchCV(estimator=model, param_grid=grid, n_jobs=-1, cv=cv,
 ↪scoring=scoring,error_score=0, refit='accuracy')
grid_result = grid_search.fit(xtr,ytr)
print("Best: %f using %s" % (grid_result.best_score_, grid_result.best_params_))
```

Best: 0.971458 using {'max_features': 'auto', 'n_estimators': 200}

```
[20]: def getPredictors(target):
          rf_classifier = RandomForestClassifier(criterion="entropy", n_estimators=
       ↪grid_result.best_params_['n_estimators'], max_features=grid_result.
       ↪best_params_['max_features'])
          categories.remove(target)
          X = recipes_full_records.loc[:, categories]
          y = recipes_full_records.loc[:, recipes.columns== target]
          xtr, xts, ytr, yts = getTestTrainData(X, y)
          rf_classifier.fit(xtr, ytr)
          predictors = []
          for i, column in enumerate(categories):
              if(rf_classifier.feature_importances_[i] > 0):
                  predictors.append(column)
          categories.append(target)
          return predictors, rf_classifier

      target = 'vegan'
      predictors, rf_classifier = getPredictors(target)
      X = recipes_full_records.loc[:, predictors]
      y = recipes_full_records.loc[:, recipes.columns== target]

      xtr, xts, ytr, yts = getTestTrainData(X, y)

      rf_classifier.fit(xtr, ytr)

      # calculating test accuracy
      yhat = rf_classifier.predict(xts)

      accuracy,precision,recall,f1_score = getMetrics(yts, yhat)
      print('Random Forest metrics')
      print('Test accuracy of {} is {}'.format(target, accuracy))
      print('Test precision of {} is {}'.format(target, precision))
```

```
print('Test recall of {} is {}'.format(target, recall))
print('Test f1-score of {} is {}'.format(target, f1_score))

print(set(predictors))
yproba_rf= rf_classifier.predict_proba(xts)
fpr_rf, tpr_rf, _ = metrics.roc_curve(yts, yproba_rf[:, 1])
```

Random Forest metrics
Test accuracy of vegan is 0.9713521772345302
Test precision of vegan is 0.9140625
Test recall of vegan is 0.75
Test f1-score of vegan is 0.823943661971831
{'atlanta', 'cucumber', 'virginia', 'ground lamb', 'deep-fry', 'breakfast',
'wedding', 'noodle', 'los angeles', 'spinach', 'coffee', 'halloween',
'friendsgiving', 'new orleans', 'frozen dessert', 'cognac/armagnac', 'lasagna',
'grape', 'avocado', 'hot pepper', 'tapioca', 'ramekin', 'cupcake', 'cumin',
