Georgia Tech

Recipe Recommendation System

Wenxin Fang   Wanjiang Zheng   Xingyu Liu   Yifei Wang   Yujia Xie   Peicheng Hua

Introduction

People always face same problem with food in their daily life: what to eat and how to eat healthy. What's more, people who love cooking sometimes struggle with how to integrate nutrition and taste into one meal. The existing food apps/website recommend recipes based on ratings or flavor preference but less on nutrition.

Our Recipe Recommendation System give users nutritious recipe choices which also satisfy their personal taste.

Approach

Step One: Different ingredient has different compound, and thus they have different flavor. We use this fact to project recipes and nutrition into one space to generate flavor network.
Step Two: Analyze recipes by clustering them using ingredient attribute. We use Silhouette to evaluate the clustering results to test the accuracy.
Step Three: Use Nutritional fact sheet to filter out the recipes to recommend to
Step Four: Developed User Interface to visualize the results.
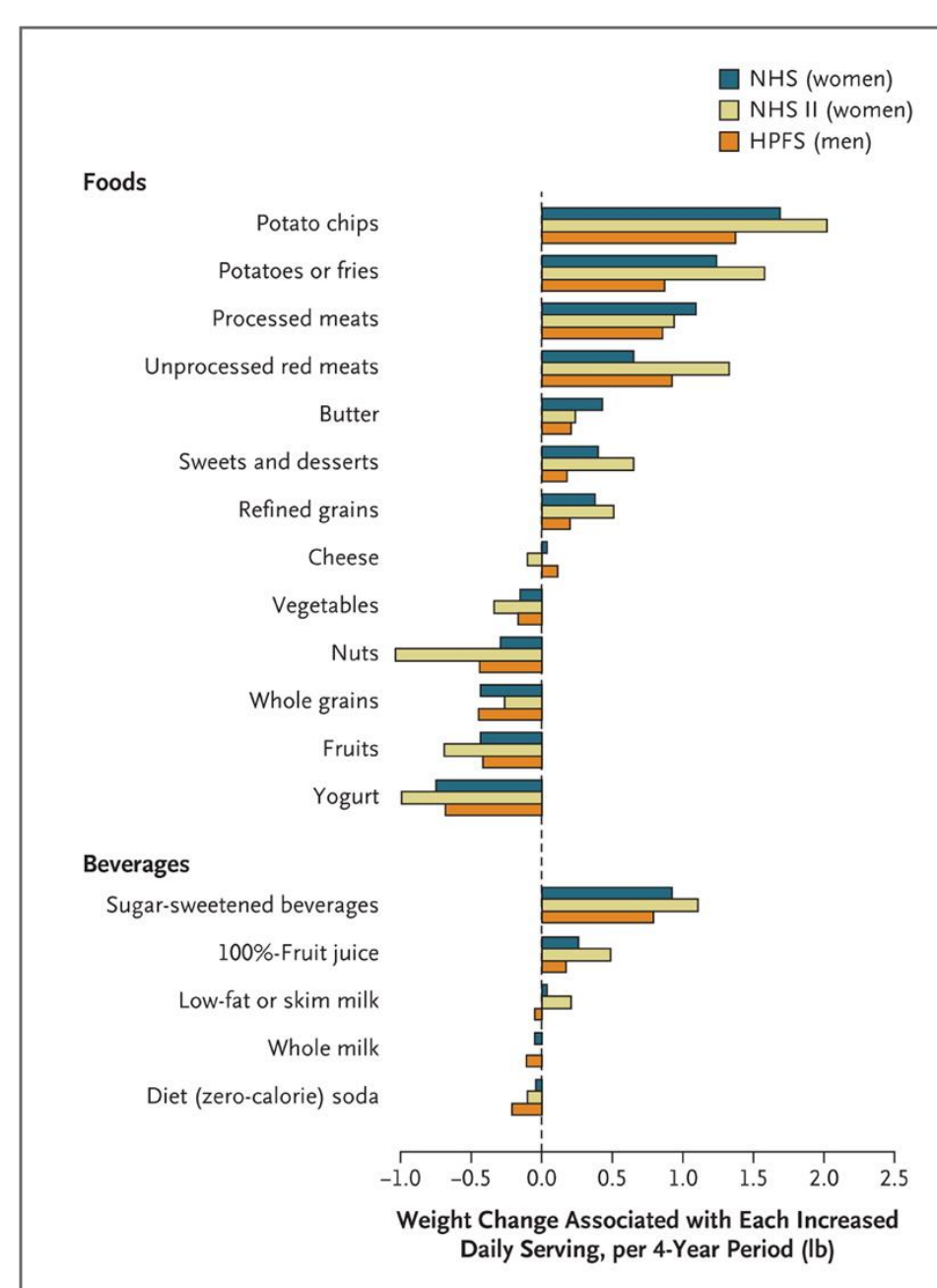
Graph 1.0: visualization of nutrition needs

