Project Proposal: A Food Categorizer for Advertisement Placement and Meal Planning

Scott Kroeger

Domain: Almost everyone enjoys a good meal, and some people even call themselves "foodies" which the Merriam Webster Dictionary defines as: a person having an avid interest in the latest food fads. Foodies often enjoy prepared and definitely enjoy eating a variety of rich and delicious foods. Pictures of the foods that foodies enjoy are sometimes referred to as "Food Porn" reddit has a forum dedicated to this topic.

This project will use Reddit's API to query and obtain posts to the https://www.reddit.com/r/FoodPorn/ I will perform textual analysis to determine the similarity of items posted to the the forum, and use clustering techniques to determine the major categories of food posts. This information could be used to build a recommender system that could be used for meal planning, and also as a way to determine which adds to place on the reddit forum or recipes sites.

An MVP will use posts from the subreddit forum and an additional provided phrase to determine cosine similarity between the subreddit posts and the provided phrase.

Data			
	Type	**Description**	**Use for Model**
Vectorized phrases from the subreddit	Sparse vectors	An array consisting of word counts of phrases	Y
Vectorized phrase provided by the user	Spase vector	An array consisting of word counts for the provided phrase	N