

EAS 507: Statistical Data Mining II

Project Synopsis

Topic: Recommender System on Nutritional Values of Different Food Items

Anand (anand6@buffalo.edu)

Rahul Patil (rpatil@buffalo.edu)

Project Overview:

COVID-19 is an unforeseen scenario the world is currently facing. From loss of economy to the job crisis, and lack of supply of food products to the markets, we're facing problems on numerous fronts. During such a crisis situation, finding substitutions to certain food items might be a difficult task. In this project, we are building a recommender system based on content-based filtering for recommendation of food products as a substitute for the original product with same or nearly same dietary fulfillments.

A recommender system, is a subclass of information filtering systems that seeks to predict the "rating" or "preference". There are different types of recommender or filtering approaches. For this project, we use content-based filtering approach. Recommender systems are used as product recommenders for variety of services like Amazon, Netflix, etc.

The dataset being used is a "Composition of Raw Foods" dataset which contains data regarding the ingredients in the manufacturing of processed foods. It consists of nutritional data of 8789 food items with details of about 65 nutrients.

The primary objectives of this project are as follows:

- 65 columns of different nutritional data would create a very complex system with less common items, so decreasing the number of nutrients to the most essential ones
- Normalize the data
- Calculate the similarity between items and display the top 10 items based on the primary nutritional elements that can be used as a substitute.

The challenge faced in this project is, there are variants of same items in the data set with different nutritional values, so the recommender built should not result in one of those variants.

We are looking to address this challenge by ranking the results of recommender algorithm and filtering out the items which are not the variant of the same items.

The expectation of this project is a recommender system based on content-based filtering approach, which would output several other food items with the same dietary fulfillments, for ex. If a person is searching for bananas and their primary dietary requirement is potassium, then a kiwi can be substituted, and so for different primary food ingredients.