Data

The recipe table, which is parsed from Yummly API(https://developer.yummly.com/), includes 17 attributes with 100,000 pieces of information. We also crawled images from Yummly webpage.
The food data, including calorie, protein and other 159 nutrition facts, is from OpenFacts (https://world.openfoodfacts.org/) website. This dataset includes 60,000 food records.
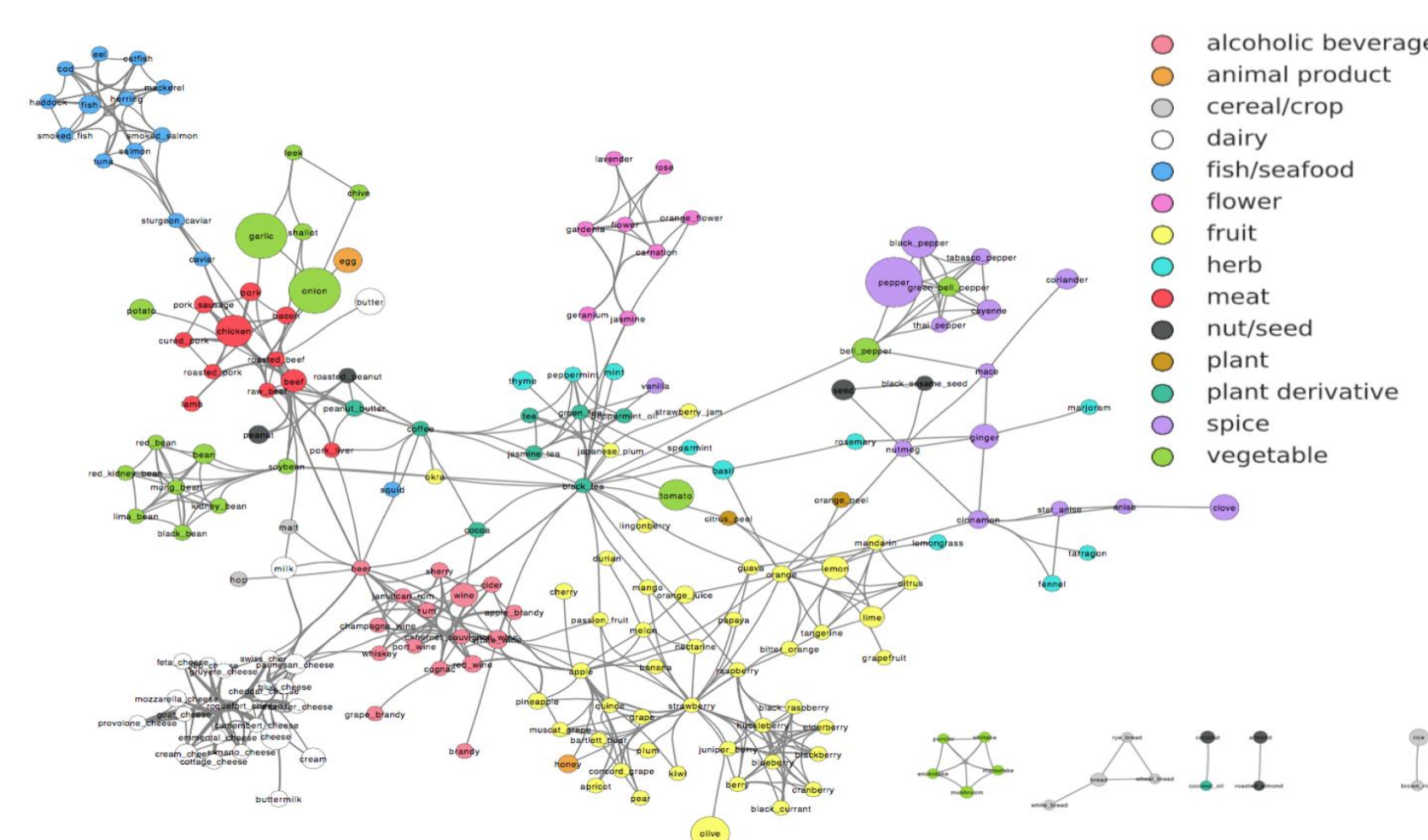Additionally, the flavor network was generated using the nutrition information.
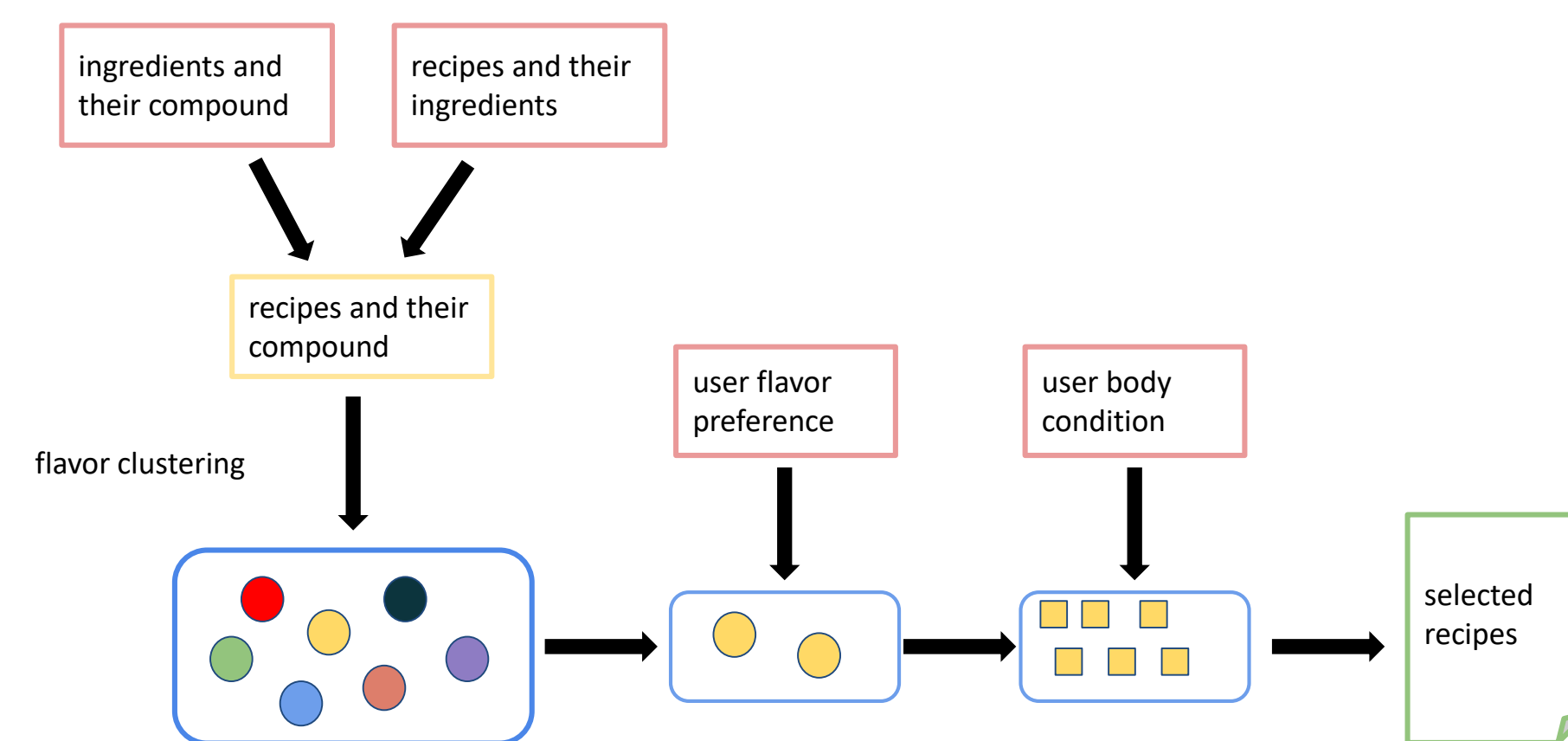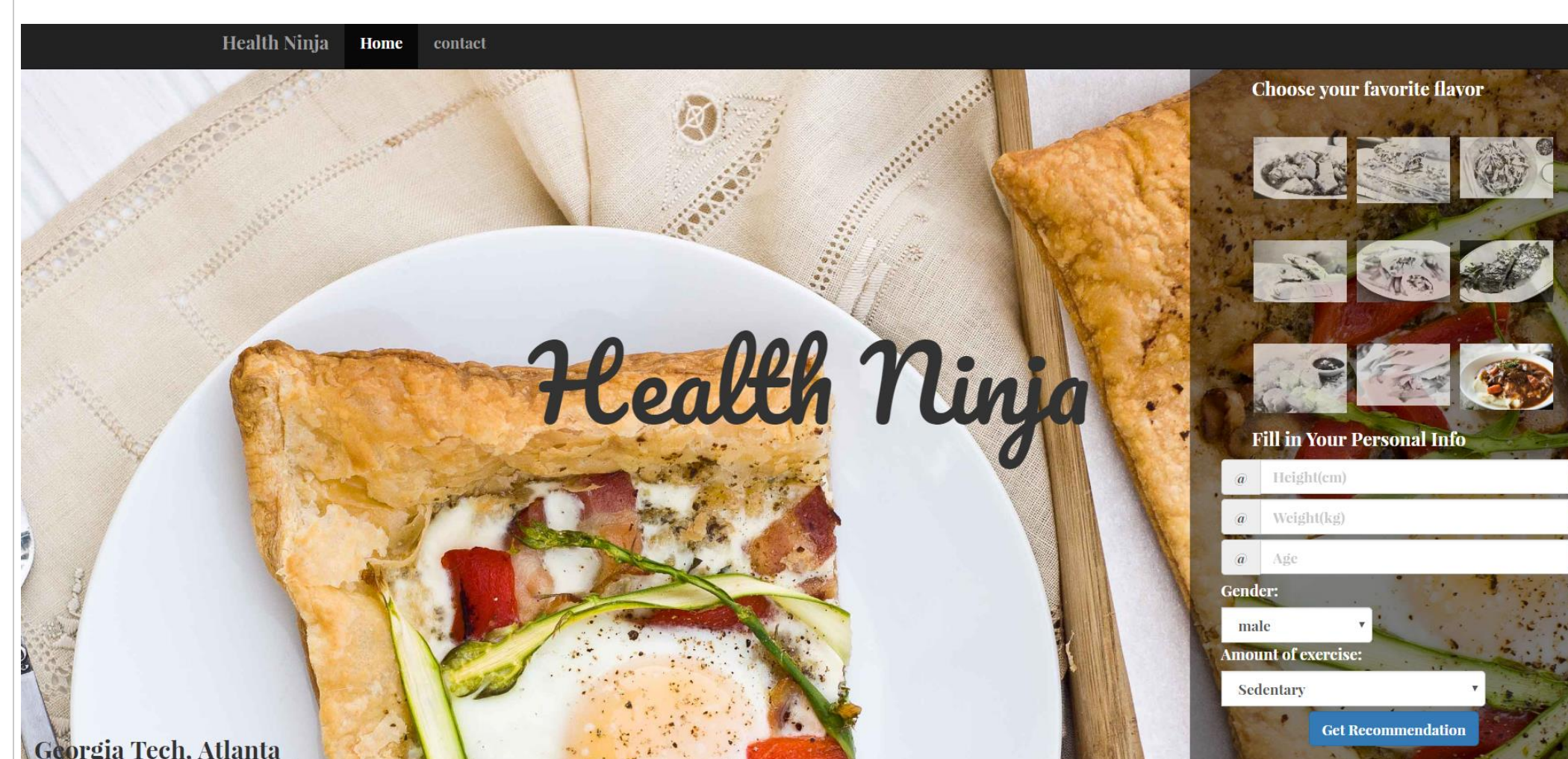

