Data Mining Project: What's cooking

Grigor Keropyan[1]

[1]Department of Mathematics and Mechanics
Yerevan State University

6 June 2020

About Dataset

```
train_json[0]
```

```
{'id': 10259,
 'cuisine': 'greek',
 'ingredients': ['romaine lettuce',
  'black olives',
  'grape tomatoes',
  'garlic',
  'pepper',
  'purple onion',
  'seasoning',
  'garbanzo beans',
  'feta cheese crumbles']}
```

Figure: What's cooking dataset

Problem Definition

Based on the the ingredients data the model should predict the cuisine
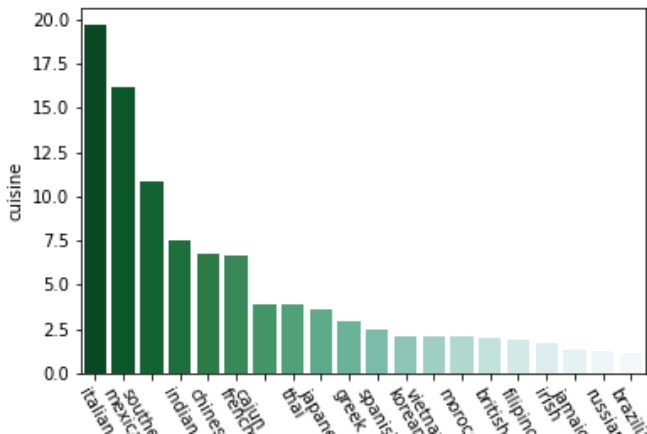
About Dataset

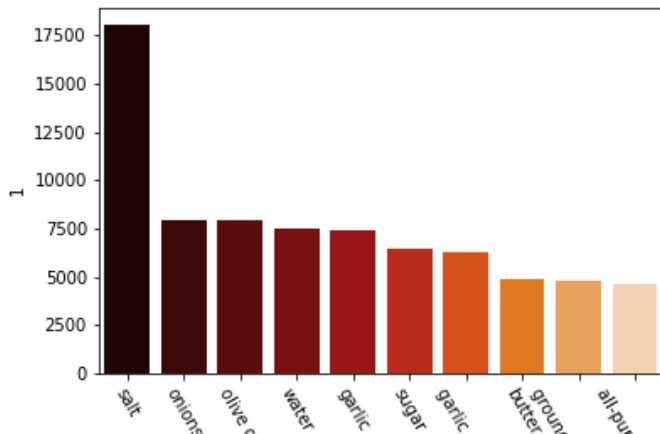Figure: BarPlot of cuisines

About Dataset

Figure: Most popular ingredients

Models

KNeighborsClassifier
SVC with linear kernel
SVC RBF
DecisionTreeClassifier
RandomForestClassifier
MLPClassifier
AdaBoostClassifier
GaussianNB

The Best Model

SVC RBF
With accuracy of 0.74 On the test set also accuracy is 0.74

Thank you !