What makes your recipe successful ?

Omar Mehio Leonardo Perrone Attila Bekker
EPFL

Motivation

- Millions of people check recipe websites everyday. Either to look-up something delicious they ate or to try a new recipe out. At the end people submit to the website if they were able to reproduce the recipe and how well it tasted.
- Usually when we search for recipes online we search for recipes rated highly by the crowd as it implies that most people enjoyed the meal.
- In this project we aim in helping all chefs in maximizing the number of good reviews for their recipes by exploring all factors that contribute in having a good rating.

Data Collection and Cleaning

- Dataset consits of dump of webiste links having recipes.
- These links can have recipes or search results.
- Built scrapers for top 3 websites as they contain 65 % of the data.

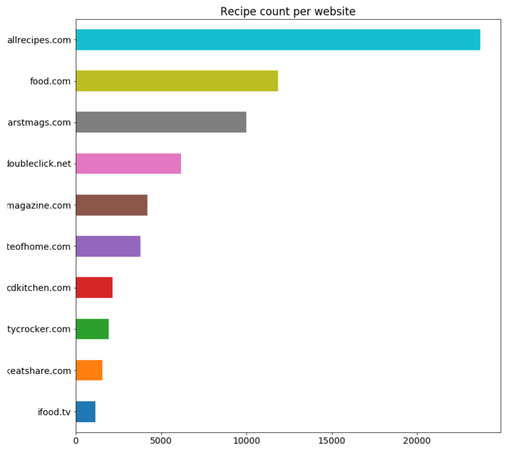

Figure 1:Number of Links per website

Datasets

Columns	AllRecipe	Food	FoodNet
Name	✓	✓	✓
Ingredients	✓	✓	✓
Preparation Time	✓	✓	✓
Cooking Time	✓	✓	✓
Servings	✓	✓	✓
Nutrition	✓	✓	✗
Review	✓	✓	✓
Rating	✓	✗	✗

Unsupervised Learning to the Rescue

- Aim is to unify the datasets into one global dataset.
- Word2vec to map ingredients into vectors that were clustered into food type clusters using k-means (silhoutte analysis to find the best k).
- Logistic regression trained on reviews of allrecipe.com and used to predict ratings from food.com revies.

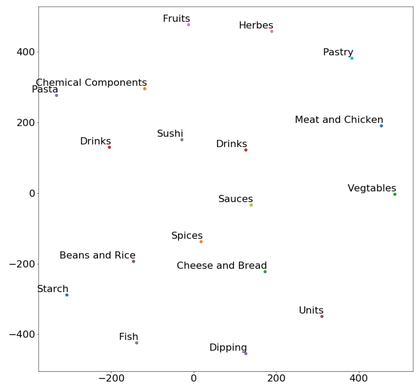

Figure 2:AllRecipe.com clusters

Data Analysis

General scheme:

1. compute average of ratings per recipe
2. bucketize dependent variable if necessary
3. perform a count of rating value per dependent variable unique values

How long should your meal be cooked ?

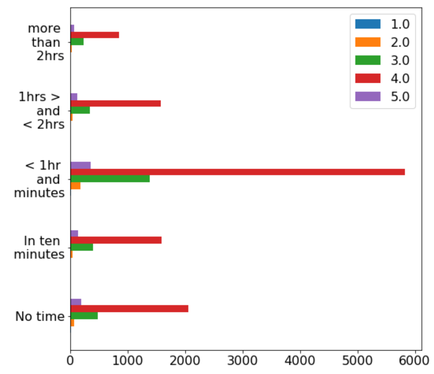

Figure 3:Rating counts per cooking time bucket

How many people should you serve?

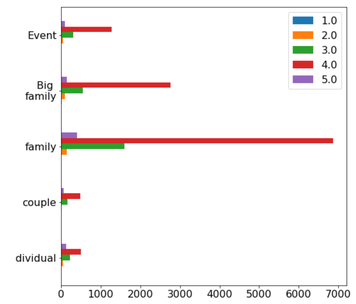

Figure 4:Rating counts per number of servings bucket

Health Scores

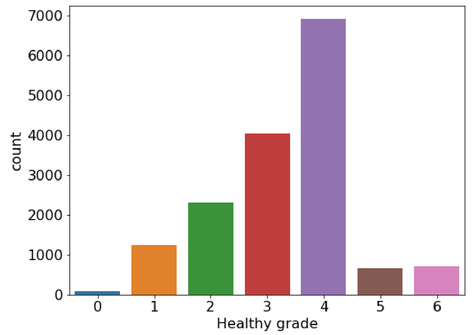

Figure 5:Count of Recipes per Health grade

What is the best ingredient composition ?

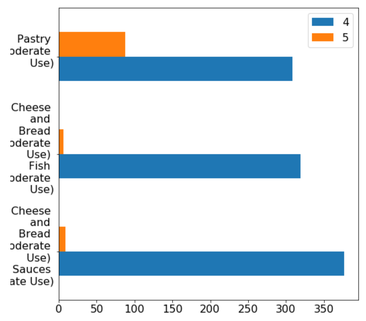

Figure 6:Ratings counts per composition

Our Advice to you

1. Design a meal to be cooked and prepared in less than an hour
2. Try to serve as much people as you can
3. Try to combine a moderate amount cheeses with bread along with some sauces
4. low levels of fats and cholesterol is negativley correlated with large numbers of servings