'kid-friendly', 'pistachio', 'poker/game night', 'taco', 'egg', 'connecticut',
'mint', 'radish', 'wheat/gluten-free', 'mozzarella', 'raw', 'tamarind',
'parsley', 'hamburger', 'butter', 'rice', 'turnip', 'sparkling wine', 'cookbook
critic', 'peanut', 'costa mesa', 'australia', 'graduation', 'curry', 'lentil',
'suzanne goin', 'milk/cream', 'vegetarian', 'non-alcoholic', 'pittsburgh',
'lemongrass', 'cranberry sauce', 'custard', 'ham', 'freeze/chill', 'egg nog',
'pomegranate', 'kosher', 'clam', 'beef', 'goat cheese', 'maple syrup',
'saffron', 'mandoline', 'kwanzaa', 'lamb', 'chili', 'one-pot meal', 'snack
week', 'liqueur', 'lobster', 'campari', 'harpercollins', 'sauté', '#wasteless',
'lemon', 'quick and healthy', 'chicken', 'beet', 'utah', 'low fat', 'sage',
'root vegetable', 'pasta maker', 'macaroni and cheese', 'sorbet', 'butternut
squash', 'calvados', 'plantain', 'squid', 'cod', 'pork tenderloin', 'scotch',
'horseradish', 'miami', 'ginger', 'israel', 'mussel', 'seafood', 'self', 'low/no
sugar', 'cardamom', 'boil', 'pineapple', 'purim', 'sweet potato/yam', 'fall',
'asian pear', 'rye', 'canada', 'onion', 'pomegranate juice', 'orange juice',
'roast', 'rosh hashanah/yom kippur', 'christmas eve', 'plum', 'cake', 'tree
nut', 'spirit', 'broil', 'nancy silverton', 'no meat, no problem', 'shower',
'birthday', 'leafy green', 'corn', 'potato salad', 'london', 'cantaloupe',
'sauce', 'semolina', 'champagne', 'pressure cooker', 'white wine', 'citrus',
'salsa', "st. patrick's day", 'nebraska', 'blue cheese', 'garlic', 'nut', 'whole
wheat', 'pennsylvania', 'wine', 'cornmeal', 'bok choy', 'bon app tit', 'gin',
'bread', 'chill', 'yonkers', 'turkey', 'bass', 'washington, d.c.', 'wild rice',
'granola', 'poach', 'switzerland', 'anthony bourdain', 'duck', 'watermelon',
'missouri', 'no-cook', 'illinois', 'yellow squash', 'cocktail party', "father's
day", 'georgia', 'gourmet', 'side', 'low sugar', 'leek', 'bulgur', 'fish',
'parsnip', 'zucchini', 'kiwi', 'bourbon', 'chive', 'caraway', 'cookbooks',
'barley', 'oregon', 'pork rib', 'meatball', "valentine's day", 'ground beef',
'mayonnaise', 'north carolina', 'kentucky derby', 'ireland', 'chard',
'appetizer', 'family reunion', 'louisiana', 'hollywood', 'soup/stew', 'apricot',
'tangerine', 'port', 'dairy free', 'walnut', 'brie', 'hot drink', 'buffet',
'maryland', 'grill/barbecue', 'capers', 'coriander', 'labor day', 'cherry',
```

