

# Instructions for the Final Project

L645; Sandra Kuebler

- You will work on a data set consisting of cooking recipes from the epicurious.com. The files contain the recipes, but also meta information, reviews (if available), and an overall, average rating based on the individual ratings. Your task will be to construct a machine learning system that will predict this overall rating, in terms of 1–4 forks.
- The data set is available on canvas site under the final project, along with a Readme, which explains the structure of the files. Note that not all files are relevant for every task.
- You should choose one task from the following list and work in groups of 2-3 people. Be prepared to let me know by Thursday, October 21 who will be in your group and which task you chose.

## 1. Adapting feature selection methods for multi-class classification tasks

In the cooking recipe data, we face a 4-class classification problem. Commonly used distribution-based feature selection methods (such as information gain, bi-normal separation) are defined for binary classification, for example we could group 1-fork and 2-fork as “bad recipes” and 3-fork and 4-fork as “good recipes”, and then select the features. Can you come up with a way to adapt these methods to this project? Are the selected features more helpful? Possible approaches include a stacked classification into good vs. bad and then a more specific classifier. Or a one-vs.-all strategy.

## 2. Grouping recipes by categories towards better in-category classification performance

Recipes on Epicurious come in many types, they can be for desserts, soups, main courses, etc. Our judgement towards soup may be quite different from the judgment of dessert or steak. For example we would love to have black pepper on a steak, but not in a dessert; we would love a crispy chocolate roll but not a crispy steak. Thus it may not be a good idea to compare across categories. In this question, we can explore grouping recipes automatically using topic modeling techniques (open to ingredients and instructions), how good is the grouping? And does the group help with in-category classification?

## 3. Modeling the topics of reviews and include them as features for rating prediction

As a baseline, the features for rating prediction are  $n$ -grams in the reviews. These  $n$ -grams tend to capture the language patterns of positive and negative comments, but may not capture more abstract meanings. For example, is there a group of words that describes the difficulty of recipe preparation? Is there a group of words that describe how much people enjoyed the dish? In this question, you will explore using the outcome of topic modeling as features, do they help with recipe prediction?

- For the ML part, it may make sense to run the experiments on an IU cluster (quarry, bigred2).
- Start early with the tasks so that you can ask questions when you get stuck. I do not expect you to investigate your domain in all details, but I expect you to have two sets of results (based on two different decisions) that you can discuss.
- Your final task is to write up your results. The final paper should be between 8 and 10 pages ACL style formatting. Style files are available from <http://acl2014.org/CallforPapers.htm>, under the section “Format”.

In the paper, you need to start by explaining the problem, then describe your approach, present an evaluation, and discuss your results.

**If you have any questions or run into any problems that you cannot answer within your group in a reasonable amount of time, talk to me!!!**

- **DUE DATE:** Wednesday, December 17.