Homebrewed beer recipes

Junjie Cai 43937780 - project proposal ITEC873

Introduction

This project is related to homebrewed beer recipes which are provided by users on the site Brewer's Friend. These recipes are uploaded by brewers who are willing to share their private beer recipes. My work is going to analyse these interesting data.

Goals and methods

For this project, it is interesting to classify beer style of 176 types according to beer details. Further, as we could guess, the color of the beer might have some correlation with other features, so I also want to find its exact relationship between the color and other features. In general, there are two main goals:

- *Can we predict the style of the beer according to given beer conditions?*
- *Can we predict weather the beer is "heavy" or "light" by given beer conditions?*

For the first question, **KNN** will be used to try and predict the style of beer based on beer conditions.
For the second one, firstly, **linear regression** will be used to build the model between the Color and other features by applying one best predictor and/or multiple predictors. Analysis will be focused on the accuracy of the model; after that, **Logistic regression** will be helpful in determining the probability of a predictor(ABV) OR multiple predictors (ABV, OG, FG, IBU) affecting the color become "heavy".

Dataset

The dataset used in this project contains over 75,000 homebrewed beers with over 176 different styles, such as Altbier, American API, Australian Sparkling Ale, Apple Wine, and so on. These beer recipes are user-reported and are categorised according to one of the 176 different styles.
Further, each recipe goes into as much or as little details as the user provided, but there are at least 5 useful features which data was provided for each: Original Gravity, Final Gravity, ABV, IBU, and Color. Other features are only provided for a part of recipes, such as Priming Method and Priming Amount, which seems like it just let the user write whatever they wanted. Thus, 5 most useful beer conditions might be used in this project, which are:

- Original gravity: the original percentage of beer wort
- Final gravity: the final percentage of beer wort
- ABV: alcohol By Volume
- IBU: International bitterness units
- COLOR: numeric levels of the color