'lime', 'pickles', 'house & garden', 'sugar conscious', 'sour cream', 'blender',
'broccoli', 'crêpe', 'banana', 'jam or jelly', 'cookies', 'lychee', 'low
sodium', 'peach', 'ricotta', 'oat', 'iced coffee', 'columbus', 'bastille day',
'kansas city', 'mixer', 'snapper', 'alaska', 'party', 'pacific palisades',
'soufflé/meringue', 'steam', 'goose', "new year's eve", 'crab', 'mango',
'3-ingredient recipes', 'endive', 'new jersey', 'engagement party', 'halibut',
'sesame', 'quick & easy', 'tortillas', 'kidney friendly', '#cakeweek',
'cocktail', 'winter', 'marinate', 'guava', 'sandwich theory', 'tart',
'digestif', 'lemon juice', 'shallot', 'yogurt', 'candy', 'ohio', 'cranberry',
'edible gift', 'sardine', 'cinnamon', 'steak', 'shavuot', 'vermont', 'quiche',
'sausage', '30 days of groceries', 'low cal', 'healthy', 'lima bean', 'paleo',
'cheese', 'chartreuse', 'hummus', 'red wine', 'bacon', 'freezer food',
'hanukkah', 'rhubarb', "new year's day", 'chickpea', 'washington',
'breadcrumbs', 'green bean', 'juicer', 'okra', 'pepper', 'ramadan', 'san
francisco', 'phyllo/puff pastry dough', 'spring', 'advance prep required',
'escarole', 'salad', "hors d'oeuvre", 'tested & improved', 'lunch',
'entertaining', 'rub', 'couscous', 'artichoke', 'collard greens', 'tomato',
'grill', 'back to school', 'pizza', 'tofu', 'almond', 'casserole/gratin',
'arugula', 'sake', 'pernod', 'brussel sprout', 'pork chop', 'no sugar added',
'drink', 'kumquat', 'parmesan', 'tuna', 'chocolate', 'brooklyn', 'south
carolina', 'cauliflower', 'meat', 'house cocktail', 'macadamia nut',
'shellfish', 'melon', 'condiment', 'arizona', 'bitters', 'fat free', 'raisin',
'honeydew', 'swiss cheese', 'santa monica', 'pine nut', 'pasta', 'florida',
'asparagus', 'yuca', 'snack', 'jalapeño', 'radicchio', 'marscarpone', 'carrot',
'ice cream machine', 'summer', 'microwave', 'chile', 'bell pepper',
'philippines', 'sandwich', 'nectarine', 'dill', 'dairy', 'coffee grinder', 'new
york', 'beef shank', 'orzo', 'tropical fruit', 'scallop', 'brown rice', 'clove',
'fig', 'anise', 'octopus', 'pear', 'quail', 'candy thermometer', 'thyme',
'poultry sausage', 'frangelico', 'punch', 'soy sauce', 'papaya', 'lime juice',
'currant', 'marsala', 'oscars', 'raspberry', 'sherry', 'herb', 'mezcal',
'legume', 'pumpkin', 'cashew', 'massachusetts', 'salad dressing', 'celery',
'lamb shank', 'orange', 'sangria', 'epi loves the microwave', 'margarita',
'smoothie', 'prosciutto', 'fortified wine', 'new mexico', 'kosher for passover',
'indiana', 'minnesota', 'braise', 'lettuce', 'colorado', 'italy', 'martini',
'fontina', 'smoker', 'apple juice', 'beef rib', 'chestnut', 'bake', 'grains',
'oregano', 'michigan', 'sukkot', 'backyard bbq', 'cuba', 'kirsch', 'mustard
greens', 'food processor', 'vanilla', 'pot pie', '22-minute meals', 'passion
fruit', 'passover', 'tailgating', 'rack of lamb', 'alcoholic', 'feta',
'oatmeal', 'hazelnut', 'brandy', 'pan-fry', 'iced tea', 'skewer', 'fry', 'lunar
new year', 'parade', 'marinade', 'peanut free', 'super bowl', 'whiskey',
'olive', 'hominy/cornmeal/masa', 'potato', 'diwali', 'sesame oil',
'butterscotch/caramel', 'blueberry', 'rum', 'mexico', 'créme de cacao', 'poppy',
'watercress', 'dominican republic', 'peanut butter', 'biscuit', 'strawberry',
'squash', 'potluck', 'beer', 'wasabi', 'tea', 'bean', 'coconut', 'cinco de
mayo', 'pie', 'thanksgiving', 'fruit', 'oktoberfest', 'pasadena', 'fruit juice',
'fritter', 'vermouth', 'aperitif', 'midori', 'tomatillo', 'cilantro',
'portland', 'low carb', 'west virginia', 'game', 'quinoa', 'texas', 'pecan',
'gouda', 'rosemary', 'persimmon', 'swordfish', 'vodka', 'dip', 'pork',

