

CS 839 Stage 1 Report

Shantanu Singhal (singhal5@wisc.edu)

Sukanya Venkataraman (venkatarama5@wisc.edu)

Yogesh Chockalingam (ychockalinga@wisc.edu)

Entity Type

- We are identifying food entities the Yelp reviews dataset. Food entity is defined as any noun that can occur in a food menu excluding generic terms like food, water, lunch etc.
- Each food entity is annotated in its entirety using XML tags `</>`
 - I ordered `<pizza/>` and `<kale juice/>`

Training and Test Set Information

We labeled 364 documents with 1978 mentions of food entities. Then, we shuffled the documents and split them as follows:

	Number of Documents	Number of Mentions
Set I	243	1402
Set J	121	576

Model Performance

To do

Rule Based Processing

To do