Final Project Proposal

Guojing Wu / UNI: gw2383

11/12/2019

Today, more and more people choose to eat outside. They think cooking at home is inconvenient and they usually have no idea of what's to cook. Eating outside somehow means more cuisines and more choices, and that's annoying for indecisive person. So, the objective for this project is to make eating out easier by recommending people with recipes of various cuisine type. In this project, we trying to analyze recipes for different types of cuisine to predict a best-matched cuisine type based on the input ingredients, or based on what they have at home.

The data is obtained from Kaggle, it contains over 12,000 recipes and 20 cuisine types. It contains three columns:

- id: unique recipes ID for each recipe
- cuisine: 20 different cuisine types: Brazilian, British, cajun_creole, Chinese, Filipino, French, Greek, Indian, Irish, Italian, Jamaican, Japanese, Korean, Mexican, Moroccan, Russian, Southern_us, Spanish, Thai, and Vietnamese
- ingredients: each recipe has its unique ingredients vector, the length varies from 1 to 110.

In an attempt to do such recommendation, we first want to perfrom exploratary data analysis, see what's the distribution for ingredients and what's the feature for each cuisine type. Then we want to perform several forms of analysis including but not limited to feature extraction (like clustering, network analysis) and multiple classification methods (like Support Vector Machine, regression, decision tree, etc).

The proposed timeline for final project:

Time	progress
Nov, 15th ~ Nov, 19th	EDA
Nov, 20th ~ Nov, 25th	feature extraction
Nov, 26th ~ Nov, 30th	multiple classification
Dec, 1st ~ Dec, 4th	result output
Dec, 5th ~ Dec, 13rd	report writing