Graph2.0: Project Architect
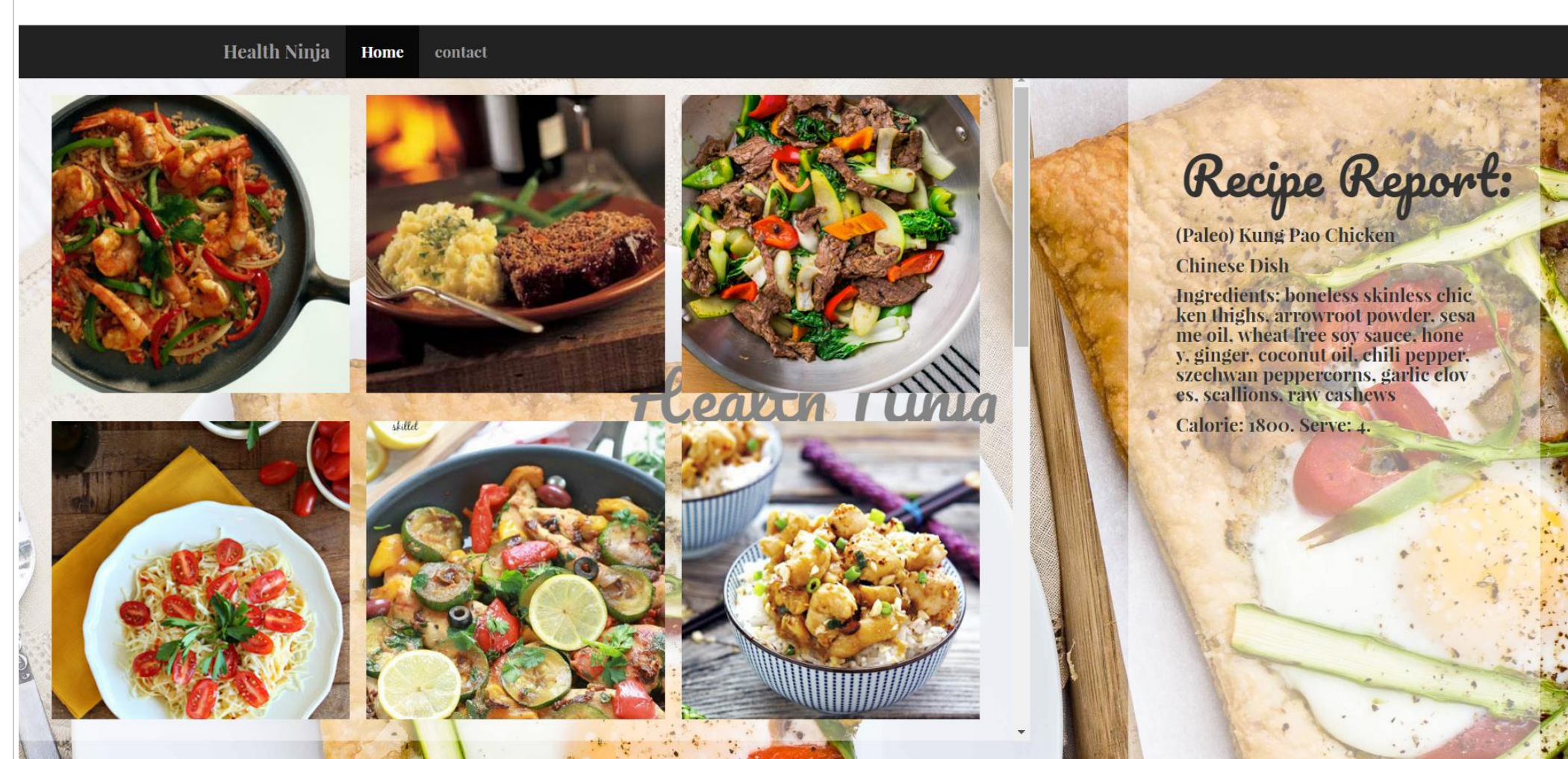
Preliminary Visualization