'providence', 'cobbler/crumble', 'bon appétit', 'sugar snap pea', 'brisket',
'poultry', 'dried fruit', 'cookie', 'christmas', 'high fiber', 'veal',
'california', "mother's day", 'kahlúa', 'beef tenderloin', 'stuffing/dressing',
'pescatarian', 'eau de vie', 'anniversary', 'stew', 'tequila', 'pastry',
'kentucky', 'lamb chop', 'organic', 'shrimp', 'chambord', 'seed', 'salmon',
'picnic', 'brunch', 'cabbage', 'flaming hot summer', 'triple sec', 'fennel',
'soy', 'spice', 'nutmeg', 'maine', 'rhode island', 'cheddar', 'paprika',
'berry', 'stir-fry', 'easter', 'dinner', 'cream cheese', 'mushroom', 'tree nut
free', 'chile pepper', 'apple', 'brine', 'dallas', 'double boiler', 'date',
'kale', 'honey', 'mustard', 'jícama', 'soy free', 'vinegar', 'anchovy', 'pea',
'dessert', 'molasses', 'ice cream', 'quince', 'tilapia', 'fourth of july',
'chicago', 'france', 'cottage cheese', 'basil', 'trout', 'green onion/scallion',
'low cholesterol', 'drinks', 'eggplant', 'broccoli rabe', 'paris', 'blackberry',
'muffin', 'grapefruit', 'wok', 'mardi gras', 'seattle', 'tarragon', 'venison',
'buttermilk', 'slow cooker', 'marshmallow', 'jerusalem artichoke', 'vegetable',
'amaretto', 'condiment/spread', 'prune', 'simmer', 'monterey jack', 'alabama',
'oyster', 'persian new year', 'weelicious', 'rutabaga', 'pancake'}

```python
from sklearn.neighbors import KNeighborsClassifier

k_range = list(range(1, 10))
grid = dict(n_neighbors=k_range)

model = KNeighborsClassifier()
target = 'vegan'
categories.remove(target)
X = recipes_full_records.loc[:, categories]
categories.append(target)
y = recipes_full_records.loc[:, recipes.columns== target]
xtr, xts, ytr, yts = getTestTrainData(X, y)

cv = RepeatedStratifiedKFold(n_splits=10, n_repeats=3, random_state=1)
grid_search = GridSearchCV(estimator=model, param_grid=grid, n_jobs=-1, cv=cv,
  scoring=scoring,error_score=0, refit='accuracy')
grid_result = grid_search.fit(xtr, ytr)
print("Best: %f using %s" % (grid_result.best_score_, grid_result.best_params_))
```

```
/usr/local/lib/python3.8/dist-
packages/joblib/externals/loky/process_executor.py:700: UserWarning: A worker
stopped while some jobs were given to the executor. This can be caused by a too
short worker timeout or by a memory leak.
  warnings.warn(

Best: 0.953958 using {'n_neighbors': 9}
```

```python
knn = KNeighborsClassifier(n_neighbors = grid_result.
  best_params_['n_neighbors'])
knn.fit(xtr, ytr)
```

```
yhat = knn.predict(xts)
accuracy,precision,recall,f1_score = getMetrics(yts, yhat)
print('KNeighbors Classifier metrics')
print('Test accuracy of {} is {}'.format(target, accuracy))
print('Test precision of {} is {}'.format(target, precision))
print('Test recall of {} is {}'.format(target, recall))
print('Test f1-score of {} is {}'.format(target, f1_score))
yproba_knn= knn.predict_proba(xts)
fpr_knn, tpr_knn, _ = metrics.roc_curve(yts, yproba_knn[:, 1])
categories.append(target)
```

```
KNeighbors Classifier metrics
Test accuracy of vegan is 0.9553093964858671
Test precision of vegan is 0.8495821727019499
Test recall of vegan is 0.6288659793814433
Test f1-score of vegan is 0.7227488151658769
```

Graph showing the comparison

[23]:
```
logreg, rf_classifier, knn

roc_auc_knn = auc(fpr_knn, tpr_knn)
roc_auc_lr = auc(fpr_lr, tpr_lr)
roc_auc_rf = auc(fpr_rf, tpr_rf)
plt.rcParams["figure.figsize"] = (10,10)

plt.plot(fpr_knn,tpr_knn,label='AUC KNN = %0.2f' % roc_auc_knn)
plt.plot(fpr_lr,tpr_lr, label='AUC Logistic Regression = %0.2f' % roc_auc_lr)
plt.plot(fpr_rf,tpr_rf,label='AUC Random Forest = %0.2f' % roc_auc_rf)

plt.legend(loc="lower right")
plt.show()
```

AUC KNN = 0.94
AUC Logistic Regression = 0.98
AUC Random Forest = 0.99