### Master Chef: Cuisine Classification and Recipe Generation

**Juhi Naik, Vinaya Polamreddi**

#### Classification

**Data**

- Distribution of Data
  - SouthEastAsian
  - WestEuropean
  - SouthAsian
  - EasternEuropean
  - Japanese
  - LatinAmerican
  - NorthAmerican
  - EasternAsian
  - African
  - EastAsian

**~6500 - Kaggle competition**

**Datasets:**
- Dataset 1: Full dataset divided 70:30
- Dataset 2: 70:30 after removing North American recipes
- Dataset 3: 2200 training examples and 550 test examples randomly sampled
- Dataset 4: 200 training examples and 50 test examples each taken from each of the 11 cuisines

**Results**

**Comparison of train errors based on datasets**

**Comparison of test errors based on datasets**

**Accuracy by Cuisine**

**Similarity between predictions of models**

**Conclusion**

- Most methods have similar performance on our data
- Stratifying the data to have a uniform data set decreased our performance the most
- Most methods were able to classify ~70% of our data correctly. ~7-11% of our data wasn't able to classified accurately by any our methods.

#### Generation

**Data**

- ~63000 recipes : scraped from AllRecipes.com

**Models for Generation**

- Each Recipe represented by a list of ingredients and instructions.
- Each Action represented as a (Verb, Ingrid) pair.

**Generate a set of ingredients**

**Examples:**
1. maple syrup
2. honey
3. cinnamon
4. banana
5. sugar
6. almond

**Generate a sequence of instructions given a set of ingredients**

**Example:**
- 1. preheat butter
- 2. mix cinnamon
- 3. mix salt
- 4. mix garlic
- 5. stir onion
- 6. stir potato
- 7. cook corn syrup

**Conclusions**

- Most methods have similar performance on our data
- Stratifying the data to have a uniform data set decreased our performance the most
- Most methods were able to classify ~70% of our data correctly. ~7-11% of our data wasn't able to classified accurately by any our methods.