

Problem Statement

Use recipe ingredients to categorize the cuisine

Picture yourself strolling through your local, open-air market... What do you see? What do you smell? What will you make for dinner tonight?

If you're in Northern California, you'll be walking past the inevitable bushels of leafy greens, spiked with dark purple kale and the bright pinks and yellows of chard. Across the world in South Korea, mounds of bright red kimchi greet you, while the smell of the sea draws your attention to squids squirming nearby. India's market is perhaps the most colorful, awash in the rich hues and aromas of dozens of spices: turmeric, star anise, poppy seeds, and garam masala as far as the eye can see.

Some of our strongest geographic and cultural associations are tied to a region's local foods. ***This playground competitions asks you to predict the category of a dish's cuisine given a list of its ingredients.***

There are three attachments with the mail- **sample submission.csv**, **test.json** and **train.json**. These attachments contains the required dataset for the event.

In the dataset, we include the recipe id, the type of cuisine, and the list of ingredients of each recipe (of variable length). The data is stored in JSON format.

File descriptions

- **train.json** - the training set containing recipes id, type of cuisine, and list of ingredients
- **test.json** - the test set containing recipes id, and list of ingredients. In the test file test.json, the format of a recipe is the same as train.json, only the cuisine type is removed, as it is the target variable you are going to predict.
- **sample_submission.csv** - a sample submission file in the correct format

*At the end of the contest mail us your submission file to bigdatachallenge.2015techsoc@gmail.com. Deadline for the submission of the solutions is **8th Oct, 2015**.*