Flavor Network Projection:

Graph3.0: Flavor Network: including 1530 food ingredients, flavor compounds and 36,781 edges

Website Demonstration

✓ Choose your flavor
✓ Enter your personal information
✓ Start your recipe generation

✓ Clustering results
✓ Gallery of recipes
✓ Recipe details and ingredients

Analysis and Evaluation

How did you evaluate your approaches?
Intuitive approach: Randomly pick 10 recipes with the same flavor group to see if they are the same flavor from human's view. And we use two cluster model (DBSCAN, Hierarchy cluster) to compare flavor groups.
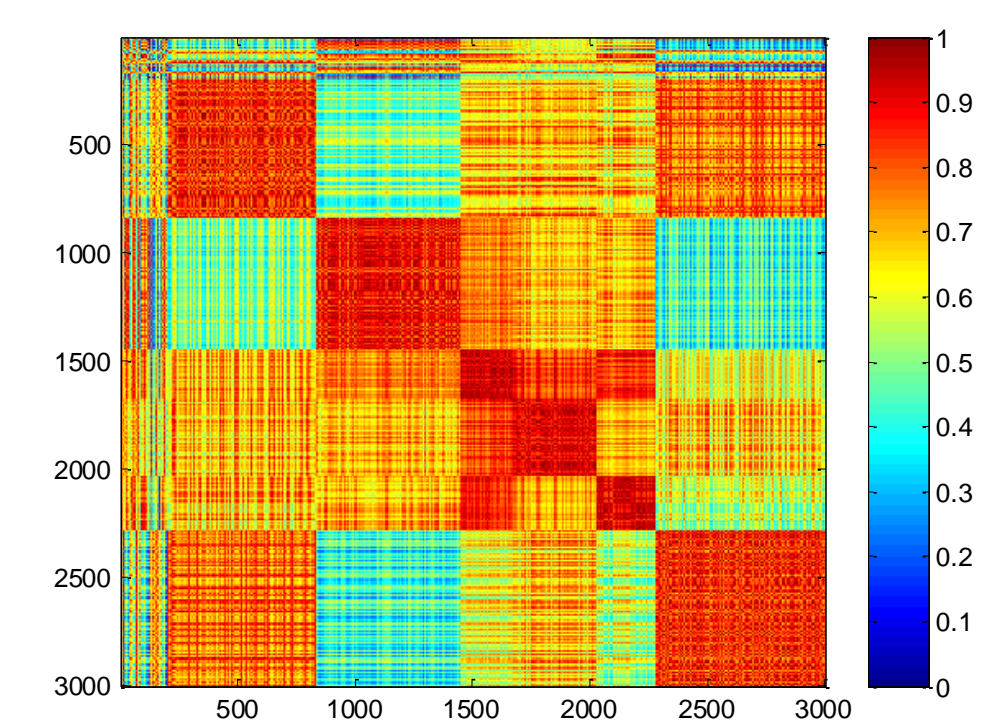
Mathematical approach:
1. Internal measurement *Silhouette*:

$$s(i) = \frac{b(i) - a(i)}{\max\{b(i), a(i)\}}$$

2. Relative comparison: cluster validation via correlation.
3. Similarity matrix.

What are the results?

Graph 4.0: example of similarity matrix

Our result can be divided into two parts. One is that we cluster our collected recipes into flavors based on ingredient compound. In addition to flavor, the second result is that we recommend recipes satisfying personal nutritional need.
Evaluation results: internal index returns result around 0.65, comparing with randomly generated data, it shows some relevancy and consistency among clusters.

How do your methods compare to other methods?
Comparing with other method using users' rating or comment to group recipe, our flavor cluster method group recipes based on compound improve the accuracy.
We however are aware that there are algorithms such as k-means that try to optimize exactly these parameters, and as such we introduce a particular type of bias; essentially this is **prone to over-fitting**.

Victoria Zheng(Victoria.Zheng@gatech.edu), Wenxin Fang(wenxin_fang@gatech.edu)